XSEDE

eXtreme Science and Engineering Discovery Environment

2011-2012 Annual Report

including Q2 2012 Report (April 1, 2012, through June 30, 2012)

XSEDE Senior Management Team

John Towns (NCSA)	PI and Project Director
Jay Boisseau (TACC)	Co-PI and Users Services Director
Ralph Roskies (PSC)	Co-PI and Extended Collaborative Support Service-Projects Director
Nancy Wilkins-Diehr (SDSC)	Co-PI and Extended Collaborative Support Service-Communities Director
Robert Harrison (UT-NICS)	Interim Co-PI
Victor Hazlewood (UT-NICS)	Interim Operations Director
Scott Lathrop (Shodor)	Education and Outreach Director
Kathlyn Boudwin (ORNL/NICS)	Manager of Project Management
Janet Brown (PSC)	Manager of Systems and Software Engineering
Kurt Wallnau (SEI)	Manager of Software Development and Integration
TBD	Chair of the User Advisory Committee
Carol Song (PU)	Chair of the XD Service Providers Forum

XSEDE ANNUAL REPORT

Contents

1 Overview

The Extreme Science and Engineering Discovery Environment (XSEDE) is the most advanced, powerful, and robust collection of integrated digital resources and services in the world. It is an integrated cyberinfrastructure ecosystem with singular interfaces for allocations, support, and other key services that researchers can use to interactively share computing resources, data, and expertise.

1.1 Project Context

Scientists, engineers, social scientists, and humanities experts around the world—many of them at colleges and universities—use advanced digital resources and services every day. Computational technologies and resources such as supercomputers, visualization systems, storage systems and collections of data, software, and networks are critical to the success of those researchers, who use them to advance our understanding of our world, and to make our lives healthier, safer, and better. XSEDE integrates these resources and services, makes them easier to use, and helps more people use them. XSEDE currently supported 16 unique supercomputers and high-end visualization and data analysis resources across the country in Program Year 1, reducing to a total of 13 by the end of the year due to retirement of some resources.

Digital services, meanwhile, provide users with seamless integration to NSF's high-performance computing and data resources. XSEDE's integrated, comprehensive suite of advanced digital services is developing and implementing tools, methods, and policies to federate with other high-end facilities and with campus-based resources, serving as the foundation for a national cyberinfrastructure ecosystem. Common authentication and trust mechanisms, global namespace and filesystems, remote job submission and monitoring, and file transfer services are examples of XSEDE's advanced digital services. XSEDE's distributed systems architecture allows open development for future digital services and enhancements.

XSEDE also provides the expertise to ensure that researchers can effectively use the supercomputers and tools. Those include:

- Extended Collaborative Support that includes teaming with individual research groups or with research communities to extend their capabilities.
- An advanced hardware and software architecture rooted in user requirements and hardened by systems engineering that allows for individualized user experiences, consistent and enduring software interfaces, improved data management, and ways for campus resources to be transparently integrated into the overall XSEDE infrastructure.
- The XSEDE User Portal, a web interface that allows users to monitor and access XSEDE resources, manage jobs on those resources, report issues, and analyze and visualize results.
- Coordinated allocations of NSF's high-end resources and digital services.
- A powerful and extensible network, in which each XSEDE service provider is connected to a Chicago-based hub at 10 gigabits per second and has a second 10 gigabit-per-second connection to another national research and education network.
- Specialized community-provided services that serve a particular function and allow for rapid innovation and experimentation.
- Advanced cybersecurity to ensure that XSEDE resources and services provide confidentiality, integrity and availability of information
- Training, Education, and Outreach efforts that expand the scope and scale of activities to foster greater community participation in XSEDE-based projects through curriculum development, live and web-based training offerings, outreach at professional society meetings, and engagement of under-represented faculty and students.

- Advanced support for novel and innovative projects.
- A fellowship program that brings Campus Champions working closely with Extended Collaborative Support Service staff on user identified challenges for up to a year.
- The Technology Insertion Service, which allows researchers to recommend technologies for inclusion in the XSEDE infrastructure and enables the XSEDE team to evaluate those technologies and incorporate them where appropriate.

1.1.1 Communities Served

The XSEDE user community absorbed significant growth in the program's first year. From 6,056 in Q3 2011, the number of open user accounts grew by 9.6% to 6,636 at the end of Q2 2012. An even larger increase of 14.2% was seen in the number of active user accounts, from 1,965 to 2,245. Gateway users consistently increase the number of active users each quarter by 33% or more, peaking at 1,580 gateway users in Q2 2012. During the year, 3,813 distinct users from 499 different institutions charged compute jobs, and 9,387 distinct user accounts were open at some point in the year.

The level of demand for XSEDE computational resources continued to outstrip the available resources. While total available resources for allocations at XRAC meetings rose by about 33%, due primarily to the arrival of the Gordon system at SDSC, the resources were more than two times over-requested at every meeting except Q4 2012, where resources were only 166% over-requested. The number of active projects hit 1,027 in Q2 2012, breaking the thousand mark for the first time.

Further details can be found Appendix E.

1.1.2 XSEDE's Integrated, Distributed Environment

XSEDE is taking on the difficult but necessary task of documenting a clearly specified architectural design for its distributed systems architecture. Given the nature of the end game of the proposal competition that ultimately resulted in the XSEDE award, the project has had to substantially redesign the architecture originally proposed in order to incorporate innovative and important elements of the previously competing proposal. While this has been difficult and has led to some confusion, the project is making progress in this area and will begin to produce design documents that specify the architecture in detail during the coming months.

1.1.3 Project Governance

The XSEDE project has established an organizational structure and governance that promotes efficient and effective project performance. As this is a distributed project involving 17 partner institutions and with many other stakeholders including NSF, and thousands of users, it was necessary to establish a governance model that balances efficiency and inclusiveness. The XSEDE governance model has strong central management to provide rapid response to issues and opportunities, delegation and decentralization of decision-making authority, openness to genuine stakeholder participation, and improved professional project management practices including formal risk management and change control.

The XSEDE governance model is geared towards inclusion of, and responsiveness to, users, service providers, and the NSF scientific community. The various stakeholders have input through three distinct advisory bodies, which have direct access to the XSEDE Project Director and the XSEDE senior management team through regularly scheduled meetings. In order to remain well informed of the requirements of the user community, XSEDE leadership receives advice and counsel from the User Advisory Committee, the XD Service Providers Forum, the XSEDE Advisory Board, and the TEOS Advisory Committee. These advisory committees are

intimately involved with XSEDE management in guiding the project towards optimal operations, service, and support for users.

The XSEDE project is managed by a senior management team consisting of the PI/Project Director as chair, the co-PIs and key leaders of major areas of the XSEDE project, the Chair of the User Advisory Committee, and the Chair of the XD Service Providers Forum. This team is constituted from those responsible for the day-to-day operation of the project and is the highest-level management body in the organization. In order to be responsive to both the user community and the set of Service Providers with whom we will collaborate, the chairs of the User Advisory Committee and the XD Service Providers Forum are members of this team.

1.2 Project Highlights

As summarized in §2 of this report, XSEDE has supported and enabled an ongoing series of scientific and engineering research and education successes. They are an ongoing testament to the importance of XSEDE to the research community supporting greater productivity, making many advances practical on an advanced time scale.

Of critical importance, we have also established key stakeholder input mechanisms for the project including: the formation of the XSEDE Advisory Board (XAB) consisting of 10 members from the community to participate in this board which will provide guidance to help XSEDE achieve the maximum impact across diverse scientific disciplines and communities; the formation of the User Advisory Committee consisting of members of the national community who represent the needs and requirements of the day-to-day users of the XSEDE environment; the formation of the TEOS Advisory Committee with a focus on the activities of Training, Education, and Outreach Services; and the formation of the XD Service Providers Forum to provide a means by which all Service Providers have input into XSEDE's management and where they can voice any concerns, make recommendations, and provide feedback.

A significant amount of PY1 activity was in initiating many process and procedures and establishing working relationships we will continue to use throughout the lifetime of the project. Throughout PY1 we have successfully put our planning and management processes in place and have been making use of them to improve the effectiveness of the program. Establishing the first technical requirements baseline was an early major accomplishment and facilitated moving forward more rapidly with the architectural design activities. Aided by the appointment of Dave Lifka to the role of Architecture and Design Coordinator, we overcame one of our most significant challenges in working through the philosophical issues remaining from the revision of the project prior to award and the unaddressed architecture issues from that revision. A clear path forward was developed and significant deliverables were planned and accomplished by the end of PY1.

XSEDE has produced a publically available high-level description of the architecture and will be turning out a series of definition documents in PY2. This is a part of a focused effort to provide transparency into the project by which XSEDE has been rolling out a growing set of public documents on the xsede.org website.

XSEDE has also developed and put to practice engineering practices and shepherding capabilities through developing a sequence of processes leading to production roll-out of new capabilities to the community. This was manifest in the delivery of several major new capabilities (GO-Data, EMS, GFFS) from the Software Development and Integration team to Operations for acceptance.

XSEDE continues to provide a high level of support of researchers with our Extended Collaborative Support Services staff engaged in over 147 active projects during PY1 covering a

variety of areas. Of particular note, we had our first example of executing our flexible hires strategy to bring on board expertise to assist prospective users that XSEDE does not already have—in this case in the area of Digital Humanities.

The Campus Champion program continues to be a phenomenal success comprising 118 campuses and 158 Champions at the close of PY1. The group continues to engage in conversations addressing challenges they have found in performing their duties as Champions and in supporting their campus researchers. What is striking is that the Champions have begun to effectively support one another with less reliance on XSEDE staff to answer questions and provide support marking an important transition for this community.

XSEDE's online presence continues to grow and be enhanced. Usage of the XSEDE web site and XSEDE User Portal (XUP) have continued to increase as new technical documentation was published and new capabilities were added to the XUP. The XUP now offers improved allocations management and user profiles capabilities.

Training has been very effective reaching over 2,000 researchers and potential researchers in PY1 via in-person, online, and webcast training workshops. In addition, we collaborated with NCSA's Blue Waters project to develop the Extreme Scaling Workshop held just prior to the XSEDE12 conference.

In our growing efforts in Campus Bridging Team, we engaged in two major activities during PY1: definition of use cases for the Architecture and Design team and a pilot project for testing the Global Federated File System (GFFS) software on campuses. A set of seven use cases for campus bridging activities have been documented covering the use of resources by campus researchers, data movement between researchers and XSEDE resources, and making local resources available to the larger community via XSEDE in exchange for other resources and are being used to drive the architectural design process.

Though this was a common achievement across all of the XSEDE areas, it was perhaps most notable in the areas of XSEDE Operations, Users Services, and Extended Collaborative Support Services, that XSEDE executed the successful and seamless transition of all software, services, support, and connectivity from TeraGrid to XSEDE. In this process, no XSEDE centralized services experienced less than 98.48% availability with most experiencing well above 99% availability. The TeraGrid network was transitioned to XSEDEnet and all Level 1 Service Providers (SP) participated. XSEDE user technical information, the XSEDE web site and XSEDE User Portal, and XSEDE ticket system were all ready to go on Day 1 of the project.

In addition, the XSEDE Operations team deployed vulnerability scanning and federated log analysis services, developed user security training, and implemented new risk-based threat assessment processes. Evaluations of two software offerings that will provide the software layer for the XSEDE Wide File System while improving documentation and service offerings were completed. Key improvements to the accounting and allocations processes were also deployed to drastically reduce the account creation time seen by the research community we support.

In PY1 the XSEDE Operations team initiated the Software Testing and Deployment (ST&D) activities, which did not exist in the TeraGrid. The ST&D group works closely with the Software Development and Integration (SD&I) group, and together with the Architecture & Design (A&D) group represent a mature and robust XSEDE software engineering approach to agile delivery of new software and services. The ST&D group provides a critical readiness check prior to deployment of new software and services within the XSEDE ecosystem and coordinates their deployment across Service Providers and campuses.

The XSEDE User Portal (XUP) has emerged as the preferred way which the research community we are supporting to accesses XSEDE systems, manages allocations, signs up for training,

submits and monitors support tickets, manages files, and updates their profiles. A new XSEDE User News system was deployed with multiple access methods: web, email, and RSS feeds. Many new training courses were created including several focused on the new NSF requirement around data management plans. In addition, training on the use of XSEDE resources was integrated into the training of other, domain-specific cyberinfrastructure projects, with a particular focus this year on the life sciences.

The Education and Outreach team accomplished their year one goals working to engage a larger and more diverse community of practitioners including researchers, educators, and students from 2- and 4-year colleges and universities, Minority Serving Institutions, and EPSCoR institutions. The team continued to leverage external partnerships with the NSF-funded Blue Waters project, NSF-funded EPSCoR projects, and the Partnership for Advanced Computing in Europe (PRACE). A new partnership with the HPC Wales project has begun with an initial emphasis on sharing training materials, developing science gateways, and launching a Campus Champions program.

2 Science and Engineering Highlights

2.1 Astronomy: Measuring Dark Halo Profiles in Galaxies (Karl Gebhardt, The University of Texas at Austin)

Fig. 2.1 *Hubble ACS Image of Elliptical Galaxy M87, one of the specific galaxies where Karl Gebhardt's team weighed the dark matter content. Dark matter exists in the halo of the galaxy (i.e., the outermost regions). This image shows a jet coming from a black hole at the center of the galaxy.*

Through indirect evidence, scientists estimate that dark matter constitutes 83 percent of the matter in the universe and 23 percent of the universe's mass-energy. To detect dark matter, researchers collect data on the motions of stars. This data drives simulations and provides a means of distinguishing the effects of dark matter on a galaxy and its distribution in space. Astronomer Karl Gebhardt used XSEDE systems at the Texas Advanced Computing Center to determine that dark matter is more distributed than predictions previously presumed. Initial results of Gebhardt's research were published in the *Astrophysical Journal* in January 2012. Gebhardt's studies provide information about the fundamental properties of dark matter, which is helping scientists substantiate existing theories and generate new findings about their function in the universe. Next year, the National Science Foundation and academic partners will deploy the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX), the first major experiment to search for the evolution of dark energy, the mysterious force causing the expansion of the universe to speed up over time. Gebhardt is one of the leaders of the project, and his work on XSEDE systems is partly driving the design and implementation of the project. Over three years, HETDEX will collect data on at least 1 million galaxies that are 9 billion to 11 billion light-years away, yielding the largest map of the universe ever produced. The map will allow astronomers to measure how fast the universe has expanded at different times in history. HETDEX will generate about 1 petabyte (1 million gigabytes) of data and require a lot of computer processing cycles. XSEDE will play an integral role in the analysis, storage, and visualization of the project.

2.2 Computer and Information Science and Engineering (CISE): Visualization of Nano-Microscopy (Chandrajit Bajaj, The University of Texas at Austin)

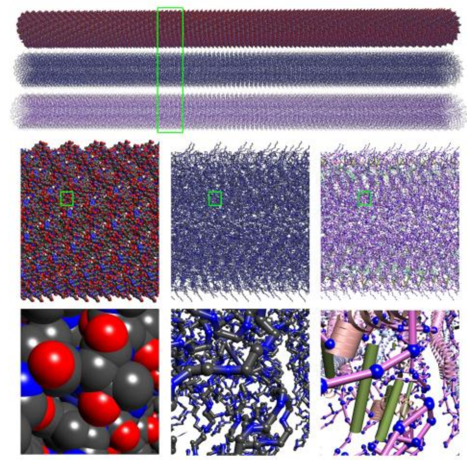


Fig. 2.2 An atomic model of the tobacco mosaic virus (PDBID: 2OM3) visualized at several scales from a full-length rod-like virus (top, tens of millions of atoms) to the atomic level details (bottom).

Computational drug discovery allows researchers to target a small group of possible molecules for therapeutic use, saving significant time and money. The process uses information from electron microscopy and knowledge about drug binding to create 3D models that can simulate the interactions of drug molecules with a target site. In the February 2012 edition of the *Journal of Structural Biology*, Chandrajit Bajaj, et al., reported on advances in image reconstruction using XSEDE systems that allow his group to detect the secondary structures of proteins from single-particle cryo-electron microscopy. This level of detail is required to accurately predict drug binding. To accomplish the image reconstruction, Bajaj uses practically every type of system available in the XSEDE computational ecosystem, including high-performance computing systems (Ranger and Lonestar at the Texas Advanced Computing Center); the world's most powerful remote visualization system (Longhorn); and super-high-resolution tiled displays (Stallion). Together, these resources help Bajaj program, visualize and debug improved image processing algorithms. Over the last decade, computer scientists have also found faster ways to search for things using computers — call it the "Googlization" of research. Bajaj has taken these insights and applied them to the problem of drug target screening. Bajaj and his research team create an ordered list of targets based on the binding energies and biochemical dynamics when two molecules come into contact, which indicates the most compatible compounds. Utilizing both CPU- and GPU-based methods, and mapping different processes to different architectures to accelerate the computation, Bajaj and his team improved the resolution and accuracy of the drug models tremendously while speeding up the docking search by an order of magnitude.

2.3 Finance: Systematic Truncation of Current Financial Data and Its Impact on Research and Policy Making (Mao Ye, University of Illinois at Urbana-Champaign)

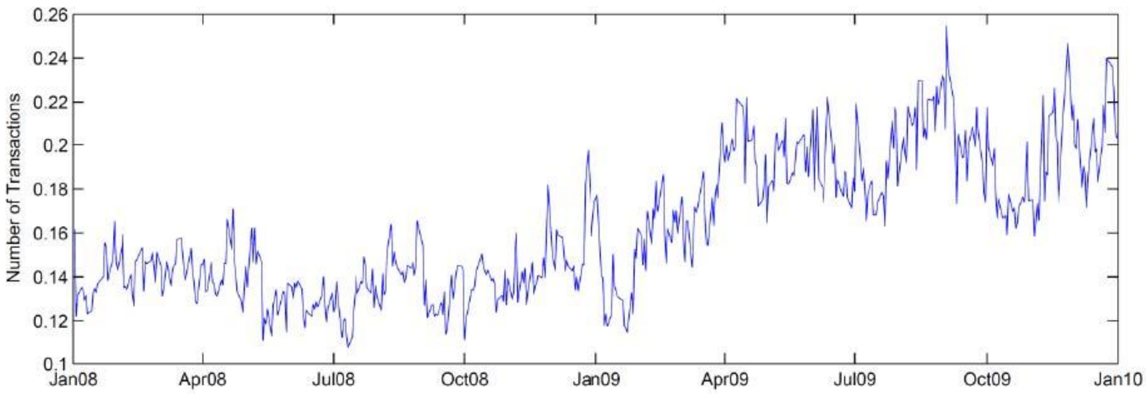


Fig. 2.3 Number of missing trades in the consolidated tape and trade and quote (TAQ) data.

So-called high-frequency trading has become commonplace in the U.S. stock market and has been considered by the press to have contributed to the "flash crash" in May 6, 2010, that sent the Dow Jones Industrial average down 998.5 points in just 20 minutes. The Securities and Exchange Commission recently said it is looking to curb high-frequency traders' huge influence on stock trading, and it is considering charging fees for buy and sell orders that are later canceled, among other options. Mao Ye, an assistant professor of finance at the University of Illinois at Urbana-Champaign, is using Blacklight at the Pittsburgh Supercomputing Center (PSC) and Gordon at the San Diego Supercomputer Center (SDSC) to sift through massive amounts of NASDAQ historical market data to better understand the practices and patterns in high-frequency trading. Ye and co-authors Professor Maureen O'Hara of Cornell and Chen Yao of the University of Illinois prepared a paper, "What's Not There: The Odd-Lot Bias in TAQ Data," that has received a "revise and resubmit" from *The Journal of Finance*, considered the top journal in this field. The study examines how a lack of transparency in odd-lot trades (trades of fewer than 100 shares) could skew perceptions of the market. XSEDE consultant Anirban Jana of PSC worked with Ye to facilitate the use of Blacklight. The study has been featured in news reports, including *Business Week* and *BloombergBusinessWeek*, and has aroused policy debate. According to Ye's research, many of these traders place an order, only to cancel it within 0.001 second or less. "We want to find out if these cancelled orders are being done to manipulate the market in some way," said Ye, whose research also has focused on odd lots, which do not have to be reported to the consolidated tape, or stock price tracker. With computer-based high-frequency trading, the size of trading data has increased exponentially — to the degree that the quantity of data stretches the ability of financial research to arrive at useful conclusions. For example, says Ye, it took several months for the U.S. Securities and Exchange Commission to analyze two hours of recent flash crash data. With Blacklight's shared memory and 500,000 processing hours of time on Gordon, Ye was able to process two years of data containing about 400 million order messages for each day and analyze a large cross-section (7,000 stocks) of high-frequency trading data. He relied on two datasets (NASDAQ Totalview ITCH and NASDAQ high-frequency trading data) that are more comprehensive than the consolidated tape and trade-and-quote (TAQ) data widely used in

academic research. His computations found that the median number of missing trades per stock is 19 percent, while for some stocks, missing trades are as high as 66 percent of total transactions. The study showed that truncation of odd-lot trades challenges the literature using trade size to proxy individual trades and biases measures of individual sentiment. Because odd-lot trades are more likely to arise from high-frequency traders, the study argues that their exclusion from TAQ and the consolidated tape raises important regulatory issues. Motivated in part by this study, the Consolidated Tape Association, a group of stock exchange executives that administers price and quote reporting, has appointed a subcommittee to look at the implications of the truncated odd-lot data.

2.4 Materials Research: Ab initio Modeling of Materials for Energy and Information Technologies (Richard Hennig, Cornell University)

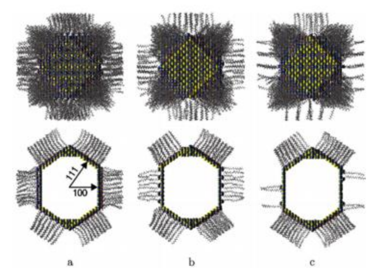


Fig. 2.4 Oleic acid ligands (gray) are shown capping a cuboctahedral lead selenium (PbSe) nanocrystal, illustrating the change in effective shape as the ligand coverage on the {100} facets decreases. The upper panel shows the nanocrystal viewed along the [001] direction, while the lower panel shows a central nanocrystal slice (the interior core atoms are not shown), perpendicular to the [110] direction.

Over the last decade, researchers have created useful nanoparticles for all sorts of industrial applications. However, a comprehensive understanding of the "design principles" behind nanotechnology continues to develop. One of the most promising areas of nanotechnology innovation lies in new energy applications, where nanoparticles can perform nano-processes — such as separating hydrogen from oxygen in water or transforming photons of light into useable power — with far greater efficiency than larger molecules. Using the high-performance computing systems of the National Science Foundation-supported XSEDE initiative, as well as those at the Computation Center for Nanotechnology Innovation at Rensselaer Polytechnic Institute (RPI), Hennig and his research team showed that the concentration and location of small molecules (ligands) on the surface of lead-selenium nanoparticles cause the particles to form different shapes with different energy potentials. The results of the study were published in *ACS Nano* in February 2012. Nanocrystals can form a range of shapes and assemble into different superstructures that are more or less efficient. Hennig's study focused on what controls the shape of the nanocrystals and what controls their assembly. By altering the concentration of ligands present when the nanocrystals form, Hennig and his team produced a range of shapes from octahedrons to cubes with cut-off corners. These results were confirmed by laboratory experiments. The simulations would not have been possible without the diverse and powerful computing systems of XSEDE. To complete the project, Hennig ran density functional calculations on computational resources at the National Center for Supercomputing Applications, the Texas Advanced Computing Center, the Louisiana Optical Network Initiative, and at RPI. By capturing the dynamics of multiple, interacting nanoparticles over time with atomic resolution, the simulations provide additional detail on how nanoparticles behave. After World War I, chemists learned how organic polymers derived from petrochemicals could be transformed through industrial processes to make a broad range of plastic products. Today, materials scientists like Hennig are working on a similar problem (albeit at a scale hundreds of times smaller) with the potential to create products with an even greater benefit to society.

2.5 Materials Science: First Principles Investigation of Materials for Energy Technologies (Gerbrand Ceder, Massachusetts Institute of Technology)

Breakthroughs in energy density are required if the nation hopes to achieve the widespread use of electric cars, or if we aim to stabilize the country's electric power grid using batteries. Researchers from MIT, led by Gerbrand Ceder, used artificial intelligence calculations and high-throughput computational searches to discover a never-before-seen inorganic compound known as lithium pyrophosphate. The compound improves upon current high-density cathode materials used in lithium ion batteries. The group used the Ranger supercomputer, an XSEDE resource managed by the Texas Advanced Computing Center, to perform diffusion calculations for the new material. The simulations provided insights into the pathways by which lithium and other elements cycle through batteries and led the scientists to an understanding of how it can be tuned further to improve its safety and take advantage of the material's unique energy characteristics. After investigating the compound computationally, the researchers synthesized and tested the material in the lab. It produced excellent energy density (the amount of power stored in a given system per unit volume), matching the simulations. The findings were reported in the *Journal of The Electrochemical Society* in March 2012. In the past, it took an average of 18 years from the discovery of a new material to its commercialization. Rapid computational search and exploration are changing that. In the last 18 months, Ceder has patented more insertion cathodes for experimental magnesium batteries than have been invented for lithium ion batteries in the last 25 years. In the past three years, Ceder has used more than 2 million computing hours on Ranger and Pople at the Pittsburgh Supercomputing Center to perform first principles investigation of materials for energy technologies. By helping scientists discover and explore new compounds, the powerful computational resources of XSEDE are playing a key role in the acceleration of materials science.

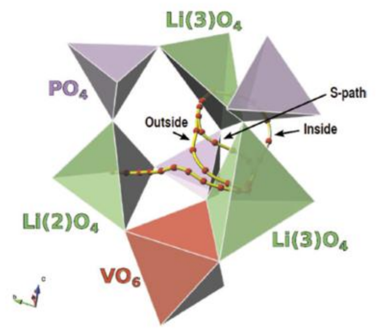


Fig. 2.5 The figure shows the different ways lithium (Li) might diffuse between sites in the structure. The calculations revealed that the rate-limiting step for Li diffusion in this case becomes Li migrating from the Li2 site to Li3 site because Li can easily go between Li3 sites using the path labeled as the 'inside' mechanism. *(This figure was generated by Charles Moore using data from calculations run at TACC.)*

2.6 Mechanics and Materials: An Inverse Stiffness Mapping Approach to the Early Detection of Breast Cancer (Lorraine Olson, Rose-Hulman Institute of Technology)

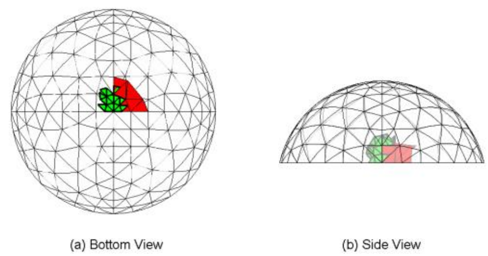


Fig. 2.6 A typical result from the stiffness-mapping algorithm for a simulated breast tumor. The predicted tumor location is shown in red while the green shows the true location. Computed using TACC's Ranger supercomputer.

One in eight women in the United States will develop breast cancer over the course of her lifetime. Lorraine G. Olson, professor of mechanical engineering at the Rose-Hulman Institute of Technology in Terre Haute, Ind., was diagnosed in 2005 at the age of 45. Fortunately, her breast cancer was caught early from a routine mammogram. Few cancer survivors are in the position to change the way cancer is diagnosed, but Olson and her husband Robert Throne, also a cancer survivor, are doing just that for breast cancer. Together, they have created a mathematical model to improve early detection efforts. Cancerous tissues are as much as 10 times stiffer than healthy tissues, but mammograms use X-rays — which are only sensitive to tissue density, not stiffness — for detection. Manual breast exams look at tissue stiffness, but, currently, there is not a good way to record the results of these exams. The ultimate implementation of Olson and Throne's system would use a robotic device to mimic manual breast palpations, enabling doctors to record accurate data about the underlying tissue. Mathematical techniques would then interpret the palpation in terms of stiffness. The device would not replace mammography, but would act as an affordable, effective, and less-invasive complementary tool. To run these simulations quickly and to generate accurate results, Olson and Throne are applying the computational power of XSEDE through resources at the Texas Advanced Computing Center. They also are working with XSEDE staff as part of the Extended Collaborative Support Service (ECSS) in an effort to improve the speed of their algorithms. For breast cancer research, speed of execution is very important, as a typical genetic algorithm must be iterated many times to produce a usable result for a complex problem. Carlos Rosales-Fernandez, an ECSS staff member based at TACC, helped Olson and Throne increase the speed of their code by replacing the solver, optimizing the displacement patterns in 2D, extrapolating those optimizations to their 3D simulations, and exploring alternative fitness functions for the genetic algorithm. With their code optimized for XSEDE's advanced computing resources, the researchers can parallelize their stiffness mapping jobs and finish simulations in minutes. Results of the research were published in the July 2012 edition of *Inverse Problems in Science and Engineering*.

2.7 Vascular Biomechanics: Patient-Specific Modeling of Abdominal Aortic Aneurysms (Endor Finol, Carnegie Mellon University)

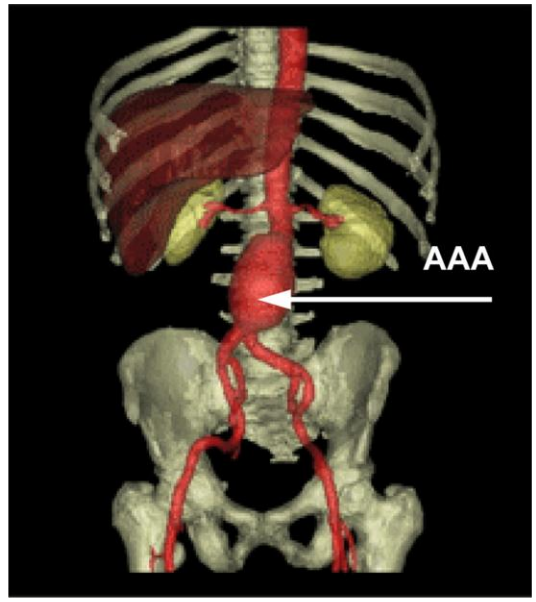


Fig. 2.7 The human aneurysmal abdominal aorta.

Abdominal aortic aneurysm (AAA), an enlargement of the abdominal aorta by 50 percent or more, occurs in more than eight percent of people over age 65. It can lead to fatal rupture and is the 10th leading cause of death for men over 50. Current medical practice lacks the ability to fully assess risk for AAA rupture, with one of the known risk factors being AAA wall stress, for which there are no reliable in-vivo measurement techniques. Many key parameters of AAAs, furthermore, show wide variation among individuals. Endor Finol's work has focused on developing computational protocols (finite-element analysis using ADINA software) for estimating patient-specific AAA features so they can be translated to reliable wall stress predictions and guide decisions about surgical intervention. Finol collaborates with Pittsburgh's Allegheny General Hospital in acquiring MRI imaging of AAA patients. XSEDE consultant Anirban Jana of PSC has provided expert advice on ADINA options and MATLAB coding and in developing a method to implement boundary conditions from patient-specific profiles. With computations on XSEDE's Pople (now decommissioned) and Blacklight, Finol has presented conference papers (with Jana as co-author) on computational solid-stress (CSS) modeling of four patient-specific AAAs. Results show wall stresses more sensitive to changes in AAA shape than material model variations, and the work also suggests that rupture risk may be characterized in relation to AAA morphology. In ongoing work with Blacklight, Finol is increasing the number of patient-specific AAA cases modeled, with the aim of completing analysis from 30 individual AAAs, each of which requires geometry reconstruction and meshing with nearly 3 million degrees of freedom for CSS simulation. Using the shared-memory version of ADINA, Jana has found that the problem (requiring about 13 gigabytes of RAM) optimizes at eight cores with up to 32 cores for faster time to solution.

3 XSEDE Project Office 1.1

3.1 Overview

This first year of XSEDE has seen a incredible amount of work go into establishing the processes and procedures, and practices the project will continue to use throughout the lifetime of the project. Throughout the Program Year 1 (PY1) we have been addressing the fact that at the start of the project some areas were very well defined during the proposal development process (e.g. project management, external relations) but others experienced difficulties in that they were either redefined or only initially defined as part of the final stages of the competition. In the latter cases, significant work has gone into defining the roles and responsibilities along with the associated processes for many practices.

Throughout PY1 we have successfully put our planning and management processes in place and as more of them are available we have increasingly been making use of them to improve how the project is managed and how we execute our work. As the project gained momentum, the management of the project acclimated to a cadence of regular activities such a risk register review and the adoption of the more formal practices in practice. As PY1 proceeded, this was facilitated by the Sciforma project management software tool, thought the procurement of this was delayed due to various bureaucratic processes required to be followed.

Establishing the first requirements baseline was an early major accomplishment and facilitated moving forward more rapidly with the architectural design activities. We have also deployed a web client to provide convenient and cost effective access to the DOORS Requirements database and have made this available to a number of XSEDE staff members and updated the content against the requirements baseline.

This was further significantly aided by appointing Dave Lifka to the role of Architecture and Design Coordinator. Given this was a new area formed as part of the final negotiations of the project plan with NSF, we did not anticipate that Towns (PI) would not have sufficient time to give this area the necessary attention to work through the philosophical issues remaining from the revision of the project and the unaddressed architecture issues from that revision. A clear path forward was developed and significant deliverables were planned and accomplished by the end of Program Year 1. While this was not on a schedule we would have preferred, it was also well recognized that this would be one of the most challenging aspect of starting up the project.

By the end of PY1, we produced a publically available high-level description of the architecture and nearly completed the hard work of defining the details of the architecture with respect to the project's efforts in campus bridging. This activity has helped to refine the processes we will be using going forward as the Architecture and Design Team tackles additional areas which have been defined for PY2.

Similarly, efforts in the Software Development and Integration (SD&I) area have made substantial progress throughout PY1 by developing its engineering practices and shepherding capabilities through developing sequence of processes leading to production roll-out. This was manifest in the delivery of several major new capabilities (GO-Data, EMS, GFFS) from the Software Development and Integration team to Operations for acceptance testing including a feedback loop from Operations resulting in improvements to GO-Data and EMS. To allow them to be more effective in their efforts, SD&I initiated an ongoing open planning activity to identify and prioritize technical enhancements to operational software, and introduced a significantly more agile approach to planning, launching and tracking software development and integration activities within SD&I and on cross-WBS activities with substantial SDI involvement (e.g., Campus Bridging pilots).

In an effort to provide transparency into tthe project, XSEDE has been rolling out a growing set of public documents on the xsede.org website. Many of these come from the XSEDE Project Office and we have made the following documents available:

- Project Definition documents
  - Original Project Summary
- Policy documents
  - XD Service Providers Forum Charter (PDF)
  - SPF Membership Application Form (Word doc)
  - Requesting Membership in the XSEDE Federation as a Service Provider (PDF) (UPDATED JULY 10, 2012)
- Project Reports
  - XSEDE Quarterly Report 2012 Q1 (PDF)
  - XSEDE Quarterly Report 2011 Q4 (PDF)
  - XSEDE Quarterly Report 2011 Q3 (PDF)
- Architecture and Design documents
  - Level 1 and Level 2 Decomposition (PDF)
  - Campus Bridging Use Case Descriptions, Year 1 (Word doc)
  - Campus Bridging Use Case Quality Attributes, Year 1 (Word doc)
  - PY2 Program Plan
- Engineering documents
  - System Requirements Specification (PDF)
- Production Environment and Operations documents
  - Software and Service Baseline (PDF)
  - Technical Security Baseline (PDF)
  - Service Provider and Campus Bridging Checklists (PDF)
  - Software and Services Summary Table (PDF)

These are all on the web site linked from the Project Documents page at: https://www.xsede.org/project-documents.

In a related effort we have been slowly making progress on establishing a digital object repository for XSEDE project objects to allow long-term preservation and sharing of these objects.

As might be expected, the XSEDE Project Office was been the driver for PY2 planning activities throughout the second half of PY1 coordinating the effort from the Project Management Team. This plan was submitted to NSF and reviewed at the first Annual Program Plan Review held at NSF June 13-15, 2012. The plan was subsequently accepted by NSF as out official plan of record.

Another significant step forward has been the formation of the XSEDE Advisory Board (XAB). As a key participant in our planning process, we selected and invited 10 members from the community to participate in this board which will provide guidance to help XSEDE achieve the maximum impact across diverse scientific disciplines and communities. The XAB membership consists of:

- Kevin Franklin, Illinois
- Phil Maechling, SCEC
- Rob Leland, Sandia
- Karin Remington, NIH
- Valerie Taylor, TAMU
- Jerry Ostriker, Princeton
- Theresa Windus, Iowa State
- Tom Lange, Procter & Gamble
- Greg Voth, UChicago

This group was be joined by the Chair of the SP Forum (Carol Song, Purdue), along with two additional SP Forum representatives (Dick Glassbrook, GaTech and Miron Livny, OSG and Univ of Wisconsin), and the Chair of the User Advisory Committee (Tom Cheatham, Utah).

3.2 Project Management and Reporting 1.1.1

The project management team has accomplished the following tasks for the quarter:

Planning:

The Year 2 Program Plan was completed, submitted to NSF, and presented to the peer review panel at the first annual review of the XSEDE project.

Risk Management:

The risk register was updated with current status of identified risks, and was updated with known retired risks.

Project Management Software Tool:

The Sciforma PM software tool completed acceptance testing and underwent customization for the XSEDE project. Initial use of the tool has commenced with the full project schedule being imported from MS Project and other project plans being developed in Sciforma such as the first annual report process. While we had anticipated a rollout for use by the project in late April, this was delayed working through some issues in acceptance and customization. We were able to do some initial hands-on training for managers at the June quarterly management meeting.

Reporting:

The third quarterly report was published and submitted to NSF via Fastlane. The annual report schedule was developed and it was determined that due to the XSEDE12 conference in July that we would delay starting the report until after the conference. NSF concurred and the schedule was agreed, with the annual report submission date target of August 31.

Quarterly Meeting:

The fourth XSEDE Quarterly meeting was planned and held June 5/6 in Knoxville, TN following the Resource Allocation Committee (RAC) meeting. The first day of the meeting was dedicated to project updates by each WBS Level 3 manager, and the second day was focused on preparation for the annual review which was held the following week at NSF. We will follow and co-locate with the RAC meeting whenever feasible to minimum travel expenses and maximize travel efficiency because a subset of the RAC meeting attendees also attend the XSEDE Quarterly meetings.

Annual Review:

The first annual review of the XSEDE project was held at NSF in Arlington, VA the week of June 11. “Lessons Learned” were captured on the XSEDE staff wiki to better prepare for next year’s review, and a set of actions/recommendations from the NSF Review Panel Report will inform adjustments to the XSEDE project going forward. Overall, the review went very smoothly and the review panel had a number of positive outcomes in their report, including a strong statement of the success of the project management activity.

Annual Report:

The PM team accomplished a number of goals in the first year of the project. Initial processes and schedules for reporting and planning were developed, and these were subsequently improved upon with lessons learned after working through these processes. The planning and execution of the quarterly management meetings likewise included enhancements as the year went on based on feedback from the meeting attendees. A significant achievement this first year was the requirements gathering, selection, procurement, and implementation of project management software (Sciforma) to suit the unique and complex needs of XSEDE.

Annual Financial Report

The period of performance for PY1 was July 1, 2011 through June 30, 2012. This financial report also includes the 90-day pre-award spending period of April 1, 2011 through June 30, 2011.

The budgeted amount for PY1 was $25.4M. PY1 expenditures totaled $20.4M plus $0.5M in pre-award spending for a total of $20.9M in expenditures to-date, leaving $4.5M in unspent PY1 funds which are carried forward.

The following tables (Tables 3.2.a and 3.2.b) provide the budget for FTE and non-FTE broken out by WBS Level 2 for PY1 and for PY2. The PY1 budget is $0.5M more than the PY2 budget, because it includes additional funding provided for redefining/replanning the XSEDE architecture as a result of merging the XSEDE and XROAD teams.

Table 3.2.a *- XSEDE Year 1 Budget by Level 2 WBS*

WBS #	WBS Name	FTE	FTE $	Non-FTE $
1.1	Project Office	30.8	$6.2M	$0.2M
1.2	Operations	26.9	$5.0M	$1.3M
1.3	User Services	14.1	$2.8M	$0.2M
1.4	Extended Collaborative Support Service - Projects	18.9	$3.8M	-
1.5	Extended Collaborative Support Service - Communities	17.8	$3.6M	$0.1M
1.6	Education and Outreach	9.5	$1.9M	$0.4M
	Total	118.0	$23.2M	$2.2M
	Head Count	~250		

Table 3.2.b *- XSEDE Year 2 Budget by Level 2 WBS*

WBS #	WBS Name	FTE	FTE $	Non-FTE $
1.1	Project Office	28.5	$5.7M	$0.25M
1.2	Operations	27.25	$5.0M	$0.475M
1.3	User Services	14.4	$2.9M	$0.2M
1.4	Extended Collaborative Support Service -	18.9	$3.8M	-

	Projects			
1.5	Extended Collaborative Support Service - Communities	17.8	$3.6M	$0.4M
1.6	Education and Outreach	11.0	$2.2M	$0.44M
	Total	117.8	$23.16M	$1.74M
	Head Count	~250		

The budget is depicted in Figure 3.2.a as percentage of spent and unspent funds, with the red unspent funds representing the variance from budget.

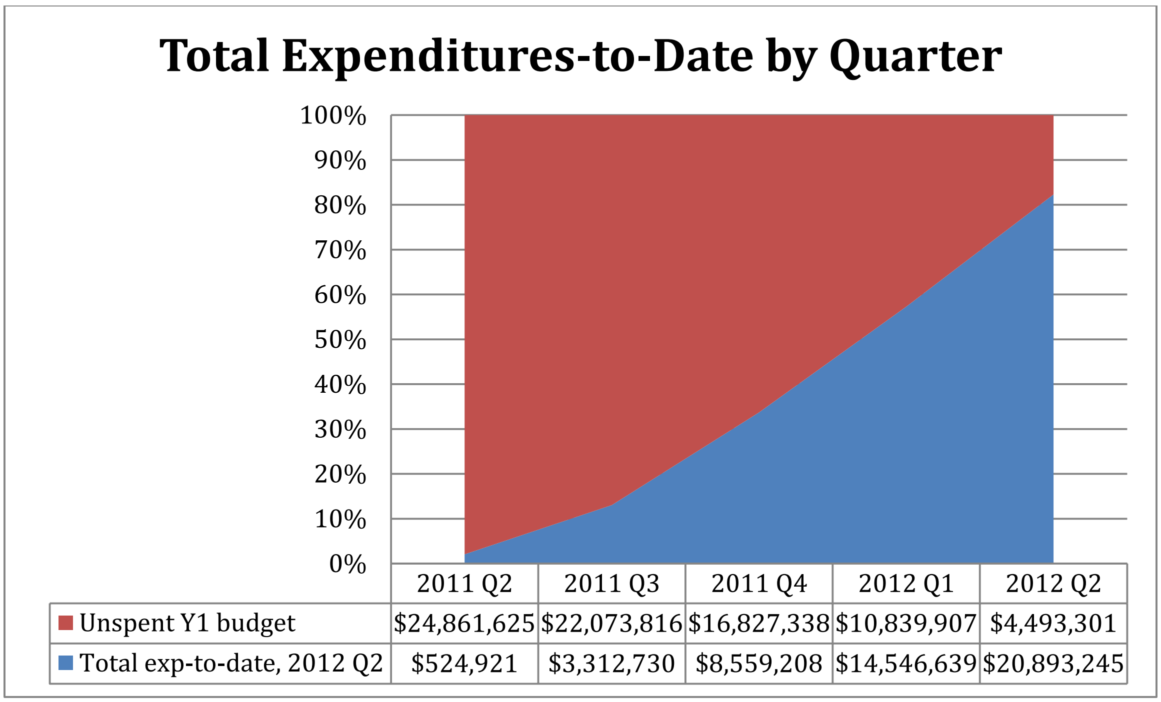


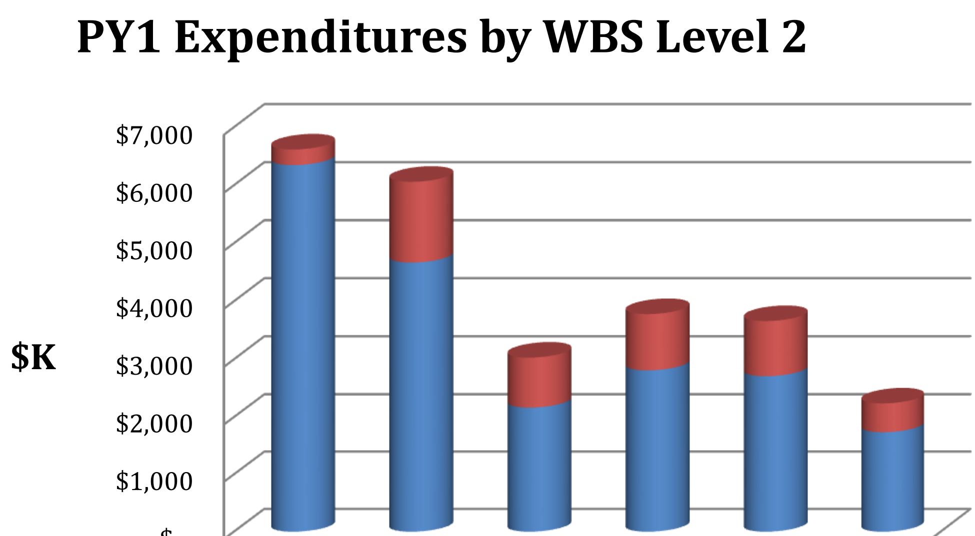
Figure 3.2.a *– Expenditures by quarter, showing the slow ramp. By 2011 Q4 the burn-rate was approaching the budgeted amount.*

An explanation of the budget variance depicted in Figure 3.2.a follows:

1) Slow ramp on FTE expenditures – began to approach budget after the 1st quarter of PY1 (after calendar 2011Q3).
   a) Several TeraGrid partner institutions were spending down the last of their TeraGrid GIG funding. The TeraGrid GIG is the predecessor NSF project that was replaced by XSEDE.
   b) Some partner institutions new to TeraGrid/XSEDE were ramping up their participation/staff.
   c) Some partner institutions are only able to begin charging to an award once their subaward is fully executed. The process for executing a subaward takes 6 to 8 weeks from award notification which was received from NSF on July 21, 2011, resulting in a 3-month delay for some institutions to initiate spending.

2) Capital equipment front-loaded and delayed.
   a) Budgeted $1.3M in PY1 for deployment of hardware, front-loaded and expected to be spent over the project lifetime to support the XSEDE architecture. Due to the delay from redefining the architecture as a result of merging the XSEDE and XROADS teams, only $175K was spent and the balance is carried forward.
3) ECSS subcontracts for specialized support
   a) Budgeted $300K, with less than $50K allocated in PY1.
4) Champions Fellows program.
   a) Budgeted $150K in PY1, however with the timing of starting up the program and selecting Fellows, the first funds were expended in early PY2.
5) Summer 2011 TEOS workshops and programs.
   a) Due to the official award notification being July 21, 2011, most of the summer programs budgeted for in XSEDE had completed by this time and were funded under the TeraGrid GIG.

Figure 3.2.b below shows the PY1 (w/o pre-award) expenditures at WBS Level 2. For the most part, the unspent funds are evenly distributed across the project. The low capital equipment expenses in PY1 described above are all within WBS 1.2 Operations, which explains the greater percentage of unspent funds in that WBS as compared to the others.



	WBS 1.1	WBS 1.2	WBS 1.3	WBS 1.4	WBS 1.5	WBS 1.6
Unspent budget ($K)	$269	$1,392	$864	$970	$951	$504
PY1 Expenditures ($K)	$6,348	$4,668	$2,166	$2,810	$2,709	$1,736

***Figure 3.2.b** – Expenditures by WBS at Level 2, showing the relative amounts of spent/unspent funds. The larger percentage of unspent funds in 1.2 Operations is due to the low capital equipment expenses in PY1.*

The PI and co-PIs can utilize the PY2 program plan development to determine a plan for unspent PY1 funds. The program plan development process captured additional activities and projects that

the Level 3 WBS Managers deemed important to the project, but that did not fit within their PY2 budget, and now some of these can be moved forward.

3.3 Systems and Software Engineering 1.1.2

During the first year of the XSEDE project, the XSEDE Systems and Software Engineering Team (SSE) has accomplished the following:

3.3.1 Requirements Analysis and Processing

Requirements management during the first year of XSEDE focused on merging and cleaning up XSEDE and XROADS requirements and starting work to design an architecture to satisfy them. The original requirements from both teams were reviewed, analyzed and merged into a single set of requirements that formed a base set with which to move forward. The consolidated list of requirements was approved by the Project Office and entered into the DOORS requirements database. Inconsistencies were removed; all requirements were given unique identifiers and now form a cohesive set.

The approved set of requirements forms the **System Requirements Specification** that is available to the project staff on the Systems and Software Engineering section of the XSEDE wiki. Since this is a lengthy document, a simplified, “spreadsheet” version of the Requirements baseline has also been developed and made available to encourage staff to look at the requirements baseline.

3.3.1.1 DOORS Requirements Database

We have installed and successfully tested a web client that provides more convenient and cost effective access to the DOORS Requirements database and have made this available to a limited number of XSEDE staff members. We also worked on developing “read only” access to this system, so that we can provide safe access to the requirements database to a broader group of XSEDE staff members.

3.3.2 Ongoing Collection and Elicitation of New Requirements and Capabilities

The approved baseline XSEDE requirements provide admirable overall goals for XSEDE. However, these requirements were developed during the proposal phase of the project and were based on user surveys and interactions that took place during 2010 or earlier, so may not reflect current user needs. SSE started working with XSEDE staff to help them update and more fully develop their requirements to reflect current activity and trends. This also enabled us to gain better insight into progress made towards satisfying project requirements.

New requirements come from many different sources: internally, from XSEDE staff members; direct requests from users, casual interactions at meetings and conferences, derived requirements, bug-fixes, and most importantly, from the XSEDE groups that involve direct user access: User Engagement, ECSS and TEOS. SSE staff began to work regularly with each of these groups to try to capture new user needs and capabilities from their respective user communities.

Precisely defining a requirement is a difficult task for staff and users alike. SSE also developed more accessible documentation to help both staff and users understand the requirements process and how they can engage with it and use it to address emerging user needs and new technologies.
SSE also reviewed the requirements gathered by the Operations group that were used to guide the new ticket system evaluation and determination, since this will be a valuable tool for collecting new user needs and capabilities.

3.3.1.2 User Requirements, Evaluation and Prioritization Working Group

We formed the User Requirements, Evaluation and Prioritization Working Group - the "UREP". This group is made up of representatives recruited from all of the major project areas of XSEDE and will serve to review and prioritize XSEDE requirements that will guide development of XSEDE going forward.

3.3.3 Documentation

One of the main challenges for the Systems and Software Engineering (S&SE) team has been to get XSEDE staff to understand that system engineering and requirements management processes are in place to make the project better and more effective, not slow things down. Although extensive documentation exists for System Engineering processes, it was written for a different audience and is full of jargon and obtuse language. This documentation is being streamlined and rewritten so that it is understandable and useful to XSEDE staff and community. We are using the XSEDE staff wiki, in addition to conventional methods such as electronic mail, to ensure wide dissemination and availability of information about S&SE processes to project staff members.

3.3.4 XDOR (XSEDE Digital Object Repository)

SSE staff investigated how best to preserve the collected works of XSEDE (key documents, web pages, videos, etc.) for future reference and to comply with the NSF Data Management initiative. We evaluated several different systems (IU Scholarworks, University of Illinois IDEALS, Cornell University Fedora Commons, ACM and IEEE Digital Libraries, OCLC) using a set of required characteristics that any suitable candidates systems must have (ease of use, accessibility, cost effectiveness, ability to handle a wide variety of media and formats, security, longevity, etc.) We selected the UI IDEALS system as the best candidate system because it met all of our needs, offers a natural tie-in with XSEDE because of the connection with UI, and because the IDEALS staff is progressive and interested in working with XSEDE. IDEALS staff have set up a work area for XSEDE and initial testing of this system has gone well. We are developing plans, policies and documentation for preparing and submitting documents into XDOR.

3.3.5 XSEDE Consolidated Glossary of Terms

XSEDE brings together staff from many very specialized areas such as system engineering, cyber-infrastructure, project management, as well as many technical specialties such as system design, operations, security, and networking. However, each of these specialties has its own unique vocabulary, which can lead to confusion for staff members who are not familiar with that specialized area. To try to help staff understand the different terminologies, we developed a consolidated Glossary of Terms to define terms and acronyms in use across the project as an aid to XSEDE staff members. Going forward, this is intended to be a shared document, updateable by anyone on the project who wants to contribute definitions and information from their area.

3.3.6 Risk Review

There are 3 main risks for Systems and Software Engineering: *architecture obsolescence*, *hindered deployment and operation of XSEDE due to architectural ambiguity*, and *lack of qualified system personnel*. We triggered the third risk - lack of qualified system personnel – within the first year of the project when a very experienced SSE left the project. This staff member had been part of project since the project was in the proposal stage and made significant contributions to SSE processes and documentation. An appropriate replacement staff member has been found and is coming up to speed on the project.

3.4 Architecture and Design 1.1.3

The Architecture and Design (A&D) team is coordinated by David Lifka and consists of the following architects:

- Felix Bachmann – Software Engineering Institute, Carnegie Mellon University
- Ian Foster – University of Chicago, Argonne National Laboratory
- Andrew Grimshaw – University of Virginia
- Morris Riedel – Jülich Supercomputing Centre & EMI
- Steve Tuecke – University of Chicago, Argonne National Laboratory

After a challenging startup at the launch of XSEDE the A&D team has spent most of the first year of the XSEDE grant working diligently to produce a detailed set of documents that describe the XSEDE architecture. These documents describe three levels of "decomposition" or details:

Level 1: three layer architectural structure (access, services & resources)

Level 2: defines major functions of each layer

Level 3: details of each function, its components, and how they interact with each other

The XSEDE architecture provides a blueprint documenting the rules of engagement for software developers, users and systems administrators to interact with XSEDE infrastructure in a consistent and secure way that includes quality of service metrics. This is an important differentiator between XSEDE and TeraGrid in the past.

The A&D architectural process is driven by user requirements expressed in terms of use cases and quality attribute scenarios. The use cases express what a user is trying to do. The quality attribute scenarios describe how the success of an architectural response to a use case can be measured. In Q4 the A&D team worked with XSEDE management to identify leaders who have relevant experience in the following architectural areas to produce the necessary use cases and quality attribute scenarios. The areas and associated leaders that have been identified thus far include:

- **Campus Bridging**
  - Rich Knepper
  - Craig Stewart
- **Science Gateways**
  - Nancy Wilkins-Diehr
  - Suresh Marru
- **Computing**
  - **High Performance Computing**
    - Mark Fahey

        - Sergiu Sanielevici
    - **High Throughput Computing**
        - Mike Wilde
        - *TBD*
    - **Scientific Workflows**
        - Ravi Madduri
        - Marlon Pierce
- **BIG Data**
    - **Data Analytics**
        - Nick Nystrom
        - Shawn Strande
    - **Data Movement, Storage, Backup & Archival**
        - Christopher Jordan
        - *TBD*
    - **Visualization**
        - Kelly Gaither
        - Sean Ahern
        - David Bock
- **Connecting Instrumentation**
    - Rick McMullen
    - *TBD*
- **Collaboration**
    - Scott Lathrop
    - *TBD*

Working with these architectural area leads, A&D has been using the following process to provide architectural responses for each architecture area:

- create & revise public facing architecture document with Level 1 & Level 2 decompositions
- work with area leads to create use cases, quality attribute scenarios, and use case/requirement matrix documents
- prepare level 3 based architectural response to use case and quality attribute scenario documents.
- review/approval of level 3 based architectural response with area stakeholders, XSEDE stakeholders and XSEDE leadership
- approved Level 3 documentation integrated into XSEDE public facing architecture document

The approval process of the architecture documents includes review by XSEDE area leads, stakeholders and leadership and involves cross-cutting analyses based on security, integration, interoperability, fault tolerance and interaction with all level 2 service layer components.

3.4.1 Use Case and Quality Attribute Wiki Page

In Q4 Lifka created a use case & quality attribute scenario wiki site. This site contains documentation, related training materials, and access to examples from other architectural areas as well as a place where area leads can contribute individual use cases and corresponding quality

attribute scenarios at any time. Although A&D will maintain its planned "focus schedule" outlined below, this wiki page will enable more efficient and asynchronous development of the XSEDE architecture by making the development components finer grained. This is in turn will support the XSEDE software engineering process by enabling a constant stream of component development from architectural design to Software Development and Integration testing to Security and Systems and Operation production readiness approvals. Development will be tracked by a component registry system currently being planned by Janet Brown of the Systems and Software Engineering team in partnership with A&D. This system will allow XSEDE leadership to track the current state of any SSE component from A&D documentation to production. It will also link requirements in the requirement registry to each component they apply to.

3.4.2 Documentation Produced

- **Public facing** architecture documentation released (see: https://www.xsede.org/project-documents)
  - **Level 1 & Level 2 decomposition**
    - Version 1 *released 2/21/2012, updated quarterly*
  - **Campus Bridging (Craig Stewart & Rich Knepper)**
    - Use Cases version 1 *released 2/21/2012*
    - Use Cases version 2 *released 5/21/2012*
    - Quality Attribute Scenarios version 1 *released 5/21/2012*
- **Internal Documents –** *in development*
  - Level 3 decomposition for Campus Bridging– *in progress, approximately 95% complete.*

3.4.3 Parallel Activities

In addition to documenting the XSEDE architecture, the architects have also been involved in parallel activities that will help them understand the technical concerns and perspectives of the SSE, SD&I, Security, SYS-OPS & leadership teams. Some of these activities include:

- Campus Bridging Pilot
- Production Globus Data Transfer Service/GO-Data and associated security reviews
- Beta Execution Management Service/EMS (Genesis/Unicore)
- Beta Global Federated File System/GFFS (Genesis)
- Cloud Computing Integration Survey

3.4.4 Current Schedule

The current schedule for completing the level 3 decompositions for the following areas is as follows:

- ***Year 1****: (Current year)*

   - ***Campus Bridging** (current - 5/3/2012) In Progress*
   - ***Science Gateways** (5/3/2012 – 8/9/2012) – Use case development started*
- Year 2:
   - **Science Gateways** (7/1/2012 – 8/9/2012)
   - **Computing** (8/9/2012 – 11/15/2012)
      - High Performance
      - High Throughput
      - Scientific Workflows
   - **Big Data** (11/15/2012 – 3/7/2013)
      - Data Analytics
      - Data Movement, Storage, Backup & Archival
      - Visualization
   - **Connecting Instrumentation** (3/7/2013 – 6/13/2013)
- Year 3:
   - **Collaboration** (6/14/2013- 9/19/2013)

Completion of a level 3 decomposition of each area will result in:

1. A use cases document for the architectural area
2. A corresponding quality attribute scenario document for the architectural area use cases.
3. A new release of the Public Facing XSEDE Architecture Document that includes updates to common components and an companion document for the area containing the application of the level 3 decomposition to each use case and corresponding quality attribute scenario.

3.5 External Relations 1.1.4

3.5.1 Accomplishments

ER staff participated in preparations for XSEDE12, as members of Key Leads, Communications, and Logistics and Local Arrangements committees. Communications Committee produced/coordinated: registration page content, website, signs, print and online schedules, awards, viz ballots, badge inserts, master XSEDE12 slides, slideshow promoting XSEDE13, flash drives, bags, lanyards, student volunteers' T-shirts, photography, hiring of sign language interpreters for general sessions, etc., and wrote and edited proceedings front matter for ACM Digital Library, wrote stories about conference events, and managed media placements.

Maintained websites for XSEDE12 conference, Blue Waters-XSEDE workshop, and HPC summer school, and promoted all three extensively to media, XSEDE staff and

users.

Made improvements to XSEDE homepage.

XSEDE's internal and external newsletters continued to be produced and distributed, with content submitted by members of the External Relations team.

Edited and assisted with distribution of announcements including: OSC new Campus Champion, PRACE-XSEDE EoI opportunity, and XSEDE12 speaker and sponsor announcements, non-profit sponsor opportunity, student program.

Tested IDEALS system for possible use as XSEDE document management repository, and created Project Documents page on public website.

Developed XSEDE booth layout for SC12.

Participated on-site in quarterly meeting, annual program review, and annual conference.

Designed and produced certificates for first cohort of XSEDE Scholars.

Created online registration form for quarterly meeting.

3.5.2 Media Hits

XSEDE was included in more than 30 news items on XSEDE-related topics, as tracked during the quarter. An Excel file containing specific media hits has been submitted for inclusion in the Appendix.

3.5.3 Challenges

Continue to be understaffed, but search process is under way.

3.6 Industry Relations 1.1.5

Due to a loss in staffing, little has been done in this area of the project as yet. Some effort has been spent in identifying a new lead for this area. The only deliverable was the ongoing outreach to industry with our training and other offerings. This has been happening to a small extent but has not been well organized. As a result of this situation, during the PY2 planning process we considered eliminating this from the program altogether.

However, at our first XSEDE Advisory Board (XAB) meeting in May 2012, the ideas we had initially developed for this area were presented to our XAB members and we were given very strong advice and encouragement to develop these ideas and move them forward. In addition, they offered to provide additional input in a subsequent teleconference call to help to develop this area of the program.

This has resulted in an initial plan for PY2 to work with the XAB to further develop the ideas and develop a more detailed plan for execution through the remainder of PY2.

3.7 Software Development and Integration 1.1.6

The original mission of Software Development and Integration (SD&I), as defined in the XSEDE Revision Narrative, was to implement the XSEDE architecture. While the Revision Narrative consolidated the best aspects of the original XSEDE and XROADS proposals, SD&I did not exist as a separate function in either original proposal. Its late introduction, while supporting the sound principle of separating design (the purview of A&D) from implementation, was an artifact of the merging of distinctly different engineering approaches. We understood that further refinement of SD&I would be needed after XSEDE was launched; the working out of such details being regarded as a kind of "technical debt" incurred by the last minute merging of XSEDE and XROADS.

Indeed, we discovered in the first quarter that SD&I's original mission, exclusively focused on implementing the XSEDE architecture, was inadequate for several reasons:

- In addition to implementing a new architecture, XSEDE needs to support, and where required to improve, previously deployed software that researchers depend on to do their work. However, there was no "system baseline" to define which previously deployed software capabilities were "in scope" and which could be retired. SD&I's mission needed to be expanded to include corrective maintenance and improvements of deployed software; SD&I and Operations jointly initiated work to document XSEDE's Operational Baseline.
- SD&I was tasked to implement several innovative technologies proposed in the original XSEDE and XROADS proposals—what later became known as GFFS, EMS, and GO-Data. However, implementation was scheduled before the respective requirements specifications and architecture specifications from XSEDE and XROADS were consolidated, which were also regarded as forms of technical debt. Rather than block waiting for the XSEDE Architecture, SD&I and Operations jointly initiated work to consolidate XSEDE and XROADS requirements and to prepare GFFS, EMS, and GO-Data for deployment.
- SD&I staff were drawn from many partner institutions, justifiably without regard for their prior proposal affiliations. Nonetheless, as a consequence, SD&I began work without the benefit of any prior planning among its senior technical contributors, and therefore without the benefit of an operating concept, agreement on engineering tools or engineering processes. Indeed, while the initial SD&I did have technical areas for resource management, data, security, networking, integration, and user information, it did not have a technical area for testing—in hindsight a rather serious, and obvious, omission.

The progress described in SD&I's 4th Quarter Update, below is remarkable in light of SD&I's rocky start and a testament to the ingenuity of the XSEDE community, and of the TeraGrid community from which it sprang.

3.7.1 *Quarterly Planning and Implementation Increments Launched in 4th Quarter*

In the first three quarters, SD&I managed to advance each of its three development efforts (GFFS, EMS, GO-Data) through a rigorous engineering and test process. While none of these efforts survived the Operation's "Acceptance Test" (which itself was initially an SD&I-defined process), this failure was a singular success for XSEDE, as it demonstrated a working principle that XSEDE software must be accepted by customers, not thrown over the fence to them – a first

in this community of practice. Nonetheless, the reasons for rejection, while consequential, were primarily a reflection of the still-missing XSEDE's Security Architecture specification and of the maturity of the new capabilities.

In the 4Q, SD&I initiated its first *quarterly* planning and implementation increment. From around one hundred possible implementation activities identified at the close of 3Q, sixteen activities were scheduled for the 4Q. In addition to addressing operational readiness concerns for GFFS, EMS, and GO-Data, the 4Q increment included activities defining consolidated testing practices and architectures, introducing security mechanisms to allow limited delegation of credentials to improve service interoperability, and improved management of system runtime and configuration information. As of 12 August 2012 significant progress has been made on most of the 16 activities. This is significant considering the increase in development tempo from Q1-3 to Q4.

3.7.2 Engineering Improvements

Q4 witnessed an accelerated adoption, consolidation and coordination of engineering practices across Architecture and Design (A&D), Software Development and Integration (SD&I), Systems and Software Engineering (S&SE), and Operations (OPS). Figure 1 summarizes practices that are routinely used in XSEDE. While SD&I did not develop each of these practices, it provided models from its own practice for OPS and S&SE, and strongly influenced A&D's adoption of view-based documentation and active design reviews (the latter scheduled for pilot use in the first quarter of year 2).

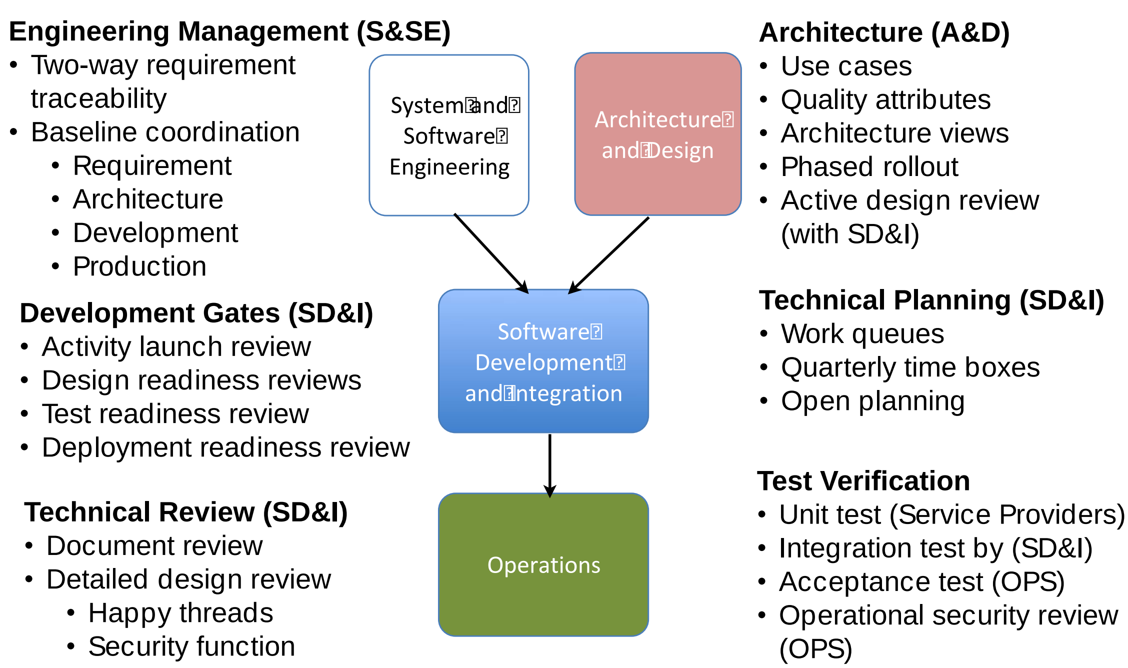


Figure 1: XSEDE Engineering Practices Consolidated in Q4

Because these practices were identified, applied, refined and ultimately fit to XSEDE by the XSEDE community itself, there is a good chance that they will become a tacit part of the XSEDE culture. While the trial-and-error emergence of these practices was often slow and occasionally frustrating, successful engineering adoption in the XSEDE community requires this kind of bottom-up, grass roots approach. The contrasting top-down approach, reflected in the original

XSEDE System Engineering Management Plan (XSEMP), never took root, even though such engineering planning documents are considered normal, if not essential, in other engineering domains. There are important lessons in this for future enhancements in XSEDE's practices.

REFERENCES USED IN THIS UPDATE

1. XSEDE Revision Narrative
2. XSEDE System Engineering Management Plan

4 XSEDE Operations 1.2

4.1 Overview

The Operations group consists of ~30 FTEs and is responsible for implementing, delivering, maintaining, and evolving an integrated cyberinfrastructure capability of unprecedented scale that incorporates a wide range of digital capabilities to support the national scientific and engineering research effort. The Operations group follows the XSEDE project management methodologies detailed in the Project Execution Plan by allocating and coordinating staff in accordance with the XSEDE work breakdown structure (WBS), scheduling tasks in the XSEDE project schedule, and identifying and reviewing risk on an ongoing basis. Operations staff is subdivided into six teams based on the WBS:

1.2.1 Cybersecurity
1.2.2 Data Services
1.2.3 XSEDEnet (Networking)
1.2.4 Software Testing and Deployment
1.2.5 Accounting and Accounts Management
1.2.6 Systems Operational Support

In PY1 the Operations group accomplished many milestones; perhaps most notably, the successful and seamless transition of all software, services, and connectivity from TeraGrid to XSEDE. In this process, no XSEDE centralized services experienced less than 98.48% availability with most experiencing well above 99% availability. The TeraGrid network was transitioned to XSEDEnet and all level 1 Service Providers (SP) participated. The 'xsede.org' Domain Name Service was created and populated, and NLR provided the private XSEDEnet network between the XSEDE SPs with an additional networking path provided to SPs through their campus connectivity or other method.

Other major accomplishments in PY1 include the deployment of vulnerability scanning and federated log analysis services, the development of user security training, and the implementation of new risk-basked threat assessment processes by the Security group. The Data Services group completed evaluation of two software offerings that will provide the software layer for the XSEDE Wide File System while improving documentation and service offerings. The XSEDEnet group succeeded in instrumenting XSEDEnet for monitoring and testing while evaluating next generation networking technologies and providers. The Accounting and Account Management group worked to improve usability and performance of services central to the accounting and allocations processes while drastically reducing the account creation time seen by users.

In PY1 the Operations team initiated the Software Testing and Deployment (ST&D) activities, which did not exist in the TeraGrid. The ST&D group works closely with the Software Development and Integration (SD&I) group, and together with the Architecture Design (AD) group represent a mature and robust XSEDE software engineering approach to agile delivery of new software and services. The ST&D group provides a critical readiness check prior to deployment of new software and services within the XSEDE ecosystem and coordinates their deployment across Service Providers and campuses.

In addition for this quarter, all project schedule items for Operations WBS 1.2 were updated for the baseline project schedule and are reflected in Appendix B. The Operations group completed all tasks scheduled for this quarter. All risk items for the group's activities were reviewed and are listed in Appendix C. No risks were triggered during this quarter, but a failure related to documentation of fail-over procedures for a centralized service has led to the documentation of a new risk. As a result and for broader mitigation, the Operation group is implementing a Critical

Document Bin that will identify operationally critical documents and specify their owner, last review date, review frequency, review process, and review committee.

4.2 Cybersecurity 1.2.1

The Security Operations group's focus is the availability, integrity, and security of XSEDE related information, data, and services. In PY1, the group transitioned and continued day-to-day security operations that include: coordination of cybersecurity across XSEDE service providers, assessment of threats and vulnerabilities, security review for new software and services, and incident response. Also in PY1, the Security group developed security training for XSEDE users, developed a risk-based security review process, evaluated vulnerability assessment technologies, deployed vulnerability scanning, evaluated and deployed federated security log analysis, and defined requirements for and received initial approval for an XSEDE Certificate Authority from TAGPMA.

The following describes the fourth quarter activities for PY1 which are more fully documented on the Cybersecurity page of the XSEDE staff wiki:

XSEDE Certificate Service Requirements - lead Randy Butler (NCSA)
On June 7-8 the Security team met with a team from the Open Science Grid to explore possible Certificate Authority collaboration opportunities. The teams agreed to further explore the potential of each organization providing fail-over services to the other.

Also in the fourth quarter of PY1, the Security group began to define options for an XSEDE wide automated host CA service with the goal of a production service by Q3PY2.

XSEDE OTP Service Coordination - lead Victor Hazlewood (NICS)
An XSEDE one time password (OTP) evaluation was completed in the fourth quarter of PY1 and will be presented at the XSEDE 12 conference. The project implementation plan was developed and is available to XSEDE SPs. A PY2 goal is to implement and test a federated OTP service between NICS and NCSA and to investigate a broader XSEDE OTP service.

XSEDE Coordinated Log Analysis - Lead Jim Eyrich (NCSA)
In the fourth quarter of PY1 the Security group deployed a log analysis solution that leverages the REN-ISAC Security Event Service whereby each of the XSEDE Service providers sends their security incident information to REN-ISAC for analysis and correlation. These logs are then correlated with a larger pool of logs from other REN-ISAC members, and the correlated data (SES data) is fed back to the Service Providers and used to configure filtering and intrusion detection systems to better identify attacks and probes in progress.

XSEDE Security Vulnerability Scanning – Lead Jim Eyrich (NCSA)
Also in the final quarter of PY1, XSEDE began utilizing a Qualys vulnerability scanning service that has already identified numerous security vulnerabilities at Service Providers. To-date five Service Providers have used this service to perform system scans on 525,440 IPs.

XSEDE Federation Risk Assessment - Lead Adam Slagell (NCSA)
This quarter ended the first year of risk-based threat analysis. Documents detailing these efforts and their result are now available on the staff WIKI. PY2 activities will include continued assessment, but focus on amending existing policies and development of needed policies. These activities were delayed as we reprioritized efforts to develop and implement the Security Design Review process detailed above.

XSEDE SD&I Security Review Process – Lead Adam Slagell
Previously the Security group collaborated with the Software Development and Integration group to develop a security review process for configuration items. In this fourth quarter, the new process was utilized to review UNICORE and Genesis II configuration items.

Q4 Response to Vulnerabilities
There were no vulnerabilities this period that were deemed to need an XSEDE wide notification.

Q4 Security Incidents
During this quarter there were no known compromised user accounts and no security issues with XSEDE resources.

4.3 Data Services 1.2.2

Data Services activities in PY1 focused on maintaining the existing operational infrastructure, including GridFTP servers, archive resources, and wide-area file systems, along with migrating and improving documentation for data resources in XSEDE. All Teragrid services and resources were seamlessly transitioned and maintained in PY1. Work also began to add new services to the XSEDE Data Services toolkit, and substantial improvements were made to the documentation of data services and data resources in cooperation with other XSEDE groups.

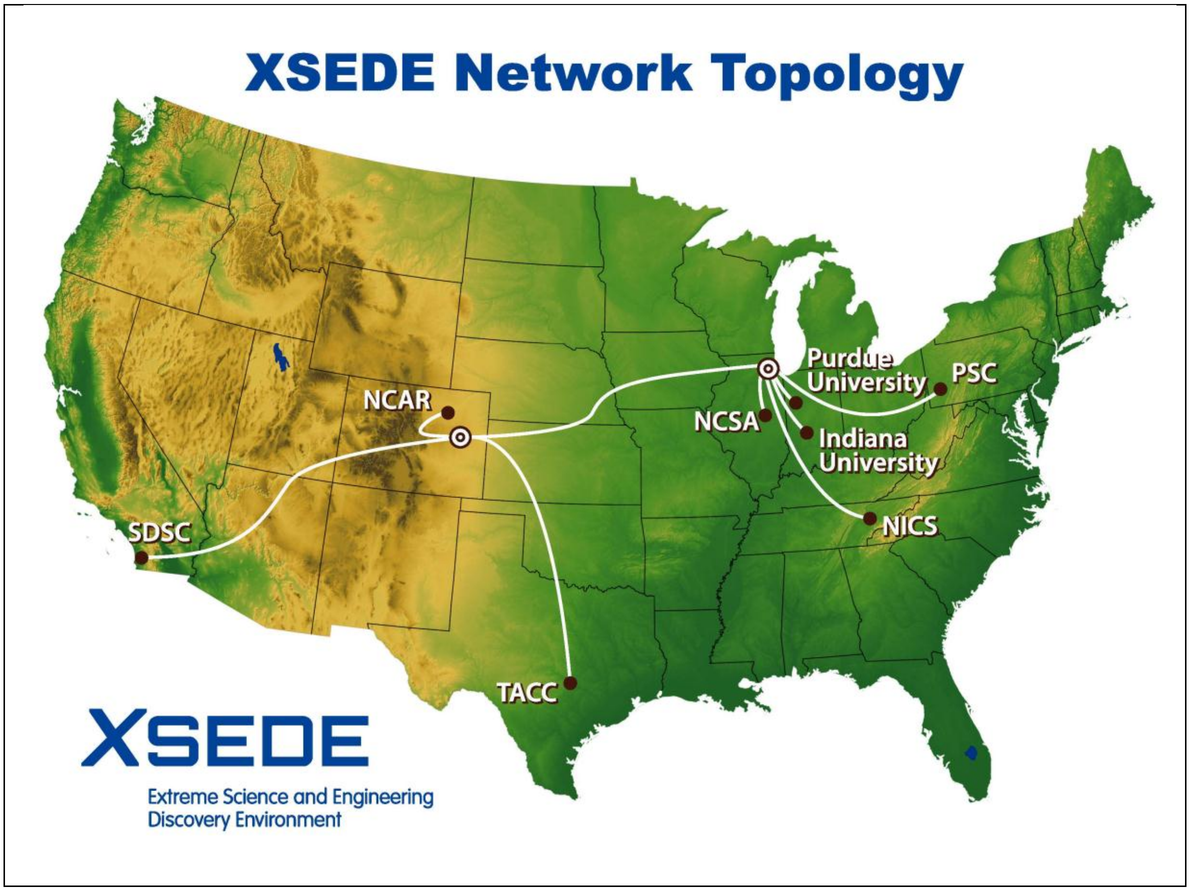
Training activities that focused on Data Services were also initiated in PY1, and three training sessions discussing data management in the context of XSEDE were offered. Data Services staff also assisted in preparing for the deployment of new services for data transfer, supporting SDI testing and planning activities for the deployment of the Globus Online Data and Global Federated File System. Both these services are now deployed in a pre-production capacity on XSEDE resources.

The most significant project for Data Services in PY1 was the examination of potential technologies for the planned XSEDE-Wide File System. A team was assembled that included relevant expertise from both within and outside of XSEDE Operations, an extensive set of requirements was developed, and three candidate technologies were closely examined in light of these requirements.

In this quarter, the XWFS report was completed and a recommendation will be made to project leadership within the Q1PY2. An additional task for this period was to examine replication frameworks and begin development of a strategy for supporting replication between archive resources within XSEDE; this activity has been completed but determination of an archival replication strategy is pending further input from the Architecture Development (AD) and Software Development and Integration (SD&I) teams.

4.4 XSEDEnet 1.2.3

The Networking team's primary focus for PY1 was transitioning from the TeraGrid network infrastructure to XSEDEnet. The updates were coordinated one SP at a time over the period June 28, 2011 – August 11, 2011. Traffic between TeraGrid and XSEDE flowed smoothly during the transition. The XSEDEnet topology for PY1 is shown below.



Operational monitoring of the XSEDE infrastructure required the purchase of a server, and the Intermapper monitoring software was selected and installed at PSC. Working with NLR and the IU GRNOC, XSEDE was able to extract network usage and availability data from the system maintained by IU on behalf of NLR. Because XSEDE sites are making use of NLR circuits for connectivity, it is essential that we work with IU/NLR to extract operational data. Some effort was necessary to interface the two systems. Following configuration and testing, operational monitoring was transitioned to the XOC in Feb 2012.

With the news that the NLR Network Operations Center (NOC) would transition organizations from Indiana University to the Corporation for Education Network Initiatives in California (CENIC) in April 2012, investigation into the impact on XSEDE operations was initiated.

Additional PY1 effort went toward identifying hardware to host perfSONAR services at all sites. Initially, one host was acquired, configured and tested at PSC. Additional perfSONAR hardware was ordered and installed at each SP. These systems were configured to perform throughput tests between XSEDE sites on a regular basis. Discussions with network service vendors were initiated

to determine if the current XSEDEnet topology is the right fit for the future requirements of XSEDE. Contracts with NLR are up for renewal in August 2012 (first quarter PY2).

During the fourth quarter of PY1, connectivity was established between the Galaxy project at Penn State and XSEDE via PSC and Three Rivers Optical Exchange (3ROX).

Traffic utilization of the Chicago-Denver link for 2Q2012 is shown below. Figure 12, Appendix E shows the peak bandwidth across the link for the period. Figure 13, Appendix E shows link utilization as a percentage with March data missing due to the transition of the NLR NOC from IU. Traffic offered load into XSEDEnet links for 2Q2012 is shown in Figure 14, Appendix E.

4.5 Software Support 1.2.4

Operations Software Testing and Deployment (ST&D) is a new activity in XSEDE that did not exist in the TeraGrid. As such, much of the initial effort in this area during PY1 was devoted to developing new processes and document templates to be used for operational reviews of new XSEDE software. The ST&D group's requirements for testing resources needed for acceptance testing were compiled and compared with available systems. Resources for testing were identified within FutureGrid, Tier 1 Service Providers, and campuses.

In PY1 the ST&D group also conducted reviews and acceptance testing for the three major configuration items (CIs) delivered by the XSEDE SD&I group: Globus Online, Execution Management Services (EMS), and Global Federated File System (GFFS).

Globus Online was reviewed and tested during Q3, but unfortunately could not be accepted for production deployment due to security concerns. EMS was reviewed and tested during Q3 and Q4, and was eventually accepted for a limited beta test deployment. GFFS was reviewed and tested during Q4, but the acceptance test report was not completed by the end of PY1. However, it is expected that GFFS will also be accepted for a limited beta test deployment. The proceeding table summarizes PY1 efforts.

	Q1	Q2	Q3	Q4	PY1
CI Readiness Reviews Performed	0	0	2	1	**3**
CI Acceptance Tests Performed	0	0	2	1	**3**
CIs Accepted for Beta	0	0	0	1	**1**
CIs Accepted for Production	0	0	0	0	**0**

4.6 Accounting and Account Management 1.2.5

During PY1, the Accounting and Account Managements (A&AM) group's initial focus was on transitioning the TeraGrid central account management database (TGCDB) and the Partnerships Online Proposal System (POPS) from TeraGrid to XSEDE. This involved numerous changes to web content, URLs, email addresses, and documentation.

The A&AM group also made a number of enhancements to POPS, including changes to simplify proposal entry and enhanced reporting capabilities. In addition, the underlying infrastructure and databases were significantly modified.

The most significant effort of PY1 was directed towards improving account creation times. This was a joint effort between the A&AM group, the User Information and Interfaces (UII) group, and the Allocations group. The existing process for account creates required manual intervention and vetting by XSEDE staff. This meant that account requests could at times take three to four days to be processed--depending on volume and timing. The new process puts the vetting responsibility into the hands of the PI or their proxy. In the first three quarters, the average time for an account request to reach the XSEDE central database (XDCDB) ranged between one and four days. This new process was put in place at the beginning of the fourth quarter and this average time dropped in that quarter to 49.4 minutes. This effort required substantial changes to the POPS infrastructure, and the underlying databases.

A requirements gathering process was also begun for upgrading the XDCDB, the Account Management Information Exchange (AMIE) system, and POPS. All three of these systems are over ten years old. This effort will be directed at addressing performance, scalability and usability. This will be a joint effort between the A&AM group, the UII group and the Allocations group. A significant part of the work in PY2 will be directed towards this effort.

Specific to the fourth quarter, POPS was modified to allow per-allocation-board/per-request-type resource limits. While there were already board level limitations, it is frequently the case that specific resources available within a particular board have lower request limits than the total allowed for that board. These limits are displayed to the person submitting a proposal to provide guidance during proposal submission, but are not actually enforced, as there are frequent exceptions that are allowed.

Also in the final quarter of PY1, a new command-line tool, xdusage, was developed. It is currently being tested, and will be installed on the various XSEDE resources in the first quarter of PY2. This tool will enable real-time command-line monitoring of allocation usage by project members. There had been a TeraGrid tool (tgusage) which provided this capability, but support for that tool ended at the end of TeraGrid. While providing similar functionality, xdusage is a completely new tool, with enhanced functionality.

Finally, during the fourth quarter, work was begun on a set of new AMIE packets to handle duplicate user occurrence within the XDCDB. One aspect of the account creation improvements mentioned above is that the possibility of duplicate user occurrence has increased. A duplicate instance of a user causes problems both for the user, as well as for XSEDE reporting, so it was determined that a programmatic solution for this problem was needed. The detection of duplicates will be a manual process, using a set of queries designed to detect these duplicates, but the subsequent merging will be handled automatically by a set of three new AMIE packets.

4.7 Systems Operational Support 1.2.6

The Systems Operational Support (SysOps) group is responsible for operating the XSEDE Operation Center and providing system administration for all XSEDE centralized services.

Significant progress was made on all WBS tasks in PY1. Accomplishments worth noting include transitioning various systems and services from TeraGrid to XSEDE, which included moving the TeraGrid Operations Center (TOC) to the XSEDE Operations Center and rebranding the TeraGrid ticket system with XSEDE.

Even though there were several planned and unplanned outages during PY1 the SysOps team maintained high overall uptime, which ensured data integrity and availability. By leveraging failover resources, where appropriate, downtime was greatly minimized. As such, no central service experienced any less than 98.48% uptime for PY1. The SysOps team continues to improve operating procedures to ensure the highest level of uptime.

4.7.1 XSEDE Operations Center

During the PY1 the XSEDE Operations Center (XOC) fielded 9,339 tickets and closed 9,036 tickets. Among these 8,652 were submitted via email to help@xsede.org, 64 were submitted via the XSEDE User Portal, and 623 were submitted via phone to the XOC. There were 4,197 tickets closed within 2 business days, which equates to 46% for the reporting period. There were a total of 8,449 tickets responded to within 24 hours, which equates to 90% for the reporting period. Table 16, Appendix E shows the ticket breakdown (opened/closed) for each major resolution center. Figure 16 in Appendix E shows tickets broken into the 7 distinct problem categories with significant representation.

4.7.2 Central Services

There were several outages, both planned and unplanned, that affected various central services during PY1. Many of these outages were the result of individual servers or sites experiencing unexpected technical difficulties or routine maintenance. Outages varied between site-specific power events, networking interruptions, system failures, and planned activities.

The following table summarizes outages by service in PY1 by showing the corresponding downtime/uptime, the nature of the outage (e.g. Planned or Unplanned), and the total number of hours down in PY1:

Service	Percentage of Uptime (Number of downtime hours)	Planned Outage(s)	Unplanned Outage(s)	Total Outage(s)
AMIE	**99.70%**		27 hours	**27 hours**
AMIE Backup	**99.84%**	6 hours	8 hours	**14 hours**
Bugzilla	**99.32%**	4 hours	56 hours	**60 hours**
Build and Test	**99.26%**	4 hours	61 hours	**65 hours**
Certificate Authority Backup	**99.90%**	8.5 hours		**8.5 hours**
Data Movement Service	**99.94%**		5 hours	**5 hours**

Globus Listener	**99.18%**	16 hours	56 hours	**72 hours**
IIS Metrics	**99.26%**	4 hours	61 hours	**65 hours**
Inca	**98.48%**	3.75 hours	129.75 hours	**133.5 hours**
Inca Backup	**99.87%**	11 hours		**11 hours**
Information Services	**99.25%**	4 hours	62 hours	**66 hours**
Karnak	**98.50%**	4 hours	126.5 hours	**130.5 hours**
Kerberos Backup	**99.94%**	4 hours	1.08 hours	**5.08 hours**
MyProxy	**99.90%**	6 hours	2 hours	**8 hours**
Openfire Jabber	**99.86%**	12 hours		**12 hours**
POPS	**99.73%**	24 hours		**24 hours**
RDR	**99.51%**	10.5 hours	32 hours	**42.5 hours**
Secure Wiki	**99.86%**	12 hours		**12 hours**
SELS	**99.86%**	12 hours		**12 hours**
Sharepoint	**99.71%**		25 hours	**25 hours**
Software Distribution	**99.26%**	4 hours	61 hours	**65 hours**
Source Repository	**99.26%**	4 hours	61 hours	**65 hours**
Speedpage	**99.94%**	4 hours	1 hour	**5 hours**
TG Wiki	**99.26%**	4 hours	61 hours	**65 hours**
Usage Reporting Tools	**99.86%**	12 hours		**12 hours**
User Profile Service	**99.30%**		61 hours	**61 hours**
XDCDB	**99.87%**	.25 hours	11.5 hours	**11.75 hours**
XDCDB Backup	**99.91%**	7 hours	.5 hours	**7.5 hours**
Knowledgebase	**99.94%**		5 hours	**5 hours**
Ticket System	**99.73%**	24 hours		**24 hours**
User News	**99.71%**		25 hours	**25 hours**
User Portal	**99.99%**		.92 hours	**.92 hours**
User Portal Backup	**99.89%**	9.5 hours		**9.5 hours**

4.7.3 INCA

In the fourth quarter of PY1, the INCA deployment was executing 983 tests to monitor

XSEDE software and services. Of these, 120 tests were running for six central XSEDE services: Inca, Information Services, Karnak, MyProxy, User Portal, and the XDCDB. The table below shows the definition of an outage for each service and the uptime percentages as detected by INCA for PY1. All services fell within acceptable limits of their high availability service definition.

Service	Definition of outage	PY1 Average Uptime
Inca	Inca status pages are unavailable or not able to fetch data from the database (i.e., test details page fails to load). Tests every 5 mins.	**98.70%**
Information Services	Information Web pages are unavailable. Tests every 15 mins.	**99.96%**
Karnak	Karnak front page fails to load. Tests every 30 mins.	**99.41%**
MyProxy	MyProxy server does not respond to credential query check. Tests every hour.	**99.90%**
User Portal	Portal front home page fails to load correctly. Tests every 30 mins.	**100.00%**
XDCDB	Connection to database refused or slow (using check_postgres.pl script). Tests every 5 mins.	**99.97%**

4.7.4 Syslog Monitoring Project

The Cornell Center for Advanced Computing (CAC) has been working to provide a novel tool for systems administration and monitoring: predictive log analysis and near real-time reactive monitoring. In PY1, CAC began using SQLstream to ingest and store the historical data from system logs of TACC's Ranger. This data was then visualized and analyzed with Ganglia. In PY2, CAC will begin testing this service with additional XSEDE partners.

5 User Services 1.3

5.1 Overview

XSEDE User Services activities continued to increase in quality and quantity throughout Year 1, providing more technical information and user portal capabilities, more comprehensive training, and more user support and engagement to the growing and diversifying XSEDE user community through an evolving allocations process. The User Information & Interfaces area greatly improved the quality and coverage of technical information—systems, policies, etc.—beyond the TeraGrid project, and the XSEDE User Portal (XUP) became a one-stop source of user information and capabilities tool for most XSEDE users. Training reached thousands of users and potential users in Year 1 via in-person, online, and webcast training workshops. User Engagement executed the first XSEDE User Survey, held numerous discussion groups with users, and identified a new ticket system solution to improve future data mining of user issues. In summary, User Services activities operated smoothly throughout the initial project year and the area is poised for addressing major milestones during project Year 2.

Highlights of the XSEDE User Services effort in Year 1 include the following:

- XSEDE User Services staff executed the near-seamless transition of the TeraGrid user community to XSEDE by having XSEDE user technical information, the XSEDE web site and XSEDE User Portal, and XSEDE ticket system ready on Day 1.
- HPC, visualization, and storage system user guides were prepared using an XSEDE-wide common template and placed on the XSEDE web site and portal. This has already drawn much praise, and other institutions have adopted the template. New user guides were developed for additional systems, including most recently Keeneland. Also, the Knowledgebase continued to add articles and update existing articles. There are currently a total of 497 KB articles available to the user community.
- The XSEDE User Portal (XUP) became used by the majority of XSEDE users, who now use it to access XSEDE systems, manage their allocations, sign up for training, submit and monitor support tickets, manage files, and update their user profiles all within the XUP, with more enhancements in progress for rollout in Year 2. POPS integration into the XSEDE User Portal (XUP) was completed, enabling allocations to be fully integrated into XUP and thereby enabling the XUP even for 'pre-users.' Also, a new XUP launch page ('My XSEDE') was created for users who are logged in, providing a comprehensive summary of their XSEDE allocations/usage, tickets, training classes, etc. XUP also updated user profile content.
- A new XSEDE User News system we developed and implemented, with a superior structure to the TeraGrid version and with multiple access methods: web, email, and RSS feeds. 'Events' are also published using the standard iCal format, so consumable by users using Outlook, iCal, Google Calendar, smartphones, etc. (Twitter-based delivery is being considered for Year 2.) Training was fully integrated into the XUP so that users can find, access, and register for training at any location in one place. A single calendar across all sites, centralized registration, and centralized listing of online training offerings are new capabilities in XSEDE, an improvement over the decentralized training information in TeraGrid. This functionality was also used to create a site for SP training coordinators to create classes that get listed in the XUP.
- Many new training courses were created including several focused on the new NSF requirement around data management plans. In addition, training on the use of XSEDE

resources was integrated into the training of other, domain-specific cyberinfrastructure projects, with a particular focus this year on the life sciences.

- Training involving XSEDE resources was integrated into genome assembly workshops of the National Center for Genome Analysis Studies, and into a series of computational biology training events provided by the iPlant Collaborative.
- A new process for addressing user support tickets that were not specific to an SP was implemented to automatically email such tickets to multiple User Services staff to help facilitate an understanding of migration issues, of using resources spanning sites, and of interests in capabilities not yet available in the project.
- User engagement implemented activities to monitor and maintain quality of service and to identify opportunities to improve tools, policies, and procedures to facilitate better tracking and evaluation of the overall consulting process. Personnel were assigned to monitor consulting activities, and quarterly data mining of help tickets was implemented.
- User engagement worked with other XSEDE staff to complete a technical evaluation of ticket systems, and RT was recommended as the system to adopt. The user engagement team submitted a cost-risk-benefits analysis and the recommendation was approved by the XSEDE management team. Configuration and implementation are in progress.
- User engagement conducted the first XSEDE user survey.
- User engagement efforts established regular opportunities for users to contribute feedback by conducting four focus groups, an open user forum, and birds-of-a-feather sessions at the TG11 and SC11 conferences. The early sessions focused on introducing XSEDE to the HPC community while gathering feedback about XSEDE activities carried over from TeraGrid and plans for future XSEDE services, while later focus groups targeted areas of specific interest such as capability computing and big data initiatives.
- XSEDE Resources Allocations Committee (XRAC) meetings and related activities within the reporting period were organized and executed. Meetings were held in September and December of 2011 and March and June of 2012. XSEDE worked with SP Forum and Senior Management team to improve reporting of allocation awards to the PI's program officer.

5.2 Training 1.3.1

The training program had an extremely successful first year, which included achieving all the milestones laid out in the project plan.

Through the past year, more than 70 training events were held, either in person or via webcast. These events reached well over the target of 2,000 attendees. In addition, 36 online modules were either ported to XSEDE from previous Teragrid content, or were created for XSEDE online training. Over 7,000 online visitors used XSEDE online training throughout the year. With a more complete online catalog, this number is expected to be even higher this year.

A number of new courses were created including several focused on the new NSF requirement around data management plans. In addition, training on the use of XSEDE resources was integrated into the training of other, domain-specific cyberinfrastructure projects, with a particular focus this year on the life sciences. Training involving XSEDE resources was integrated into genome assembly workshops of the National Center for Genome Analysis Studies, and into a series of computational biology training events provided by the iPlant Collaborative.

A major achievement of the last year was to improve the user experience for finding, accessing, and registering for training by centralizing support for training XSEDE wide through the user portal. Support for training in the portal includes a single calendar across all sites, centralized registration, and centralized listing of online training offerings.

While this new functionality improves the user experience, it also has major impacts on the project's ability to track the effectiveness of training. The user portal allows the project for the first time to systematically track course attendance across sites, and to tie attendee information to other information XSEDE possesses about the user, such as allocations, support requests, etc. This new information is already being used to track metrics about training effectiveness and user persistence. For instance, since tracking was put in place (after the first project quarter), more than 250 training attendees came back to XSEDE for additional courses. The use of this information to develop progress metrics will continue over the next project year.

5.3 User Information & Interfaces 1.3.2

Both the XSEDE web site and XSEDE User Portal (XUP) released new features and met deliverables for this quarter. The XSEDE web site continued to expand and online content was reorganized by creating new HPC, HTC, Networking and Visualization pages to be consistent with other areas of the site. The UII team also created updated user guide templates and new user guides for Keeneland and Quarry. The web site continued to improve documentation and navigation based on both internal and user feedback including a new and improved XSEDE Home page.

New features in the XSEDE user portal include expanding a user's profile information to accommodate biography, interests, social media information, title and more. Users now have the ability to decide which profile features are to be public or private. The XUP team also improved the usability of the software search and released a new version of the XSEDE OAuth service. The team also released a new GSI-SSH download page to enable users to GSI-SSH from their desktop.

Complementary work included the XUP team release of REST services for user news to allow gateways and other services to access XSEDE User news feeds and deployment of a JAAS Module to enable other XSEDE services (such as projects.xsede.org) to authenticate using XUP credentials to other XSEDE services for single sign on. The team also helped create communities to support education and outreach and campus bridging efforts in XSEDE.

The overall usage of both the web site and user portal continues to increase with almost 2.68 million hits on the web site and over 1.46 million on the user portal. Furthermore, almost four thousand XSEDE users logged in to the user portal and out of 2,295 users running jobs on XSEDE almost 70% of them are logging in to the user portal. Over 21K file transfers occurred within the reporting period, accounting for over 20TB of data transferred via the portal file transfer service.

The Knowledgebase (KB) team expanded and improved the KB articles by adding 19 new KB items and updated existing articles to insure accuracy for a total of 497 documents.

5.4 User Engagement 1.3.3

XSEDE User Engagement is tasked with both proactively and reactively supporting users as they interface with XSEDE. The associated activities of the group are classified as either feedback or consulting activities based on the primary mechanism by which actionable information is obtained. Feedback efforts focus on gathering, recording, and mining information obtained from direct interactions with the scientific and engineering communities, while consulting efforts focus on ensuring seamless user support across XSEDE and inferring information from interactions with users during user support processes. Requirements derived from the information obtained through feedback and consulting activities feed the self-improvement processes within XSEDE.

5.4.1 Feedback

Feedback activities during the first year of the XSEDE project focused on establishing a steady stream of opportunities to directly talk with a variety of XSEDE users about their needs and opinions concerning XSEDE and on developing effective approaches to mine the gathered data for actionable information. The primary approaches employed to collect feedback were focus groups, forums, and the annual user survey, supplemented by individual interviews for clarification when necessary. The user engagement team also created feedback@xsede.org to collect and track feedback from users and XSEDE staff on demand via a special queue in the XSEDE ticket system. Finally, online user forums were established within the XSEDE User Portal in hopes of creating a support community in which XSEDE users could directly assist each other. To simplify the mining of data collected through all of these activities, individual event reports were broken down into specific items of feedback that were then submitted to feedback@xsede.org to produce feedback tickets that were routed to appropriate areas within XSEDE for review.

To establish regular opportunities for users to contribute feedback, the user engagement team conducted four focus groups, an open user forum, and birds-of-a-feather sessions at the TG11 and SC11 conferences. The early sessions focused on introducing XSEDE to the HPC community while gathering feedback about XSEDE activites carried over from TeraGrid and plans for future XSEDE services, while later focus groups targeted areas of specific interest such as capability computing and big data initiatives. All of these events were scheduled at conferences, except for the focus group on big data, which was conducted via webcast. User attendance at general events was low, despite offering incentives and scheduling to minimize conflicts when possible; but, attendance at the targeted events for capability computing and big data initiatives was relatively high, indicating that future efforts should primarily focus on specific topics of significant interest that are delivered online, when possible.

Efforts to establish always available conduits for feedback had mixed results. The creation of feedback@xsede.org was clearly worthwhile based on the amount of easily acquired feedback and on the usefulness of the associated ticket tracking mechanism. However, efforts to establish online user forums as a useful mechanism for discussion and support were unsuccessful due to difficulties in establishing a critical mass of interested and engaged users to populate the forums. Feedback from users and staff concerning the lack of online forum activity provided several ideas for improving participation that are currently under consideration for implementation in the second year of the XSEDE project.

5.4.2 *Consulting*

Initial consulting efforts in the first year of the XSEDE project focused on ensuring a smooth transition from the TeraGrid consulting process to the XSEDE consulting process. Leveraging existing support infrastructure and tools, the user engagement team worked with XSEDE Operations to rapidly and seamlessly transition from TeraGrid consulting services to XSEDE consulting services in the first few days of the XSEDE project. All on-going consulting activities, including open help tickets, were transferred from TeraGrid to XSEDE without issue, and no complications emerged following the transition.

Following the successful transition to XSEDE consulting services, focus shifted to ongoing activities to monitor and maintain quality of service and to indentifying opportunities to improve tools, policies, and procedures to facilitate better tracking and evaluation of the overall consulting process. Personnel were assigned to monitor consulting activities, and quarterly data mining of help tickets was implemented. The XSEDE ticket system was identified for potential upgrade due to difficulties in conducting data mining, inadequacies in tracking performance metrics, and dissatisfaction expressed in feedback provided by user support staff across XSEDE. Additionally, an initial investigation into the capabilities of the XSEDE ticket system and several potential replacement systems identified significant opportunities to enhance consulting services and feedback handling through the upgrade or replacement of the ticket system.

To carefully investigate the benefits and feasibility of upgrading or replacing the XSEDE ticket system, an evaluation team (composed of user engagement staff and representatives from across XSEDE and the service providers) conducted a comprehensive evaluation of the XSEDE ticket system and potential replacement systems. The evaluation criteria were based on a set of derived requirements distilled from general requirements collected from across XSEDE and the SPs. Based on the evaluation, the user engagement team identified significant advantages to replacing the existing ticket system with a federated deployment of RT (Request Tracker) and recommended such following a comprehensive analysis of the associated costs, risks, and benefits. The recommendation was approved for implementation by XSEDE project management, and user engagement began working with XSEDE Operations to configure and deploy the new ticket system along with an associated set of revised policies and procedures that are expected to significantly enhance the tracking and handling of help and feedback tickets.

5.5 Allocations 1.3.4

This objective encompasses the allocations process, both for Startup, Education and Campus Champion allocations as well as the merit-review XRAC Research request process, the POPS system for request handling and management, mechanisms by which allocation PIs manage allocations through transfers, extensions and so on, and interfaces by which allocation PIs manage the users who are authorized to use their allocations. Operationally, this objective includes the XRAC review process, the Startup allocations review and decision process, and the maintenance and operations of the POPS system.

The table below shows the overall allocations management activity handled by POPS and the allocations staff for the reporting period. Note that for Transfers, the table shows only the positive side of the transaction to show the total transfer level; there is a corresponding negative amount, adjusted for resource exchange rates.

POPS Requests and Awards

	Research				Startup				Education				Campus Champions			
	# Req	SUs Req	# Awd	SUs Awd	# Req	SUs Req	# Awd	SUs Awd	# Req	SUs Req	# Awd	SUs Awd	# Req	SUs Req	# Awd	SUs Awd
New	62	162,848,177	52	63,206,036	182	26,989,270	154	18,399,646	12	4,170,014	11	2,770,004	15	11,066,065	14	9,312,009
Prog. Report	2	1,170,050	1	607,050	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Renewal	83	558,945,553	80	315,617,305	24	2,160,612	15	1,324,928	1	200,000	1	200,000	7	4,841,209	7	3,610,017
Advance	37	33,070,136	26	7,419,505	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Justification	0	0	2	24,780,012	0	0	0	0	0	0	0	0	0	0	0	0
Supplemental	22	12,706,499	17	6,470,013	25	7,851,845	21	1,751,845	2	100,000	1	20,000	7	980,000	7	753,000
Transfer	94	36,417,738	81	18,715,901	38	2,311,851	35	1,540,987	0	0	0	0	1	2,600,000	1	190,000
Extension	62	n/a	57	n/a	33	n/a	37	n/a	3	n/a	3	n/a	0	n/a	0	n/a

The June quarterly allocations meeting, XRAC, was planned and held in Knoxville, TN. The next two XRAC meetings have been scheduled for Champaign-Urbana, IL, in September 2012 and Orlando, FL, in December 2012.

Requests totaling 723M SUs were reviewed at the June 2012 XRAC meeting. Reviewer recommendations totaled 504M SUs with 379M SUs available. A large number of moves were needed to bring recommended awards in line with the individual resource available limits. However, the XSEDE Allocations staff and XSEDE site representatives are faced with a difficult situation of continued requests for supplements, transfers, and startups but no significant pool of SUs to satisfy the user community requests!

The XSEDE Allocations staff received 463 tickets within the reporting period, a 44% increase from the previous quarter. Most, if not all, were addressed and a high rate of user satisfaction achieved.

Lastly, XSEDE Allocations staff along with the XSEDE Operations/accounting group, led by Steve Quinn, are working on improving reporting of not only PI awards to their respective program officer but also to the entire XSEDE community via quarterly award announcements in conjunction with the XSEDE ER team.

6 Extended Collaborative Support Service – Projects 1.4

6.1 Overview – ECSS Projects WBS 1.4

The Extended Collaborative Support Service (ECSS) pairs members of the XSEDE user community with expert ECSS staff members for an extended period to work together to solve challenging science and engineering problems through the application of cyberinfrastructure. In depth staff support, lasting weeks to up to a year in length can be requested at any time through the XSEDE allocations process. Expertise is available in a wide range of areas, from performance analysis and petascale optimization to the development of community gateways and work and data flow systems. ECSS staff members also participate in reviewing adaptive proposals associated with XRAC meetings.

We divided ECSS efforts in two, one designated as Projects, headed by Ralph Roskies; the other, designated as Communities, headed by Nancy Wilkins-Diehr. These groups have very close interactions, with common Project Management support. All told, ECSS consists of 37 FTEs, spread over ~80 people at about a dozen sites.

ECSS-Projects consist of ESRT (Extended Support for Research Teams) and NIP (Novel and Innovative Projects).

Primarily, the major successes of ECSS-Projects this past year included (these successes are shared with ECSS Communities):

1) a seamless transition (as far as users are concerned) from TeraGrid, with no projects suffering any serious disruption, although the consultants on certain projects did change.
2) exceeding the quantitative goals we had set for the year
3) effective engagement of new communities that had previously not been users of TeraGrid or XSEDE resources (see NIP report)
4) Much tighter project management
5) ECSS staff members participating in reviewing adaptive proposals associated with XRAC meetings
6) The NIP team often identifies disciplines where the current ECSS team does not have the requisite expertise to assist prospective users. For this, we budgeted for flexible hires for a year at a time. In Q3, we identified such an area (Digital Humanities) and made our first flexible hire.
7) By reshuffling staff assignments at the major XSEDE sites, we were also able to cover emerging areas like genomics without additional hires.
8) Planning a successful workshop with Blue Waters on Extreme Scaling held July 15-16, just before the XSEDE12 conference in Chicago (this will be discussed more extensively in next quarter's report since the workshop took place in that quarter, but the planning occurred previously).

Detailed quantitative accomplishments this past year and quarter are to be found in the metrics section.

We found several tasks more challenging than expected. Primary among them was the difficulty of developing workplans. We have now developed suggestions for ECSS consultants for expediting their formulation. The delay in the choice and then customization of the Project Management software (Sciforma) forced us to operate with less than adequate tools for properly managing the 80+ consultants at about a dozen

different sites. The recent customization should lead to even more effective project management.

ECSS-Projects and ECSS-Communities are very gratified that our efforts were rewarded by the accolades received in the PY1 review held June 13-15 at NSF. These included:

"The Extended Collaborative Support Service (ECSS) has arguably been the most successful component of XSEDE during Year 1."

"It is clear to the panel that ECSS is one of the primary reasons why XSEDE had made such a promising start in supporting the research community through, in many cases, a seamless transition from TeraGrid."

Project Management Activities: Over the past year, ECSS's Project Managers Karla Gendler and Natalie Henriques have been successful in establishing standard operation procedures for handling XSEDE projects that request ECSS support. This has included streamlining the process by which projects are assigned to ECSS staff through L3 managers. They've set up spreadsheets to handle the tracking of projects until formal project management software can be implemented. They have set up and maintained the ECSS wiki and mailing lists, keeping all updated with meeting notes, action items, and staff changes throughout the year. They have coordinated the gathering of quarterly and final reports for each project that is considered active in the last year. They have actively participated on the committee in charge of purchasing the and testing the PPM software with Gendler taking an active leadership role to ensure ECSS needs are met in regards to the PPM software.

Over the last three months, ECSS's Project Managers have continued their tasks of maintaining spreadsheets for project requests, active projects, project assignments and staffing in lieu of formal project management software. They continue to review and track workplans, entering staff allocations on each project and quarterly objectives into the spreadsheets. They manage and attend ECSS meetings and XSEDE PM meetings, posting notes and action items to the ECSS wiki once the meeting has concluded. They provide ECSS information to the XSEDE PM office and relay information from the PM team to ECSS. They also maintain the ECSS wiki and mailing lists. They coordinated the gathering of information for the Q4 report and have published all of the information to the wiki. The purchase of the project management software has been completed and both Gendler and Henriques have taken roles to demonstrate the software at the quarterly meeting to L2 and L3 managers. They are working to transfer the information contained in their spreadsheets and workplans enter in all the information regarding staffing and projects from the last year. Once complete, they will begin to train staff on how to interact with the software. Gendler has become an administrator on the software and is working to customize it to fit the needs of ECSS.

6.2 Extended Support for Research Teams WBS 1.4.1

An ESRT project is a collaborative effort between an XSEDE user group and one or more ECSS staff members, whose goal is to enhance the research group's capability to transform knowledge using XD resources and related technologies. Typical ESRT projects have a duration of several months up to one year and include the optimization and scaling of application codes to use 100,000 nodes or more per job; aggregating petabyte databases from distributed heterogeneous sources and mining them interactively; or helping to discover and adapt the best work and dataflow solution for simulation projects that generate ~100 TB of persistent data per 24-hour run. The first year of the XSEDE ESRT program is also managing projects transitioned from the TeraGrid ASTA program, and all of the ASTA projects will be completed by the end of June 2012.

A request for ESRT support is made by the principal investigator (PI) of a research team via the XSEDE resource allocation process. If the request is recommended by the reviewers and suitable to be an ESRT project, and if staff resources are available, a statement of work for up to one year will be developed by in collaboration by the PI, the ESRT team leader, and the ESRT manager and project manager. The work plan will include staff assignments from the pool of available advanced support experts who have the necessary skills. The ESRT team leader, working with the ECSS project manager, will be responsible for project tracking and reporting and for requesting additional resources or assistance from XSEDE management as needed.

Metrics that quantify ESRT requests and projects for this quarter are provided in Table 1 and Table 2.

Table 1 ESRT project metrics for this quarter

Metric	**XRAC**	**Startups/Edu**
Number of requests	8	12
Number of projects initiated	8	12
Number of work plans completed	1	2
Number of work plans completed for previous quarter(s)	8	3

Table 2 ESRT project breakdown

Metric	**XRAC**	**Startup/Edu**	**TG ASTA**
Number of projects active	29	20	1
Number of projects completed	0	0	15

As of July 2012, there are 12.8 FTEs assigned to ESRT from NCSA, NICS, PSC, SDSC, and TACC.

Table 3 summarizes the number of requests, unjustified/rejected requests, workplans completed (and as such projects in progress), and work plans still in process.

Table 3 ESRT project metrics for the year.

	Requests	Not justified	Workplans completed	In process
Jun-Aug startups	12	8	4	0
Aug XRAC	10	2	8	0
Sep-Dec startups	12	9	3	0
Dec XRAC	11	6	4	1
Jan-Mar startups	11	4	6	1
Mar XRAC	14	6	7	1
Apr-June startups	11	5	2	4
June XRAC	10	2	2	6
Totals	**91**	**42**	**36**	**13**

For the first year, there has been one main challenge with managing the ESRT program - the management of projects and people. ESRT has a large distributed staff and many different projects with different start/end dates. The project management staff is excellent and has helped tremendously, but we eagerly await the project management software to replace email and spreadsheets.

The following sections highlight a few projects that provide examples of the kind of work that is being done in the ESRT program.

6.2.1 Large-Scale High Performance Computation of Aerodynamics, Free-Surface Flow and Fluid-Structure Interation (Bazilevs, UCSD)

ECSS Project Team: Amit Majumdar (SDSC)

This was the first quarter of the ECSS project. The overall objective of the project is to analyze the performance of the parallel code on XSEDE HPC machines by profiling the code and suggesting any modification that can result in improved performance. In the original Fluid Structure Interaction (FSI) parallel code the fluid simulation part is implemented in MPI and the structural simulation part is serial.

During the course of the quarter, PI became more interested in understanding the performance of the serial structural code and this is within the scope of the workplan. ECSS staff worked with a graduate student of the PI, and profiling of the serial structural code was done using gprof on a local cluster (2 quad-core Intel Nehalem 2.4 Ghz processors, with 24 GB of memory); the local cluster was used for ease of access as the initial gprof profiling would not vary much depending on the machine used. Of the many routines used by the code, the top few routines' gprof profiling result is shown below:

```
Flat profile:

Each sample counts as 0.01 seconds.
  %   cumulative   self              self     total
 time   seconds   seconds    calls   s/call   s/call  name
 68.25     44.39    44.39    97363     0.00     0.00  sparseprod_shell_
```

```
 12.09     52.26     7.86         6     1.31     8.71  sparsecg_bdiag_shell_
 11.76     59.91     7.65    149526     0.00     0.00  e3lrhs_klshell_
  1.39     60.81     0.90                              __c_mzero8
  1.36     61.70     0.89    299052     0.00     0.00  dersbasisfuns_shell_
  1.04     62.37     0.67                              pgf90_mm_real8_str1_mxv_
```

Based on these results, and in discussion with the PI, the first quarter's plan was modified a bit and preliminary implementation of OpenMP was taken up as a task for the first quarter. Based on profiling results, the matrix product of the solver routine was implemented using OpenMP and various test cases were run. Test cases were run for both a small problem (with 1,800 elements) and a larger problem (with 18,000 elements) on the Triton machine. Tests were also run on the Trestles machine. The following table shows the results obtained on the Triton machine; the results on the Trestles machine did not show such speedup behavior and that is being investigated currently.

	1 OpenMP Thread	2 OpenMP Thread	4 OpenMP Thread	8 OpenMP Thread
1,800 elements	61.5 sec	40.6 sec	29.4 sec	26.2 sec
18,000 elements	3259 sec	2334 sec	1743 sec	1573 sec

During the the first quarter's work the focus shifted to OpenMP implementation and optimization and this will be noted in an updated workplan that will be sent separately. The FSI code is running on both Trestles and Lonestar and that milestone is also satisfied. In addition, the preliminary OpenMP implementation of the serial structural code, which was not one of the milestones, has also been completed. This task developed naturally after the initial profiling of the serial structural code and further effort in understanding OpenMP performance will continue and will be part of the next quarter's updated workplan.

Publication: Y. Bazilevs, A. L. Marsden, F. Lanza di Scalea, A. Majumdar, M. Tatineni, "Towards a Computational Steering Framework for Large-Scale Composite Structures Based on Continually and Dynamically Injected Sensor Data," Procedia Computer Science, Vol. 9, 2012, Pages 1149-1158.

6.2.2 DNS of Spatially Developing Turbulent Boundary Layers (Ferrante, U. Washington)

ECSS Project Team: David Bock and Darren Adams (NCSA)

The objective of this project is to provide advanced support for the University of Washington DNS team as they implement a 2D decomposed version of the current 1D decomposed MPI code. This includes work to (1) reach Petascale scaling and deployment of DNS spatially developing turbulent boundary layers and droplet-laden homogeneous turbulence on NSF XSEDE supercomputers, (2) do visualization tests per simulation run, (3) do final production quality visualization imagery (for journals, conferences, posters), (4) implement high-level parallel IO library layer built on HDF5, and (5) provide petascale HDF5 parallel IO library (dubbed H5DNS) with functionality for both data creation(via petascale simulation runs) and data analysis.

There has been considerable effort in software engineering. The software started as a single Fortran file--about 20,000 lines of code. Over time the code was split, and refactored to Fortran 90 standard (many memory bugs etc. fixed along the way). Assistance was also provided in making the code portable. This effort ultimately took the form of the cmake build and test implementation they are currently using.

The HDF5 I/O implementation has been specifically developed to fit the characteristics of the existing science code. This software provided methods and a HDF5 file structure that is a self-describing and capable of storing multiple timestep dumps of three-dimensional scalar and vector field data. The current I/O library is written as a simple C library on top of the parallel-HDF5 and MPI (PHDF5) library stack. The library has both C and Fortran interfaces so that it can be used by both the Fortran simulation code and various C or C++ based visualization tools including the VisIt platform developed at LLNL.

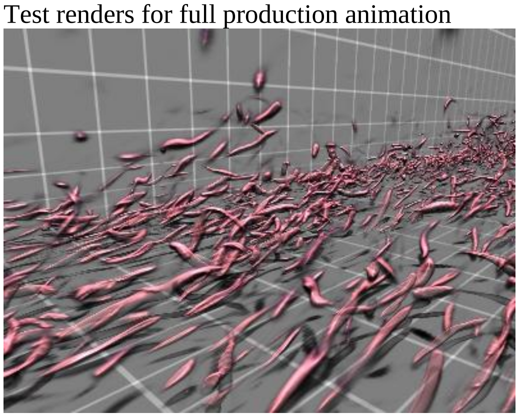
Data provenance is preserved by storing input parameters, software version, build information, and other metadata automatically, for all runs as the attributes "light data" at various levels in the HDF5 file structure. This is a crucial step to both preserving provenance information of the scientific data as well as providing needed metadata about the storage layout, timestep parameters etc. to construct useful visualizations.

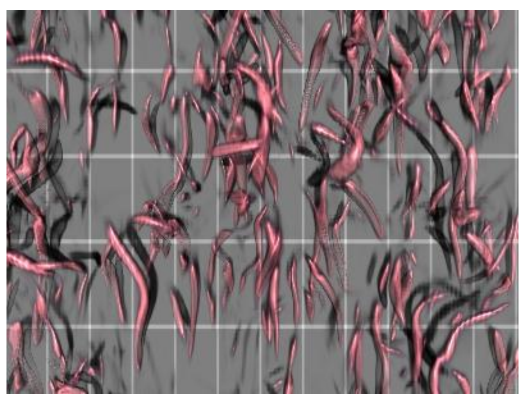
For the large "heavy data", some data-reordering and buffering is performed to store data in a standard visualization-ready data set. The end result is reasonably good parallel IO performance that is portable to all existing XSEDE resources and usable by C and C++ Visualization software as well as the Fortran-based simulation code. The I/O library routines and HDF5 file format layout performed well enough that they have also been used to carry out the large field restart dumps.

On the visualization component, accomplishments to date have been to generate visualization tests to verify simulation and HDF5 DNS reader by rendering images and test movie sequences to test most recent simulation run. In addition, full production visualizations work has begun with scene construction and tests for full production animation run using latest test simulation data.

The visualization work is journaled at http://lantern.ncsa.uiuc.edu/~dbock/Vis/Ferrante/.

Figures:
Test renders for full production animation





6.2.3 Modeling ship propellers (Chryssostomidis, MIT)

ECSS Consultant: Vincent C Betro, NICS

In order to properly model propeller "crashback" physics (wherein a propeller quickly reverses direction while still moving ahead) within a reasonable time constraint, a highly optimized hybrid OpenMP/MPI code is sought. The modeling is based on high-order spectral/hp element solver NEKTAR. NEKTAR was developed at Brown University and has been successfully employed to simulate unsteady 3D flow in complex geometry and at high Reynolds numbers. The conjugate gradient loop and many of its sub-functions have already been parallelized, and many loops have been replaced by calls to functions by each thread. These functions work on different chunks of data, wherein 16 byte memory alignment is preserved for each chunk and SSE code is generated. This has resulted in cutting the CPU time per time step roughly in half; by applying these parallelization techniques to more candidate sections of the code, the intention is to cut the CPU time in half again.

With ECSS' Vince Betro's effort being between 15-20% and an equal commitment from MIT's Leopold Grinberg, the work to parallelize pre-determined candidate sections of Nektar using OpenMP and direct functional decomposition along with the original MPI is nearly done.

In preparing a paper for SC12, which unfortunately was not in the 20% of papers accepted, several tests were run in timing the effects of different numbers of MPI ranks and OpenMP threads per rank as well as differing modifications to loop fusion/pragma use in the conjugate gradient solver. The best performance was seen at 2 threads per rank, and the best performing "hybridized" routines were the DAXPY, DSPMV, and parallel gather sections of the CG solver.

Runs have been done at several sizes, including 7,400; 14,400; and 28,800 cores, and we have used three different physical problem setups with polynomial interpolants of order 1-8 and differing initial conditions to test Nektar rigorously.

With this background work having been completed, it is hoped that the remainder of the project will go smoothly and entail testing at scale and working out any I/O issues.

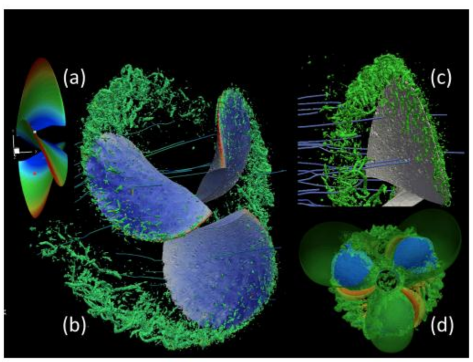


Figure 1: HPC in ship design: simulations of a flow around a realistic 4119 propeller. Geometry of the propeller blades (a) is represented with Immersed Boundary Method. λ_2 iso-surfaces (b,c) show the turbulent flow generated due to the fast rotating blades. (d) iso-surfaces of pressure close to the propeller. Visualization: J. A. Insley (ANL).

6.2.4 New Scoring, Assembly and Evaluation Techniques for Protein Structure (Xu, U. Missouri)

ECSS Consultant: Bhanu Rekepalli, NICS

This project wanted help parallelizing next-generation sequencing analyses software' and protein structure prediction (MUFOLD). The project team needs support to run them on Kraken or Nautilus in the parallel computing environment and the resulting code would be more efficient for generation of large-scale results.

MUFOLD is currently targeted toward a single-user desktop environment. The software operates by executing a workflow of internal and external tools. Some of these tools rely on the output of one another. MUFOLD uses the system() function to execute the external tools. Then, the output is generally moved, copied, or transformed to another location – for the next tool to use

There were several path/filesystem issues that needed correcting, essentially since it wasn't written to be portable to other systems, basically symlinking and aliasing, and sometimes changing hard-coded locations for e.g. BLOSUM*, hmm, nr databases, etc. Also, changing the makefile for Kraken and figuring out what modules were needed had to be done. The list of missing files was extensive, e.g. the cheng/betacon files, all of the neural network weights files, db files, their hmm database, and the still-missing CallMuAligner.pl. The person of contact from University of Missouri is taking quite a bit of time to send the required databases and missing scripts after repeated requests, which is slowing down the progress. The code uses perl scripts, and since perl is not supported on Kraken compute nodes, we had to compile a statically-linked version of perl.

The code uses many shell commands (cp, rm, find, etc.) that also don't work on Kraken, so I added a copy Busybox that had been compiled for Kraken, and added symlinks for all of those commands. There was also some error-handling added, as when one step fails and generates an empty output, another step does an fgets on that empty output, but doesn't check for feof, and instead loops endlessly on the empty fgets string, waiting for it to change to some value. There are many examples of this; we only fixed the ones that were infinite-looping in our tests. There were other small things, e.g. the arguments in the instruction's example don't match the way the code is written, that had to be figured out. The progress is slow but we are fixing errors as we go and waiting for the missing databases and scripts that did not come bundled with the release which was provided by the PI's students. We are still working on milestone 2, we are nearing to running a test run with MUFOLD in coming weeks as we are close to final steps of porting the code to Kraken.

6.3 Novel and Innovative Projects 1.4.2

The mission of the Novel and Innovative Projects (NIP) team is to provide proactive, sustained efforts to jump-start XSEDE projects by non-traditional (to HPC/CI) users. Activities range from initial contact to the development and execution of successful projects, including those that receive extended collaborative support. The scope of NIP includes disciplines whose practitioners have rarely availed themselves of HPC/CI resources in the past. It also includes demographic diversity, such as researchers and educators based at MSIs and EPSCoR institutions, and SBIR recipients. Bringing these communities to XSEDE leads to the consideration of applications and programming modes that have not been the focus of HPC in the past, such as those necessary for data analytics and informatics, and of innovative technologies such as streaming from instruments, mobile clients, and the integration and mining of distributed, heterogeneous databases. The implementation of campus bridging processes and technologies will be particularly important for these communities.

Biweekly teleconferences and the use of the project wiki and email list have been successful in catalyzing communications among team members, who benefit from each other's contacts and expertise and share best practices.

In the first year of the program (July 2011-June 2012), the NIP team (5.3 FTEs) has initiated and executed 16 outreach events; engaged 38 groups of potential XSEDE users; contacted 47 XSEDE user groups; led 9 ECS project planning efforts; and is currently involved in the technical execution of 9 ECS projects.

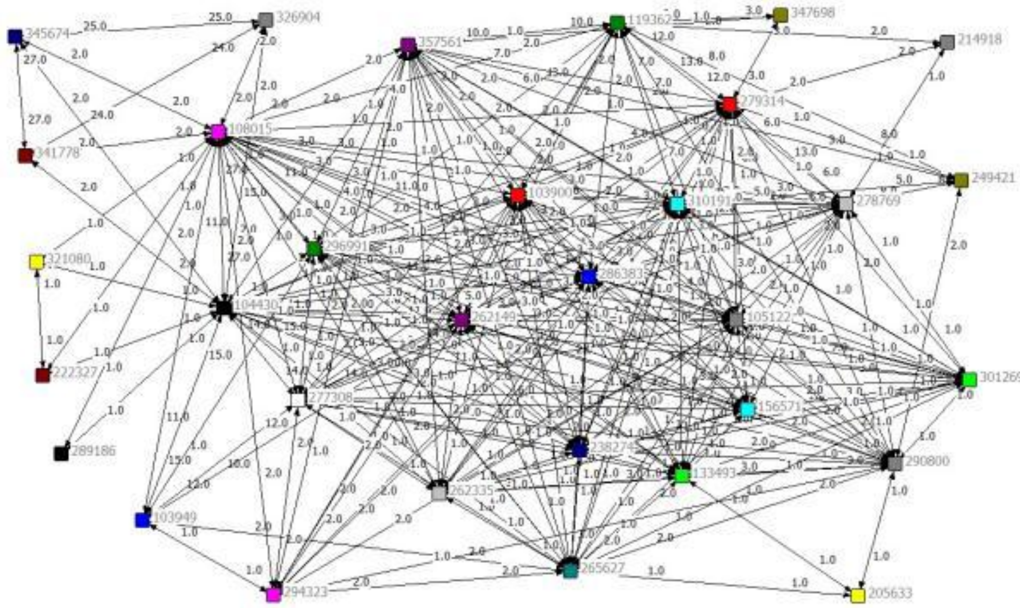
Our main challenge is to balance our ability to *attract* novel projects with our ability to help them *succeed* in becoming efficient users of XSEDE resources. One aspect of this is matching the expertise of ECSS staff to the characteristic requirements of NIP, which often differ from those of traditional HPC applications. To do this, we completed the first ECSS flexible contract hiring process by recruiting Alan Craig of I-CHASS to become our Digital Humanities expert. To bolster our effort level in the short term, and to grow digital humanities expertise in the longer term, we secured the participation of one of the first four ECSS Campus Champion Fellows, Dirk Colbry of Michigan State University, in the ECS project *"Interactive Large Scale Media Analytics"* (PI Virginia Kuhn, USC). For the ECS project *"Data Mining and Network Analysis for Library Collections"* (PI Harriett Green, University of Illinois), we have enlisted the support of the Nautilus (RDAV) project team at NICS.

Another strategy is to partner with organizations and programs that develop and provide tools and services already widely used by the target communities, or with a mandate to do so. We have had considerable success doing this in the Genomics community, where our partnership with the Broad Institute focused on their community codes *Trinity* and *Allpaths-LG* has generated 26 new Startup allocations on the Blacklight system at PSC, between December 15, 2011 and June 30, 2012 alone. We then developed a *Community Capabilities* ECS project to optimize these codes for suitable systems in XSEDE; we are leveraging the *Trinity* effort undertaken by the National Center for Genome Analysis Support (http://ncgas.org/) and their partners. We have also partnered with the *Galaxy* team at Penn State and Emory (see http://galaxy.psu.edu/) for the Science Gateway project "*Galaxy Bioinformatics Platform on XSEDE"* to enable jobs submitted by end-users to the Galaxy main site to execute on XSEDE resources.

With our domain expert Alan Craig, we are now pursuing similar partnerships in the humanities, arts and social sciences (HASS). Discussions are underway with the Humanities, Arts, Science, and Technology Advanced Collaboratory (HASTAC, http://hastac.org/) as well as with the Hathi Trust Digital Library (http://www.hathitrust.org/htrc).

Several of the projects fostered by NIP in our first year are beginning to generate results and publications. For example, the XSEDE 12 proceedings paper *Comparing the Performance of Group Detection Algorithm in Serial and Parallel Processing Environments* discusses the progress and findings of the group led by PI Marshall Poole (U. Illinois) supported by NIP team member Dora Cai (a co-author of the paper). The paper points out that developing an algorithm for group identification from a collection of individuals without grouping data has been getting significant attention because of the need for increased understanding of groups and teams in online environments. This study used space, time, task, and players' virtual behavioral indicators from a game database to develop an algorithm to detect groups over time. The group detection algorithm was primarily developed for a serial processing environment and later then modified to allow for parallel processing on *Gordon.* For a collection of data representing 192 days of game play (approximately 140 gigabytes of log data), the computation required 266 minutes for the major steps of the analysis when running on a single processor. The same computation required 25 minutes when running on *Gordon* with 16 processors. The provision of massive compute nodes and the rich shared memory environment on *Gordon* has improved the performance of this analysis by a factor of 11. Besides demonstrating the possibility to save time and effort, the study

also highlights some lessons learned for transforming a serial detection algorithm to parallel environments. The group network detected by their algorithm is visualized below.



Dynamic social network in a massive multiplayer online game.

The XSEDE 12 science track session in which this paper was presented also featured presentations resulting from the NIP mentored projects *High performance mining of social media data* (PI Judith Gelernter, CMU) and *Interactive Large Scale Media Analytics"* (PI Virginia Kuhn, USC, with NIP team members Ritu Arora and Alan Craig as co-authors and co-presenters).

7 Extended Collaborative Support Service – Communities 1.5

7.1 Overview

Requests continue to come in through the XSEDE allocations process for Extended Support for Community Codes (ESCC) and Extend Support for Science Gateways (ESSGW). 18 ESCC projects and 15 ESSGW projects are currently in progress. Work in both of these areas is also internally initiated by XSEDE. In ESCC, these include projects to optimize commonly used community codes. This work is reviewed and approved by the newly formed XSEDE User Advisory Committee. In ESSGW, internally initiated work includes development of science gateway use cases for the XSEDE architecture team.

ECSS staff demonstrate their leadership through Extended Support for Training, Education and Outreach (ESTEO) activities. They present at prestigious conferences (domain science, computer science and computational science). They review for prestigious journals and they deliver tutorials at high profile events. This reporting period included deadlines for submission to both SC12 and XSEDE12.

The ECSS Symposium continues each month and is open to the public to highlight work going on in ECSS projects and allow ECSS staff to learn from one another. There were six speakers this reporting period.

- Scaling the Knowledge Base for the Never-Ending Language Learner (NELL): A Step Toward Large-scale Computing for Automated Learning. Presenter: Joel Welling (PSC). PI: William Cohen (CMU)
- Use of Global Federated File System (GFFS). Presenter: Andrew Grimshaw (UVa)
- Visualizing Residual Based Turbulence Models for Large Eddy Simulation. Presenter: Mark Vanmoer (NCSA), PI: Arif Masud (UIUC)
- Running WIEN2K on Ranger with both Coarse and Fine Parallelism. Presenter: Hang Liu (TACC), PI: Luis Smith (Clark University)
- Prediction and Control of Compressible Turbulent Flows. Presenter: Robert McLay (TACC), PI: Daniel Bodony (UIUC)
- P3DFFT Introduction. Presenter: Dmitry Pekurovsky (SDSC), PI: P.K. Yeung (Geogia Tech)

These talks attract both the user community, Campus Champions and staff.

The Campus Champion Fellows program application process opened and closed during this reporting period. Applicants chose from 23 projects where they can gain expertise in many different areas of cyberinfrastructure, including sequencing analysis, performance analysis, hybrid parallelization, I/O patterns and workflows. Selections will be highlighted in the next quarterly report.

7.2 Extended Support for Community Codes 1.5.1

Extended Support for Community Codes (ESCC) efforts are aimed at deploying, hardening, and optimizing software systems necessary for extensive research communities to create new knowledge using XD resources and related technologies. ESCC projects are focused on helping users with community codes and tools on XSEDE systems.

ESCC projects may be proposed in several different ways. Over the past year, most ESCC projects, like other ECSS projects, were initiated as a result of requests for assistance submitted by principal investigators during the resource allocation process. Projects that request assistance with community codes are assigned to ESCC rather than another ECSS area such as ESRT or ES-

SGW. Two projects were initiated internally due to a recognized need for wide spread support. In the future, these projects will be reviewed by the newly formed User Advisory Committee (UAC).

Although the annual goal for ECSS was to support 10 projects, by the end of the fourth quarter, 18 projects were initiated and assigned to a consultant. Of these, two were initiated internally.

The first project, *AMBER Sustainability*, is a trial project to support an installed code across several SP's in which the consultant is a liaison between the SP and the development team. The long-term goal of this project is to provide users with up-to-date versions of AMBER, system specific documentation, and test cases for each system.

The second project, *Collaboration with the Broad Institute*, was initiated due to a large number of requests for help using Trinity or Allpaths-LG on Blacklight. Since these requests had the same goal, they were consolidated into a single project to port Trinity and Allpaths-LG first to Blacklight and then to other XSEDE resources. This single project is supporting 15 separate research groups.

Over the past quarter, three new ESCC projects were initiated and two projects completed. Also, the ESCC effort was presented to the newly formed User Advisory Committee (UAC) at the May, 2012 meeting. Current projects and future plans were covered in the initial UAC meeting. In addition, a request was made for UAC input regarding community codes the UAC considered important. ESCC updates will be provided at the next UAC meeting in August.

Below are representative summaries of one ESCC project that completed in Q2 2012 and the progress of the initiated AMBER project.

7.2.1 The CaterPillar Project -- Exploring the Dark Matter Substructure of Milky Way Galaxies (Frebel, MIT)

ECSS Project Team: Doug James (TACC)

Executive Summary: The CaterPillar research team, directed by Anna Frebel at the Massachusetts Institute of Technology, runs large-scale cosmology simulations using GADGET, an open-source community code developed and maintained by Volker Springel at the Heidelberg Institute for Theoretical Studies. The CaterPillar team and their collaborators have had great success running simulations on the Ranger cluster at the Texas Advanced Computing Center (TACC). On TACC's Lonestar cluster, however, they have encountered problems so severe that until recently they had largely abandoned Lonestar for GADGET-based simulations. Anxious to return to Lonestar, the CaterPillar team partnered with TACC under this Extended Collaborative Support Services (ECSS) project to identify and correct the problems. The collaboration proved successful: in coordination with the GADGET lead, the CaterPillar-TACC collaborators fixed the problem that had prevented use of Lonestar. As a result Lonestar is once again available to the worldwide community of GADGET users. More significantly, the update eliminates a fundamental flaw in the newest version of GADGET that would have caused problems going forward on all but the oldest of clusters.

Technical Details: The original problem occurred early in a 33 hour experiment on 1020 cores. Phillip Zukin (CaterPillar Team) was able to reproduce the problem on a 60 core test problem that exercised the same physics, and provided James with an even smaller related problem that James could probe with tools designed to detect memory leaks. James used all three problems to explore the behavior of GADGET, and was able to duplicate Zukin's results for all test problems.

Eventually Zukin and James located the point in the code where the problem occurred. After gathering as much information as they could about the conditions under which the code failed,

they contacted Volker Springer, the lead developer of GADGET. The problem was new to Springer, but he immediately recognized its cause.

At issue was an unsafe call to the MPI function MPI_Allgatherv. This particular MPI function allows each task to communicate information to every other task, and allows the message size to vary from one sending task to another. Each task stores its sent message in a data structure informally called the send buffer, and receives the results from all other processors in a data structure informally called the receive buffer. The MPI_Allgatherv calls were using overlapping send and receive buffers without the proper MPI argument.

Springer proposed to James a short-term fix pending the formal release of the next version of GADGET. In late April, James implemented and tested the fix and supplied a temporary update to the CaterPillar team. Zukin and his colleagues validated the update's correctness and stability in their own tests on Lonestar.

There are two easy fixes for this problem. One approach, available only when using an MPI-2 library, is to use an alternative method of calling MPI_Allgatherv that allows safe use of overlapping buffers. This alternative call employs a flag called MPI_IN_PLACE to direct the library to use overlapping buffers. The MPI library then proceeds under assumptions that follow from this information, and in fact manages the details.

The other fix is simply to avoid overlapping buffers completely. This is the approach that the GADGET lead developer recommended as a temporary fix for the CaterPillar team, and presumably is the one he will implement in the upcoming release of GADGET 3. This approach has an important advantage for the GADGET community: it allows GADGET to run with older MPI libraries based on the MPI-1 standard.

Impact: We are pleased that this ECSS collaboration makes it possible for the CaterPillar team and its collaborators to have access once again to the TACC Lonestar cluster. There is no doubt that the CaterPillar project will benefit going forward from the added computational power and increased flexibility. We are delighted, however, that the potential benefits exceed the short term goal of running GADGET on Lonestar. The stability, portability, and correctness of GADGET are vitally important to the larger cosmology community; we are proud to know that this collaboration has improved GADGET for all its stakeholders.

7.2.2 AMBER Installation and Testing on XSEDE (Internally initiated project)

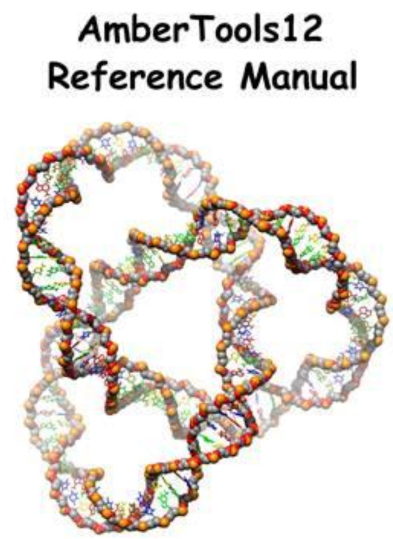
ESCC Team: Ross Walker (SDSC)

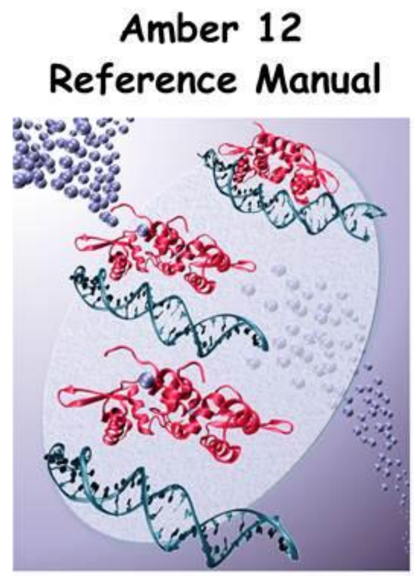
Executive Summary: The objective of this project is to install, test, validate and maintain the AMBER software on various XSEDE resources. This includes the installation of AMBERTools and AMBER in serial, parallel and GPU version as appropriate. The goal is to install the latest version of AMBER on each XSEDE resource. In addition to installation, this effort will fully test installations and optimize based on a choice of compilation parameters, MPI versions and compiler versions.

Technical Details: This project began in Q1 2012, which has nicely correlated with the release of the latest version (v12) of the AMBER and AMBERTools Software packages. The first three months of this project were spent beta testing the AMBER 12 release candidates on various XSEDE resources including, SDSC Trestles, SDSC Gordon, GT Keeneland, NCSA Forge, TACC Ranger and PSC Blacklight. A complete test installation was run each machine and the configure and test scripts were modified as necessary to fix compiler and machine specific issues such that AMBER 12 supports these machines out of the box. AMBER 12 was released in April 2012. In Q2 a blessed installation was created on each of the major XSEDE machines. Q2 has been spent trying to debug an issue we have running on the new Keeneland to make sure

AMBER 12 is ready for the machine being in full allocated production (This has also included trying to debug a number of issues on Titan as well, although that is DOE but is the same architecture as Blue Waters so it is critical to get AMBER running properly on there. This will also be useful for other XSEDE machines that might include GPUs). This work included trying to figure out why we get random lockups on M2090s.

We have things working well now so there will shortly be a large scale patch released for AMBER 12 that is being put together for the beginning of Q3. This debugging work has put a halt on optimizing the installations on the other XSEDE machines plus given the patch is quite detailed and also addresses some MPI performance it has been key to complete this first in order to avoid having to reinstall and repeat all the benchmarks on the various machines. The plan for Q3 is to release the patch and then conduct a thorough benchmark on all the XSEDE resources. This will include putting together a webpage with all the information on as a central resource for people to use when submitting XRAC proposals (and also for XRAC reviewers to be able to check proposers reported numbers).





7.3 Extended Science Gateways Support 1.5.2

The Extended Support for Science Gateways (ESSGW) is tasked to provide assistance to researchers wishing to access XSEDE resources through web portals and science gateways. The group assists both new and advanced groups and has experience in the use of web technologies, grid software, fault tolerance, complex workflows, and security and accounting aspects of the program.

The gateways developers within the ECSS Science Gateways group have worked with Architecture team and drafted the use cases. This exercise is resulting in an extensive survey of gateway architecture in regards to XSEDE integration. The effort was coordinated through contributed use cases from bi-weekly teleconferences.

The ESSGW team has prepared science gateway tutorial material within this quarter to assist XSEDE user community in developing new gateway projects.

7.3.1 ECSS SGW projects

7.3.1.1 *Einstein Genome Gateway:*

The Einstein Genome Gateway is targeted to provide large-scale computational interface between XSEDE compute and storage resources and the national community with large-scale computational needs in genomics and related fields.

The Quarry gateway hosting infrastructure is set up as a development testbed and targeted production platform. The project team is transition to a new gateway infrastructure at which point the ECSS effort will gain momentum in integrating job management and workflow abstractions to seamless migrate from Einstein medical school's computational resources to XSEDE.

7.3.1.2 *VLAB:*

The Virtual Laboratory for Earth and Planetary Materials (VLAB) project provides petascale computations in mineral physics with the Quantum ESPRESSO codes. The ESSGW support is helping the VLAB project team to build workflow infrastructure to perform first principles calculations of unprecedented magnitude and scope in mineral physics. The project will optimize the execution of the Quantum ESPRESSO software for materials simulations in XSEDE systems.

7.3.1.3 Ultrascan Science Gateway:

The major component of ECSS support involved enabling new Ultrascan SOMO analysis code on XSEDE resources. The effort lead to improving the Ultrascan Gateway user experience by incorporating new XSEDE resources such as SDSC Trestles and Gordon. XSEDE is considering new servers for job management and data movement. The team will explore those options to take advantage of new services.

Ultrascan developers deployed codes on Trestles and Gordon. ECSS staff members helped to test the code using PBS scripts, figure out "work" and "scratch" directory locations and test grid services. There were several issues to resolve with Grid services, including debugging and fixing GridFTP on Trestles.

7.3.1.4 Social Science Gateway:

ECSS personal Tom Uram and Lars Vilhuber have worked together to brainstorm requirements for the Social Science Gateway (SSG) and formulated a work plan. The workplan has now been approved. The SSG seeks to simplify large-scale statistical analysis using an interface between statistics packages SAS, Stata, and SPSS (proprietary packages which are not available on XSEDE resources) and R (which is available on XSEDE resources). By doing so, the SSG will enable users to scale from their desktop to SSG to XSEDE.

While developing the workplan, Lars has documented one of the SSG workflows, Tom installed Swift on SSG resources and set up example Swift workflows to run on SSG and XSEDE resources, and Tom has used the NX service to access the SSG resource. In the next quarter, with an approved workplan, the team will begin work on the interface to R, execution on SSG and XSEDE resources, and investigate means of simplifying data transfer between SSG and XSEDE.

7.3.1.5 CyberGIS Science Gateway:

CyberGIS Gateway continued to engage two social science domains: culturomics and political science. In culturomics, a heat map analytical approach has been integrated into the CyberGIS service infrastructure and started serving community use. In political science, the redistricting

research code has been converted to software package for CyberGIS integration. In the domain of spatial analysis and modeling, we continued to develop scalable algorithms for large- and multi-scale parallel agent-based models using XSEDE (Ranger and Trestles).

The following research publications have been produced:

Wang, S. (Accepted). "Special Issue on CyberGIS: Blueprint for Integrated and Scalable Geospatial Software Ecosystems." International Journal of Geographical Information Science.

Shook, E., S. Wang, et al. (In Review). "A parallel communication framework for spatially-explicit agent-based models." International Journal of Geographical Information Science.

For the next quarter the project will continue to work with researchers in culturomics and political science for integrating and outreaching the developed analysis methods as community services. A parallel version of the redistricting problem solver will be developed to leverage massive processor cores on Ranger (up to 16,384) to handle fine-grained redistricting research needs (at block level).

7.3.1.6 SimpleGrid:

The effort focused on finishing the development of the new version of SimpleGrid with the following achievements:

- Successfully developed SimpleGrid gadgets using Apache Rave by closely working with the Indiana University team
- Upgraded SimpleGrid Toolkit to manage more XSEDE resources by adjusting to the dramatic change on XSEDE middleware
- Applied for a hands-on tutorial at XSEDE'12 and got approved
- Developed tutorial materials, including software, tutorial servers, and documents, for the hands-on tutorial at XSEDE'12

Next quarter will conduct a hands-on tutorial for building XSEDE science gateway applications at XSEDE'12. Based on the feedback, we will further improve the usability of SimpleGrid Toolkit. We will explore other venues for the outreach and education of science gateway development using SimpleGrid Toolkit.

7.3.2 Other ESSGW Efforts

Yan Liu provided technical support for allocated projects that were awarded with Open Science Grid (OSG) resources. Two tickets have been received from the University of Delaware and the Northeastern University. A sample application package was developed to provide a sandbox solution for illustrating seamless usage of several XSEDE clusters (Trestles, Steele, and Ranger) and the XSEDE OSG cluster. User requirements on third-party library dependency were captured in the package to demonstrate the cohesive compiling and running practice on these XSEDE resources. This package will efficiently resolve similar requests as users allocated with OSG resource are often confused with the difference of XSEDE and OSG.

During next quarter the effort will be focused on applying the XSEDE-OSG sample application package to the two identified use cases and help users run their codes on both XSEDE and OSG seamlessly. Since running jobs on XSEDE and OSG are significantly different, we expect multiple rounds of support and close coordination with the OSG team. We will continue to support helpdesk tickets related to the XSEDE science gateway program.

7.4 Extended EOT Support 1.5.3

Over the first project year, ESTEO has contributed to many tutorials, mentoring opportunities, meetings, as well as presented numerous talks and presentations at scientific and high performance computing conferences. Many staff were also preparing for local summer institutes this quarter.

ESTEO staff initiated the contributions, developing and delivering the content at a number of venues including the Seismological Society of America and FEniCS 12 (a collaborative project for the development of innovative concepts and tools for automated scientific computing, with a particular focus on automated solution of differential equations by finite element methods). During this period, staff had content accepted for XSEDE12 and submitted to SC12. ESTEO staff reviewed submissions for many conferences as well, including 12th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP-12), 12th International Conference on Computational Science (ICCS 2012) and XSEDE12. Some highlights of the activities include special focus topics on accelerators (ranging from introductory to advanced), introduction to new compute resources such as the upcoming Stampede resource at TACC, as well as broader topics for users new to XSEDE and high performance computing.

ESTEO work for the period continued the intense focus on preparing materials for upcoming conferences. XSEDE12 and SC12 are targets of several tutorials and presentations by our staff. Phil Blood has been coordinating the technical program for XSEDE12; this quarter included significant staff participation reviewing the proposed papers, posters, tutorials, and presentations, as program chairs and/or program committee members. Some upcoming tutorials will cover:

- Accelerator programming: OpenACC and intermediate CUDA
- Introduction to XSEDE
- Introduction to Unix/Linux
- Eclipse IDE and the Parallel Tools Platform

During the period, staff continued to provide outreach and training activities even as they prepare new material for large events later in the year. A number of staff participated in developing and delivering content for the XSEDE-PRACE "2012 European-US Summer School on HPC Challenges in Computational Sciences", in Dublin, Ireland on June 24-28. Frank Willmore taught part of Scientific and Technical Computing at the University of Texas, in collaboration with Jay Boisseau.

A number of staff, including Ross Walker, Phil Blood, Yifeng Cui, and Lars Koesterke mentored students this quarter. This also included mentoring activities at tutorials, as needed.

We also continue to review web tutorials and provide new content.

The full list of training courses for the period is now available online at https://www.xsede.org/web/xup/course-calendar.

8 Education and Outreach 1.6

8.1 Overview

The Education and Outreach team accomplished their year one goals working to engage a larger and more diverse community of practitioners including researchers, educators, and students from 2- and 4-year colleges and universities, Minority Serving Institutions, and EPSCoR institutions. The team engaged under-represented faculty and students from Minority Serving Institutions and majority serving institutions, and continued to leverage external partnerships with the NSF funded Blue Waters project, NSF funded EPSCoR projects, and the Partnership for Advanced Computing in Europe (PRACE).

A new partnership with the HPC Wales project has begun with an initial emphasis on sharing training materials, developing science gateways, and launching a Campus Champions program. Members of the HPC Wales team attended XSEDE'12 to advance the collaboration.

The Training, Education and Outreach Services (TEOS) formed and engaged a TEOS Advisory Committee, conducted an annual Community Requirements survey, and worked with the External Evaluators from I-STEM to conduct formative evaluations of the TEOS programs. These efforts have helped to improve the resources and services provided by TEOS.

TEOS and ESTEO have agreed to conduct monthly conference calls to conduct longer-term planning and coordination of events requiring assistance from ESTEO staff, and to include ESTEO recommendations in the TEOS planning of activities and events.

During the next quarter, XSEDE TEOS will be working with the new Blue Waters Community Engagement activities that will begin in October to coordinate activities among the two projects. Details of the coordination will be documented in the next three to six months as the Blue Waters activities are solidified.

8.2 Education – WBS 1.6.1

Education activities for the fourth quarter continued all of year one activities: 1) conducting educational programs for faculty and students, 2) development of competencies for computational science programs, 3) initiating collaborations with universities interested in starting or expanding formal computational science degree and certificate programs 4) Creation of computational science materials for pre-service science and math educators.

8.2.1 Educational Programs for Faculty and Students

In the fourth quarter, we continued to work on several efforts to expand the reach of our education efforts. The lectures from the Berkeley parallel computing course, CS267 was captured in Virtual Workshop format, including video of all lectures and the presentation materials. After the course was finished, a sample module was created from a portion of the first lecture. The sample includes the unabridged material in both video and text formats. The video primarily shows slides, sometimes cutting to the speaker. The slides used in the module however are native, i.e., they are not the video of the screen during the lecture. The recorded video is only used to occasionally show the lecturer, and for the audio of course. For the sample slides, the complete recording was used; in later sections, excerpts might be more appropriate. All Virtual Workshop features, e.g. page notes and bookmarks, are functional in this sample module as well, but only if you are registered and logged in. We will be working with the Berkeley faculty in the coming months to add quizzes and other feedback mechanisms so that we can pilot an online version of the course this coming autumn or winter.

We continued planning for a number of workshops for faculty the summer and ran several workshops in the spring. We completed two one-day workshops providing an overview of

computational science educational resources and tools for community college faculty in Nevada as college sites in Las Vegas and Elko. Those workshops were co-sponsored by the Nevada EPSCoR office. There were 12 attendees at the College of Southern Nevada, and 7 at Great Basin College. We are exploring other opportunities to work with those campuses and the Nevada EPSCoR office on other joint activities.

We completed the arrangement for five summer workshops on a variety of topics and opened registration that are listed in Appendix H. A scholarship program was initiated that subsidized faculty attendance at the workshops giving priority to those working on formal academic programs with XSEDE. Two of the workshops were in June. A summary of the workshop results will be provided in the first quarter of 2012.

Planning is also underway for the second offering of the parallel programming summer bootcamp at University of California, Berkeley August 15-17. Last year that workshop attracted the participation of several hundred people from around the world.

Development of Competencies

Work continued on the development of model competencies for both undergraduate and graduate programs in computational science. The current version of the competencies can be seen at http://www.rrscs.org/competencies. We also have drafted a set of competencies for graduate students in physics and chemistry that we will be reviewing with a cross-section of faculty this autumn. Based on the competencies, we have developed a new metadata framework to index materials on computational science education topics. Along with the Shodor Education Foundation, we will be working to add this framework to the HPC University site (http://www.hpcuniv.org) coupling it with a review mechanism that will provide a comprehensive list of existing training and education materials to share with the national community. We expect a pilot test of the revised site by November 2012.

8.2.2 *Initiating University Collaborations*

Over the past quarter, we have continued work with the campuses interested in formal computational science programs. We have been engaging with the campuses through audio conferences and campus visits. These sessions have included a description of the model programs and competencies, discussion of the current expertise on each campus, and establishment of working groups to help create or enhance formal computational science programs. We also participated in two regional conferences at minority serving institutions and reviewed efforts and approaches to computational science education programs. Table 1 summarizes the activities for the spring quarter.

Table 1: Institutions with Interests in Computational Science Programs

Institution	Status and Interests
Kean University	Campus visit on April 26 attended by 12 faculty 2 staff, and 4 administrators. Commitment made by administrators to support a broad based computational science minor program and a possible graduate program.
Nashville Regional Conference	Regional conference including faculty and graduate students from a number of Tennessee institutions. Attended by 11 faculty 11 staff and 23 graduate students. Introduced them to XSEDE resources and the potential for a formal computational

	science program.
Norfolk State University	Interested in graduate programs in computational science. Visit made on April 17, 2012 attended by 15 faculty and 5 administrators.
North Carolina	Follow up video conferences held with North Carolina State (6 faculty) and North Carolina Asheville (8 faculty). Discussion of minor programs and how they fit into the current curricula at each institution.
North Carolina Central	Participated in regional conference on May 14-15 attended by 18 faculty. Introduced them to XSEDE resources and the potential for a formal computational science program.
Southern University	Interested in both undergraduate and graduate concentrations and new graduate degree. Visit completed on April 2, 2012 and working committee formed to start program preparation.
Stockton College	Expanded program endorsed by campus administrators. Campus visit was made on April 27 attended by 12 faculty 2 staff and 4 administrators. Consensus to build a broad based undergraduate minor program was reached.
The Ohio State University	Working on integration of modeling and simulation into pre-service science teacher education program. Worked to create model exercises in chemistry, physics, and biology that will be tested in a summer course for pre-service teachers.

We continued to work with Dr. Karen Irving in the OSU College of Education on the preparation of example lesson plans that include computational modeling. Three modules were developed that will be tested in a methods class for pre-service teachers in the summer of 2012.

8.2.3 EU-US HPC Summer School

From more than 230 applications, 60 graduate students and postdocs from four continents -- including 30 percent women -- were selected from higher education institutions across Europe and the United States to attend the 3rd Annual Summer School on Computational Challenges in High-Performance Computing, June 24-28, 2012. The event was hosted by the Partnership for Advanced Computing in Europe (PRACE) and the Extreme Science and Engineering Discovery Environment (XSEDE). The students met at the Royal Marine Hotel in Dublin, Ireland, where the Irish Centre for High-End Computing (ICHEC) -- the Irish member of PRACE, took care of local coordination and hospitality. Nationalities of the students spanned several continents including Asia, Europe, the Middle East, and Central and North America.

This was the third in an ongoing series of summer schools jointly organized by PRACE and XSEDE with funding from the European Commission (EC) and the United States National Science Foundation (NSF). The goals of the summer school series are to expand the knowledge of the attendees about high-performance computing (HPC) and its applications in multiple fields of science, technology, engineering and mathematics. The summer school stimulates international collaborations and friendships among the attendees and presenters through the unique global setup of the school.

Attendees were immersed in a mix of presentations and hands-on sessions held by more than 20 leading researchers and HPC professionals from both sides of the Atlantic. The program placed

particular emphasis on how HPC is being applied to meet current and future computational challenges in various scientific disciplines, as well as the relevant tools and techniques for tackling different problems. Meals and social activities were coupled with mentoring to form an integral part of the school's program, which was designed and organized to promote interaction among the participants.

8.2.4 *Extreme Scaling Workshop*

XSEDE and Blue Waters collaborated on the 5th in a series of Extreme Scaling Workshops on July 15-16, in Chicago. The workshop addressed algorithmic and applications challenges and solutions in large-scale computing systems with limited memory and I/O bandwidth. The presentations and discussions were intended to assist the computational science and engineering community in making effective use of petascale through extreme-scale systems across the spectrum of local campus-scale to national systems. The workshop was conducted in coordination with ACM SIGHPC. The workshop engaged 45 participants. A workshop report will be produced in the next quarter, and the report and the papers submitted for the workshop will be posted on the ACM Digital Library. The workshop web site is https://www.xsede.org/web/xscale/.

8.2.5 *Other Activities*

A number of visits were made this quarter to engage faculty and students in computational science activities. The visits are summarized in Table 2.

Table 2: Visits Made		
Location and Date	Description	Attendance
April 17 Dept of Defense Dependent Schools (DoDDS), Stuttgart, Germany	Math Teachers: modeling skills in the sciences	9 teachers
April 18-19 University of Freiburg (Germany)	Workshop on System Dynamics and Modeling in Biology	9 faculty 3 post docs 3 grad students.
April 20 Dept of Defense Dependent Schools (DoDDS), Wiesbaden, Germany	Math Teachers: modeling skills in the sciences	7 teachers
April 21-28 DoDDS-Europe, Oberwesel, Germany	STEMposium, week of computational modeling and simulation tutoring, workshops and mentoring	90 students, 17 teachers
May 10 East Carolina University	Half day workshop on Computational Science	4 faculty
May 11 Pamlico Schools, Bayboro NC	Introduction to Computational Thinking half day workshop	3 administrators, 6 teachers
May 22-25 Universidad de Sagrado Corazon, San Juan, PR	Intro to Computational Thinking workshop	15 faculty
May 30-31 Union College, Schenectady, NY	Half day workshop: Computational Thinking across the Curriculum	6 faculty

June 5-9 University of Colorado, Boulder	AgentSheets teacher Institute, Agent Modeling in Science (2 full days)	30-40 teachers attended all or most of this part of the workshop
June 11-14 Southern University	Computational Thinking with a Parallel Perspective	21 faculty and graduate students
June 18-22 Harkers Island, NC	Modeling your World in partnership with The Bridge Downeast	7 high school girls
June 25-29	Modeling your World in partnership with The Bridge Downeast	17 middle school kids
June 21 West Carteret High School, Morehead City, NC	Modeling in Marine Science	43 middle school students, 4 teachers

We also participated in the planning for the XSEDE12 workshops helping to solicit papers for the training and education track, reviewing papers, and planning for that portion of the agenda.

8.2.6 Challenges

The main challenge continues to be is the slow pace of change for programs at the universities. Faced with very tight budgets at all higher education institutions, the prospects for creating new programs may become even slower. Nevertheless, an interdisciplinary program like computational science is being well-received on the campuses and makes it likely for us to succeed in establishing new programs in the long run.

8.3 Outreach – WBS 1.6.2

8.3.1 Under-represented Engagement

The Southeastern Universities Research Association (SURA) team has completed their outreach target for program year one, and submitted requested events for program year two.

For XSEDE’12, they supported 7 minority community Campus Champions to attend the annual conference. SURA also conducted a face-to-face meeting of the Minority Research Community, from which a report will be produced in the next quarter.

Rice University’s first cohort of 40 students involved in the XSEDE Scholars Program have completed their year-long involvement during program year one. The Rice University team has selected a second cohort of 40 students for the XSEDE Scholars Program for program year two. The students received on-line training by completing CI-Tutor sessions on Linux and C programming. The students attended XSEDE’12, anchoring the conference Student Program. The XSEDE Scholars participated in the Student Programming Contest, with one team (out of seven Scholar teams and a total of 11 teams) took first place in the event. The winning team members were: Manuel Zubieta, Justin Peyton, David Manosalvas, Nancy Carlos, Melissa Estrada, Grace Silva with Coach: Alice Fisher.

Rice held the first meeting of the Minority Faculty Council at XSEDE'12, led by Richard Tapia. A report will be produced during program year two.

8.3.2 *Speaker's Bureau*

The team is developing a list of candidate events for program year 2 that include:

- SIGGRAPH
- American Chemical Society
- NOBCChE's 39th Annual Conference
- SACNAS National Conference
- Grace Hopper Celebration of Women in Computing
- AAAS
- American Physics Society
- Richard Tapia Celebration of Diversity in Computing
- Emerging Researchers Network
- Joint Annual Meeting (NSF Diversity Programs)

ESTEO has agreed to more aggressively recruit staff and researchers from within XSEDE to help staff these events. Outreach Services will provide exhibitor-level registration support for XSEDE attendees who help staff the booth.

8.3.3 *Campus, Regional and Champion Events*

Campus Champions have been encouraged to consider hosting an Information Day, seminar, or workshop on their campus in program year two. Outreach Services will help staff the event as needed. The goal is for 5-7 campuses to host a local or regional campus event in program year two.

XSEDE Outreach has been approached by the Blue Waters program to more closely coordinate outreach events for the purpose of presenting a more unified CI environment for computation STEM research. Blue Waters' funding for this expansion of their outreach activities is anticipated to begin October 1, 2012. XSEDE and Blue Waters outreach will meet during the next quarter to establish a plan of action for making this leveraged presence effective for both projects.

8.3.4 *Student Engagement*

All 15 of the students in this summer's Student Engagement program participated in XSEDE'12. All of the students were required to present a brief overview of their projects and status at a conference session on Tuesday, July 17. The full slide stack is on the staff wiki in the Outreach - Student Engagement space: https://www.xsede.org/web/staff/staff-wiki/-/wiki/51230/Outreach#section-Outreach-StudentEngagement

Six students presented posters of their summer projects:

Authors	**Title**
Kurt A O'Hearn, David E Hudak, John Eisenlohr and Karen Tomko	"Parallel Application Development for the MIC Architecture"
Matthew Clark, Ross Walker and	"An extensible interface for ab initio QM/MM

Andreas Goetz	molecular dynamics simulations with AMBER”
Nicholas Woodward and Weijia Xu	“Towards Dynamic Provisioning and Optimization of a Hadoop Cluster: A New Approach for Data I/O in HDFS”
Graham Lopez	“Keeneland Performance Analysis”
Huanhuan Xu, Pietro Cicotti and Laura Cervino	“Dose Verification in lung cancer radiotherapy”

Kurt O’Hearn received the second highest score in the Student Poster contest in the Undergraduate category. Nicholas Woodward received the second highest score in the Student Poster contest in the Graduate Student category. One student presented a poster based on his thesis work: Pedro Da Silveira, Renata Wentzcovitch and Suresh Marru: “Science Gateway Job Scheduling Framework to Support Community Account Fair-Share Policy”

The XSEDE Student Engagement Program has begun discussions with Blue Waters to expand and coordinate efforts in workforce development, based on the Blue Waters “Community Engagement Project”, which was recently approved, with funding anticipated to begin October 1, 2012. The XSEDE Student Engagement activities will be coordinated with the Blue Waters Internship Program, to provide petascale-level opportunities to undergraduate students through a mature program in place at Shodor.

XSEDE Student Engagement is also in discussion with Clemson University to submit an REU Site proposal to NSF (due September 12, 2012). The goals of this award would be:

- Extend the Student Engagement experience to encompass a 12-month program, as originally proposed; and
- Provide a competitive stipend to students participating in the program.

Including the successful Clemson EPSCoR student engagement program will leverage their expertise in mentoring and coaching, supplementing the project activities (Student Engagement, XSEDE Scholars) provided by the XSEDE program.

8.3.5 Campus Champions

A major highlight of the year is the formation of two working groups of Campus Champions, and a significant increase in communications and community building among the Champions. Another major highlight of the year is that the Campus Champions now cover 48 states and the Virgin Islands. Discussions are underway with to establish Campus Champions in Nevada, Maine, and Puerto Rico. The HPC Wales team has begun discussions with the XSEDE Campus Champions to replicate the success of the US program in Wales. Finally, the User Services team has begin conducting regular calls and inviting the Campus Champions to attend.

8.3.5.1 Campus Champions Working Groups

Phil Blood has been assisting the Campus Champions with the launching of two working groups – Outreach and User Assistance. Biweekly meetings were held over the span of multiple months. A Face to Face meeting of both working groups was conducted at Notre Dame on May 23-24. The working group meetings resulted in the development of a “Welcome Packet” that was assembled for distribution at XSEDE’12, and a set of “Best Practices” that were formulated for review at XSEDE’12. Campus Champions working groups (Outreach and User Support)

completed the Champions Welcome Packet (printed) and Champions Survival Guide (USB) for the XSEDE'12 conference. The working groups are currently accepting and processing feedback on distributed materials

8.3.5.2 Campus Champions Fellows

Three out of four Fellows were at XSEDE'12. They met with their project teams and are starting to get oriented to the work. MOUs are all signed by Fellows, Laura (coordinator). Delivered to Nancy (ECSS coordinator) to continue processing.

XSEDE announced the four selected Fellows:

- Dirk Colbry, a research specialist at the Institute for Cyber-Enabled Research at Michigan State University,paired with Ritu Arora, a research associate in high-performance computing at the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. Colbry brings many years of high-end computing expertise to the program and hopes to learn more about large-scale data analysis. Colbry and Arora are supporting PI Virginia Kuhn's Interactive Large-Scale Video Analytics project. Kuhn is associate director of the Institute for Multimedia Literacy and assistant professor in the School of Cinematic Arts at the University of Southern California (USC). The project focuses on indexing, tagging, and searching vast media archives in real time, applying a hybrid process of machine analytics and crowd-sourcing tagging. SDSC's data-intensive Gordon supercomputer will be used as a resource for this project.
- Nasseer Idrisi, an assistant professor at the University of the Virgin Islands' Center for Marine and Environmental Studies, paired with Kwai Wong, a computational scientist with the National Institute for Computational Sciences (NICS) at Oak Ridge National Laboratory. They are supporting principal investigators John Bryant Drake and Joshua Fu at the University of Tennessee Knoxville, both of whom are doing climate simulation. Working with Wong, Idrisi will expand his knowledge in parallel programming while contributing a domain viewpoint to the projects.
- Liwen Shih, professor and computer engineering chair at the University of Houston-Clear Lake, paired with Yifeng Cui, a research scientist at SDSC. Their project focuses on Cui's work on physics-based seismic research in collaboration with SDSC and the Southern California Earthquake Center (SCEC). Thomas Jordan at USC is the SCEC principal investigator. Phil Maechling is SCEC's information technology architect and also a member of the XSEDE Advisory Board.
- Jack Smith, a research staff member with Marshall University and cyberinfrastructure coordinator at the West Virginia Higher Education Policy Commission, paired with Yaakoub El Khamra, a research engineering/scientist associate with TACC. Smith has extensive experience in many programming areas, as well as chemistry and material science, and would like to use knowledge gained in the Fellows program to grow research programs in new areas. They are working on the projects of PI Ronald Levy (Rutgers) and PI Thomas Bishop (Louisiana Tech). Both involve the use of SAGA software to manage large numbers of loosely coupled bioinformatics calculations.

8.3.5.3 New Campus Champions

Five new Campus Champion institutions were added in the recent quarter:

- The Citadel
- Clark Atlanta University (Minority Serving Institution)

- George Mason University
- Northwestern State University
- Carnegie Mellon University

8.3.5.4 Web Presence

Champions team met with Web Team to discuss updates to web pages.

8.3.5.5 Outreach

The Campus Champions program had a strong presence at the annual Great Plains Network (GPN) Conference. Kay Hunt led a Champions BOF discussing Regional Champions. Paper presentations/panel were given by several Champions: Jeff Pummill, Dana Brunson, Dan Andersen and Rick McMullen.

Phil Blood organized the Campus Champions program at the XSEDE'12 Conference. Jeff Pummill led the Champion dinner and program. Tim Still let the Champions focus group session. Phil Blood led a Campus Champions panel session. Kay Hunt organized a breakfast meeting with the HPC Wales visitors.

8.3.5.6 Campus Champions Training

Kim Dillman organized a survey of topics that was distributed to Champions at XSEDE'12; priorities will be set based on survey results. Champion-specific training will be developed within the Champions program; XSEDE-specific and general topics will be made available through XSEDE Training services.

The following training sessions were offered:

- New Champion Orientation – June 12 – Kim Dillman
- Presented "Selecting and Using XSEDE Resources Effectively" to Clemson student program – June 26 – Kim Dillman

8.3.5.7 OSG

Dan Fraser provided a presentation on OSG to the Campus Champions during the June monthly call. Kim Dillman organized a tutorial at XSEDE'12. Kim Dillman prepared presentation about XSEDE that was presented by Alain Roy at the OSG Summer School on June 26. Kim Dillman connected University of Michigan users with Mats Rynge to set up code for use on OSG platform.

8.3.5.8 Campus Champions Picture/Bio Repository

Jeff Pummill and Scott Lathrop collected bios and photos from the Campus Champions and XSEDE staff to help Champions and the XSEDE staff get to know one another during the annual conference.

8.3.6 <u>Outreach/Cross-group Interactions</u>

8.3.5.9 External Relations

Outreach and External Relations (ER) had a very productive cross-team staff meeting at XSEDE'12. All of the Outreach activities were represented and most of the ER group was

present. There is agreement that ER should be the source for handouts and collateral for Outreach events, and that the materials need to be tuned a bit more specifically for the communities we are meeting with (computational chemistry materials for chemistry conferences, eg). Also, Aaron Dubrow (TACC) will serve as the ER liaison between ER and Outreach. This will be similar to the role that Jay Alameda has for ESTEO.

8.3.5.10 Allocations

Cross-group issues: Champions are still having questions and issues with Allocations. We will need to schedule a CC Allocations refresher session in the current or next quarter.

8.3.7 XSEDE'12

During the last quarter, many of the Education and Outreach team members were extremely engaged in helping to plan for XSEDE'12, starting with Craig Stewart as the Conference Chair.

As of this writing, the XSEDE'12 Conference has been completed, but a full report will be included in the next quarterly status report. There were a lot of breakout meetings held by the various groups and everyone is currently processing notes and developing plans for follow-up.

8.4 Community Requirements – WBS 1.6.3

8.4.1 Highlights

The TEOS Community Requirements activities proceeded on track as proposed. The TEOS team has benefitted from three major sources of feedback during program year one. The first major source of feedback was the TEOS Community Requirements Survey Report, submitted to the XSEDE Leadership Team on June 30, 2012 by Edee Wiziecki, and subsequently posted to the XSEDE wiki.

The TEOS Advisory Committee (AC) provided valuable feedback as well. Initial members included:

- Carlos Castillo-Chavez, Arizona State University
- Roscoe Giles, Boston University
- Gwen Jacobs, Montana State University
- Rubin Landau, Oregon State University
- Wilf Pinfold, Intel
- Nora Sabelli, SRI
- Valerie Taylor, Texas A&M University

Valerie Taylor, a member of both this committee and the XSEDE Advisory Board (XAB), provides a liaison between the two committees.

During the first quarter of 2012, TEOS Advisory Committee members received status reports from TEOS Managers on their focus areas and goals, and activities conducted to date. The second AC meeting on April 24, 2012 was teleconference focused on the external evaluator matrix. The AC members responded with questions, comments, and suggestions documented in the minutes from that meeting, which were posted on the wiki. Soon thereafter AC members received a report analyzing responses to the Community Requirements Survey, which focused on current planned activities and community interest in other potential activities. TEOS Level 3 Managers met with the Advisory Committee in Chicago on May 22, 2012. At the meeting, Level 3 managers presented their initial responses to the survey as well as status updates. The Advisory Committee generated a written report that will be analyzed during the next quarter to improve TEOS plans and activities.

The third major source of feedback was the annual NSF panel review report which will be analyzed during the next quarter to enhance TEOS goals and activities.

8.4.2 Next Quarter Focus:

The Community Requirements Manager synthesized the three reports' major points for a meeting at XSEDE 12. A condensed synthesis of the essential elements is currently posted to the wiki for responses from Level 3 Managers. During the next quarter, the Community Requirements Manager will compile, communicate, and coordinate the Level 3 Managers' plans, milestones, and timelines in response to the three reports.

A quarterly conference call of TEOS Advisory Committee and Level 3 Managers will be held in October, 2012. The Community Requirements Manager will oversee collection and compilation of Level 3 manager responses to the three sets of guidance, their presentation to the TEOS Advisory Committee, and communication of plans across XSEDE groups, to maximize coordination, and minimize redundancy.

8.4.3 Risks

Any proposed changes to programs carry risks of redundancy and confusion. The Community Requirements Manager will work with TEOS level 3 managers on a communication plan to extend across TEOS and XSEDE to mitigate that risk.

8.4.4 External Evaluation

8.4.4.1 Training

The Training evaluation suite has been approved and the Training Session Survey was made live on the XSEDE User Portal (XUP) in April 2012. Participants who register through the XUP for Training sessions receive an invitation through the portal to complete the survey. This data is currently being analyzed. The response rates for these surveys are considerably low. The evaluation team will explore additional methods for increasing the response rates during the next quarter.

8.4.4.2 Education

Evaluators continue to compile and analyze Dr. Steven Gordon's completed Observation Guides to ultimately build case studies on HPC program implementation in higher education. The external evaluation team is also involved in the evaluation of the 2012 Short Course on Parallel Programming at the University of California at Berkeley from August 15-17, 2012. With program administrator input, evaluators created an online survey for participants to complete at the end of the course to assess the added value of the program. Participants will receive an email invitation to complete the survey. The email will contain the relevant IRB approved consent form to participate in order to utilize the data for research purposes.

8.4.4.3 Outreach

8.4.4.3.1 ***XSEDE Scholars Program***

The evaluation has collected and analyzed data from the Intro to Linux and Intro to C Programming webinars that the XSEDE Scholars Program (XSP) has sponsored. These webinars were offered to provide a basic introduction to the topic areas so that the second cohort of scholars may have a richer experience at XSEDE12. Interim Reports for each survey administered have been generated and submitted to program coordinators. In general, participants found these webinars to be a useful introduction to the topics.

The evaluation team will also be involved with Dr. Richard Tapia's newly formed Minority Faculty Council by sharing relevant data with the group and attending meetings.

8.4.4.3.2 *SURA*

Lorna Rivera attended the May 7-8 workshop at Vanderbilt University to conduct observations and interviews. An online survey was administered after the event. 100% of the respondents had never attended an XSEDE workshop or event and 96% agreed or strongly agreed that they are interested in learning more about the resources and opportunities available through XSEDE. An Interim Report for this workshop has been generated and submitted to program coordinators.

8.4.4.3.3 *Campus Champions*

Lorna Rivera attended the first face – to – face working groups meeting on May 23 & 24 at the University of Notre Dame. To obtain feedback on how the Campus Champions Working Groups Meetings might be improved in the future, Campus Champions and leadership team members (excluding project coordinators and XSEDE program managers/leads) were surveyed. A paper and pencil survey consisting of 4 open-ended questions was administered at the end of each working group meeting to capture the immediate impressions of the participants. 100% of the participants responded to the survey on both days. Overall Champions felt as though the meeting was a good use of their time and contributed to moving the working groups forward. A more detailed Interim Report on this survey has been generated and submitted to program coordinators.

8.4.4.3.4 *Student Engagement Program*

Lorna Rivera attended the Student Engagement kickoff event at the Pittsburgh Supercomputing Center (PSC) May 30 & 31 to conduct observations and interviews with students and supervisors. An online survey was released following the kickoff for students to complete. 73% (11/15) of students responded to the online survey. 60% of respondents agreed or strongly agreed that the kickoff met their expectations. During the interviews, supervisors clearly expressed the value of having summer students and their desire to extend the program. A detailed Interim Report on the survey and interview data has been submitted to program managers. A post-participation survey will be generated and released this quarter to determine the impact of the program.

8.4.4.3.5 *Campus Bridging*

Evaluators continue to work with Craig Stewart and Rich Knepper of IU to formalize the evaluation plans. Lorna Rivera will be attending regular Campus Bridging phone calls to conduct observations and solidify the evaluation plans based on developing program activities and timelines.

8.4.4.3.6 *XSEDE12*

XSEDE12 was evaluated by the TEOS external evaluation team. The evaluation team administered paper and pencil surveys to conference attendees. These surveys included an overall Conference Survey, a Student Survey, and a Tutorial Survey. Evaluators also conducted observations and informal interviews at the conference. An evaluation report will be submitted to Craig Stewart, John Towns, Nancy Wilkins-Diehr, and Scott Lathrop by the beginning of September 2012, in time for review and discussion at the next quarterly meeting.

8.4.4.3.7 *Extreme Scaling Workshop Evaluation:*

The 2012 Extreme Scaling Workshop organized by Blue Waters and XSEDE was evaluated by the TEOS external evaluation team. The evaluation team administered a paper and pencil survey to workshop attendees on the final day of the event. The survey was designed to provide useful information to program administrators for future workshop improvement and to offer insight into the workshop's efficacy. Evaluators also conducted observations of the workshop and informal interviews with participants. An evaluation report will be submitted to Scott Lathrop by August 20, 2012.

8.4.4.3.8 *Plans for Next Quarter*

In addition to the future activities already stated in this report, evaluators will continue to further analyze data collected during the summer of year 1 and produce formal reports including the annual evaluation report. Evaluation plans for the Campus Champions Fellows Program and the Campus Bridging GFFS Pilot Program will also be outlined.

8.5 E&O Infrastructure 1.6.4

The Education and Outreach Blog was grown by 41 entries in the 4th quarter, for a total of 132 newsworthy, vetted items for the year, nearly all events targeted at educators and students. Events featured in the Blog originate with XSEDE TEOS members, and the scrounging work done by Ange Mason and Jim Ferguson to find events of interest to share with the community.

During the 4th quarter, XSEDE External Relations granted direct access to TEOS Infrastructure to do limited editing of the E&O web page on the XSEDE site. Some of the E&O pages that needed updates were taken care of in the 4th quarter, and others that would like an overhaul will get early in PY2, done by TEOS Infrastructure staff. TEOS Campus Champions edit their own section of the XSEDE web site.

Facilities provided by the XSEDE Portal for use in registration of students for E&O programs in the student engagement and education areas have proven inadequate for the job in the first PY. Communications between TEOS personnel and the Portal team are ongoing, with the Portal have been made aware of the problems and required functionality needed of the Portal. If the Portal team cannot meet the requirements in a reasonable amount of time, alternative (non-Portal) avenues of achieving functionality will be identified and developed for use by the TEOS teams.

TEOS Infrastructure staff are closely involved with the HPC University effort, which meshes with the goals of the group. The staff have consulted on a new design and organization of the site to better serve students who may be directed there to learn about computational science in the high-performance computing area. Educational and Training materials created and collected by the XSEDE project will be collected and organized at HPC University.

New spaces on the XSEDE portal were requested and created for the XSEDE Campus Bridging pilot projects. Members of these projects from different non-partner universities will have access to these spaces along with XSEDE staff who are working on various aspects of the XSEDE Campus Bridging effort. The space was only visible to the participants to start, but is likely to be a public resource eventually, after the pilots have helped identify issues and solutions to problems that will inevitably arise.

8.6 Campus Bridging - WBS 1.6.5

The critical and overarching theme for program year one was transforming campus bridging from a very amorphous state of "sounds like a great idea . . . what is it?" to a point where we have a clear definition of what the XSEDE Campus Bridging team would like to accomplish over the course of the 5 year XSEDE award. The Campus Bridging team engaged in three major activities during this year – which were iterative and overlapping:

- discussion of campus bridging within XSEDE and within the national cyberinfrastructure community
- preparation of documents and presentations describing what campus bridging is
- work on pilot projects for testing the Global Federated File System (GFFS) software on campuses

8.6.1 *Discussion of campus bridging within XSEDE and within the national cyberinfrastructure community*

One of the side effects of the very great deal of public input and public attention that surrounded the activities of the National Science Foundation Advisory Committee for CyberInfrastructure was that community expectations about XSEDE's Campus Bridging efforts ran far ahead of actual planning work related to campus bridging within XSEDE, or the preparation of products useable in a campus bridging context. To address this, a tremendous amount of effort was spent addressing the question "what does campus bridging mean to you?" Such discussions were held in BOFs and panels at SC11, Internet2 meetings, XSEDE organizational and planning meetings (including several meetings with Campus Champions), computing organizations (e.g. Coalition for Academic Scientific Computation, EDUCAUSE), participants in and leaders of the Open Science Grid, and many other community members.

Formal liaisons were established to foster on-going communications as follows:

- Open Science Grid (with Dan Fraser assigned as liaison from OSG to XSEDE Campus Bridging)
- EDUCAUSE ACTI (Rich Knepper assigned as liaison from XSEDE Campus Bridging to EDUCAUS ACTI)
- XSEDE Architecture & Design (Craig Stewart assigned as primary liaison to A&D; Knepper secondary).

The Campus Bridging team also set up discussion areas in the community forums of the XSDE User Portal. The Campus Bridging team provided seven presentations describing campus bridging and facilitating discussions about Campus Bridging in XSEDE, that are included in Appendix H.

8.6.2 *Definition of use cases and quality attributes for XSEDE's Campus Bridging Efforts*

Working with the Architecture & Design team, the XSEDE Campus Bridging team finalized two technical reports on XSEDE's Campus Bridging use cases. These documents serve two purposes. First, they define the functions that the XSEDE Campus Bridging team would like XSEDE as an organization be able to fulfill during the 5 year XSEDE award. Second, they have now become the model for the use case development process within XSEDE. These documents are:

- Stewart, C.A., Knepper, R., Grimshaw, A., Foster, I., Bachmann, F., Lifka, D., Riedel, M. and Tuecke, S. Campus Bridging Use Case Quality Attribute Scenarios, 2012. Available from: http://hdl.handle.net/2022/14476.
- Stewart, C.A., Knepper, R., Grimshaw, A., Foster, I., Bachmann, F., Lifka, D., Riedel, M. and Tuecke, S. XSEDE Campus Bridging Use Cases, 2012. Available from: http://hdl.handle.net/2022/14475.

The Campus Bridging team is working with Architecture & Design on Level 3 decomposition documents, which are nearly complete as of the end of the current reporting period.

With these use cases now defined, the Campus Bridging team put a great deal of effort into defining a proposed plan of action for the remaining four years of the initial XSEDE award. Discussions and a variety of informal and formal presentations gave rise to the paper "*What is Campus Bridging and what is XSEDE doing about it?*" which was accepted for inclusion in the XSEDE12 conference. A tremendous amount of effort and time went in to the process of presenting drafts of this document, and parts thereof, to constituencies within and beyond XSEDE, and refining and editing the paper as a result. Presentations were given to the XSEDE Advisory Board, the XSEDE TEOS Advisory Board, the NSF review panel that conducted the first annual review of XSEDE, and a teleconference of the Campus Champions.

The completion of the paper the paper "*What is Campus Bridging and what is XSEDE doing about it?*" – which was presented at XSEDE12, after the end of the reporting period but distributed in preprint form during the current reporting period, is a major milestone. It constitutes a suggested work plan, which may now be used as input into the normal XSEDE prioritization and budgeting plan for the remaining four years of XSEDE. It presents a succinct and clear answer to the questions "what does XSEDE mean by campus bridging?" and "how does the Open Science Grid relate to campus bridging?"

In addition to this paper, the Campus Bridging team proposed two panels that were accepted for presentation at XSEDE12:

- Security for Science Gateways and Campus Bridging (Jim Basney, Randy Butler, Dan Fraser, Suresh Marru, Craig Stewart)
- Campus Bridging and the GFFS Pilot: Pilot site reports (Jin Ferguson, Guy Almes, Toby Axelsson)

The XSEDE12 conference was very definitely a defining moment for XSEDE Campus Bridging. We have a plan: now the questions are 1) where do our plans fit into the overall picture of XSEDE plans, priorities, and constraints; and based on the tasks and goals that are approved and set by XSEDE leadership as prioritized and funding Campus Bridging activities, can we have positive impact on how XSEDE supports the national open research community – and how that community works together to support itself?

8.6.3 GFFS Pilot Project

Working with the A&D team, SD&I team, and operations team, the XSEDE Campus Bridging team initiated a pilot program to test the XSEDE Global Federated File System software (GFFS). Proposals for projects were solicited. The four pilot sites and main themes of their work are as follows:

- Texas A&M University: using GFFS to move between users on campus and the Brazos file systems and to the TACC Ranger system.
- University of Kansas: using GFFS for cosmology, molecular modeling, and polar research to share data between KU, NCSA, Purdue, and Indiana University.
- University of Miami: using GFFS to simplify data transfer between Miami and XSEDE resources as well as data sharing within UM.
- City University of New York: using GFFS to facilitate researcher use of CUNY Center resources. Data will be shared via GFFS between the College of Staten Island, Miami, and Delaware with the CUNY Center to transfer large files. CUNY's team also plans to make use of GFFS to expand the capacity of existing CUNY Center file systems.

Just as the reporting period ended the latest version of the GFFS software passed XSEDE readiness review, so we will pursue completion of these pilot projects during PY2.

8.6.4 Challenges in Program Year Two

The XSEDE 12 conference, held just after the end of Program Year One, was a watershed event for Campus Bridging: we presented a clear vision for what we wanted to do, and got a very positive reaction. Now the challenge in Program Year Two will be to demonstrate some tangible and positive change in the way researchers in the US do their research as a result of XSEDE Campus Bridging efforts.

9 TAIS/Audit Services 1.7

9.1 Technology Audit Services Goals:

The Technology Audit Service (TAS) will provide both quality assurance and quality control for TG/XD to help maximize the positive impact of the infrastructure investment on science and engineering research. TAS will provide quantitative and qualitative metrics of performance rapidly to all stakeholders, namely, NSF, service providers (SP) and the TG/XD user community represented by the Science Advisory Board, interested in the optimal usage of TG/XD resources. Coupled with a strong user needs and usability analysis program, TAS will be leveraged to help ensure that users' needs are met, with particular emphasis on improving the effectiveness of major users, bringing novice users up to speed rapidly and attracting non-traditional users. By identifying user needs early on and providing the tools and information that they need, TAS will allow them to utilize TG/XD resources more efficiently and effectively.

9.2 Project Staffing

University at Buffalo:

We are pleased to report that the TAS staff has remained together for the last two years including: Dr. Thomas R. Furlani, who serves as the PI and oversees the entire TAS program, Dr. Matthew D. Jones who oversees the technical aspects of the program as well as the application kernel development; Mr. Steven M. Gallo, who leads development and implementation of the XDMoD portal and backend infrastructure, Mr. Andrew Bruno, developer of the UB Metrics on Demand (UBMoD) portal, who plays a central role in development of the XDMoD portal, Mr. Ryan J. Gentner, scientific programmer who works on the XDMoD portal development, Dr. Charng-Da Lu, computational scientist who is working on application kernel development, Mr. Amin Ghadersohi, scientific programmer who focuses on the XDMoD portal development and the XDMoD Data Warehouse, and Dr. Robert L. DeLeon who serves as the TAS Project Manager.

Indiana University (IU) Subcontract:

The Indiana University subcontract is led by Dr. Gregor von Laszewski supported by newly hired Dr. Fugang Wang. Efforts included analyzing the POPS allocation data and bringing it into XDMoD.

University of Michigan (UM) Subcontract:

The University of Michigan subcontract effort was led by Dr. Thomas Finholt now supported by newly hired Mr. Jonathan Brier. UM is responsible for determining user assessment of XDMoD and obtaining critical user feedback to optimize XDMoD to best meet user needs.

9.3 Technical Progress

The data addressed by XDMoD has been expanded to include more than simple usage. This has led to adding a number of data realms to the XDMoD data warehouse. Realms within XDMoD are used to group distinct types of data and metrics. Each realm consists of a set of data and a number of metrics derived mainly from that data. Five realms are currently supported: Jobs, Allocations, Accounts, Grants, and Proposals. XDMoD populates these realms with ingested data from both the XDcDB and POPS databases.

Realm	Description
Jobs[1]	Provides data and associated metrics for jobs run on XSEDE resources. Metrics provide details on many aspects of user jobs including wait times, job sizes, SUs consumed, resource utilization, and queue expansion factor. Data can be aggregated by user, PI, institution, resource, service provider, field of science, NSF directorate, gateway, queue, and allocation.
Allocations[1]	Provides data and associated metrics for active and expired allocations across XSEDE. Metrics provide details on burn rate, usage rate, total SUs allocated and total SUs used. Data can be aggregated by PI, resource, field of science, NSF directorate, field of science, and allocation type.
Accounts[1]	Provides general account metrics such as number of accounts created, opened, and closed for any given time period.
Grants[2]	Provides general information regarding grants awarded in support of XSEDE allocations including number of awards and total amount awarded. Data can be aggregated by field of science, funding agency, and NSF directorate.
Proposals[2]	Provides basic information regarding awarded XSEDE projects and proposals. Metrics associated with the proposals realm can be aggregated by field of science and NSF directorate.

1: Data for this realm comes from the XSEDE Central Database (XDcDB)

2: Data for this realm comes from the POPS database

9.3.1 XDMoD Portal Progress:

XDMoD (**X**SE**D**E **M**etrics **o**n **D**emand) provides a role-based web portal for mining HPC metrics data and performing statistical analysis of XSEDE usage. During the previous year, a great deal of progress was made on XDMoD including: load testing of the XDMoD application to ensure adequate response under load from multiple users, addition of the Usage Explorer and the Application Kernel Explorer interfaces, implementation of role-based access to XDMoD, increased functionality in the public role and release of XDMoD 1.6, the first publically available version open to all XSEDE users with an active XSEDE allocation. Please note that for the general public and those users lacking a current XSEDE allocation, the functionality of the public role, which requires no XDMoD account, has been greatly expanded. The remainder of this section provides greater detail on each of these areas as well as other notable achievements.

User Authentication using XD User Portal credentials:

In preparation for the first release of XDMoD available to all XSEDE users, a new authentication mechanism leveraging the existing XSEDE user portal authentication (OAuth) has been put into place. The new login screen is shown in Figure 1. The user will have the choice of logging in using local XDMoD credentials, if they have them, or using their XSEDE user portal credentials. Their choice of authentication mechanism will be remembered on subsequent visits to XDMoD.

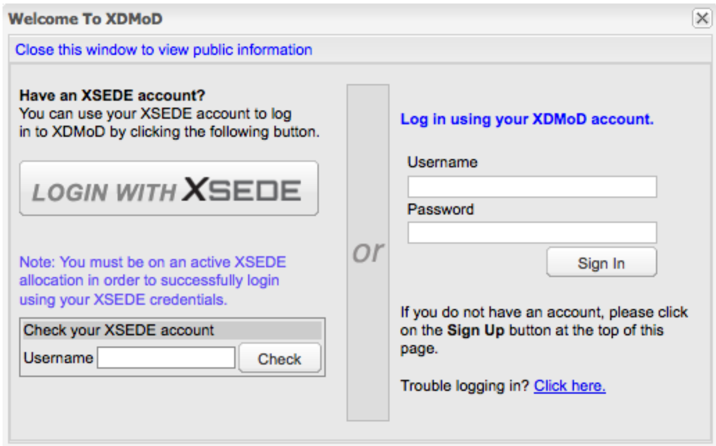


Figure 1. XDMoD login screen showing the ability for a user to login using their XSEDE user account or an XDMoD user account. The majority of XSEDE users will login via their XSEDE credentials.

Usage Explorer:

The Usage Explorer has been greatly improved both in capability and usability. Particularly noteworthy is the ability to plot multiple dependent variables (y-axes), as shown in Figure 2, which shows a screen shot of the upgraded Usage Explorer This powerful tool allows users to

- plot multiple dependent variables as data series (y-axes)
- create custom plots containing data from multiple realms (Jobs, Allocations, etc.)
- exert fine-grained control over properties of each variable
- save multiple queries for retrieval
- customize font sizes, titles, and other aspects of the chart
- add filters to each chart to restrict the displayed values

Displayed in the window is a plot that shows total XSEDE usage in XD SU's charged for the first quarter of 2012 grouped by job size on the primary (left-hand side) axis and average wait time per job on the secondary (right-hand side) axis. The plot was created by selecting "Job Size" from the "Groupings" heading and "Total XD SUs Charged" and "Avg Wait Hours Per Job" from the "Metrics" heading – three clicks in total. Under the "Chart Options" heading, the grouping and metric(s) that are plotted are displayed. The actual data upon which the plot is created is displayed in the window below the plot, which can be hidden from view if desired. The plot can also be made available for the custom report generator by dragging and dropping the chart to the "Report Generator" tab or by clicking the box that reads "Available For Report". It can also be exported in either PNG (portable network graphics) or EPS (encapsulated postscript) format. The data itself can be exported in either CSV (comma separated values) or XML (extensible markup language) format. As shown in Figure 3, the data can be filtered in a variety of ways to display only a desired subset of the data. For example, the plot shown in Figure 3 was filtered to display

only the “NICS-KRAKEN” data. Taken in its entirety, the Usage Explorer provides a powerful and flexible interface to facilitate analysis of the XSEDE usage data.

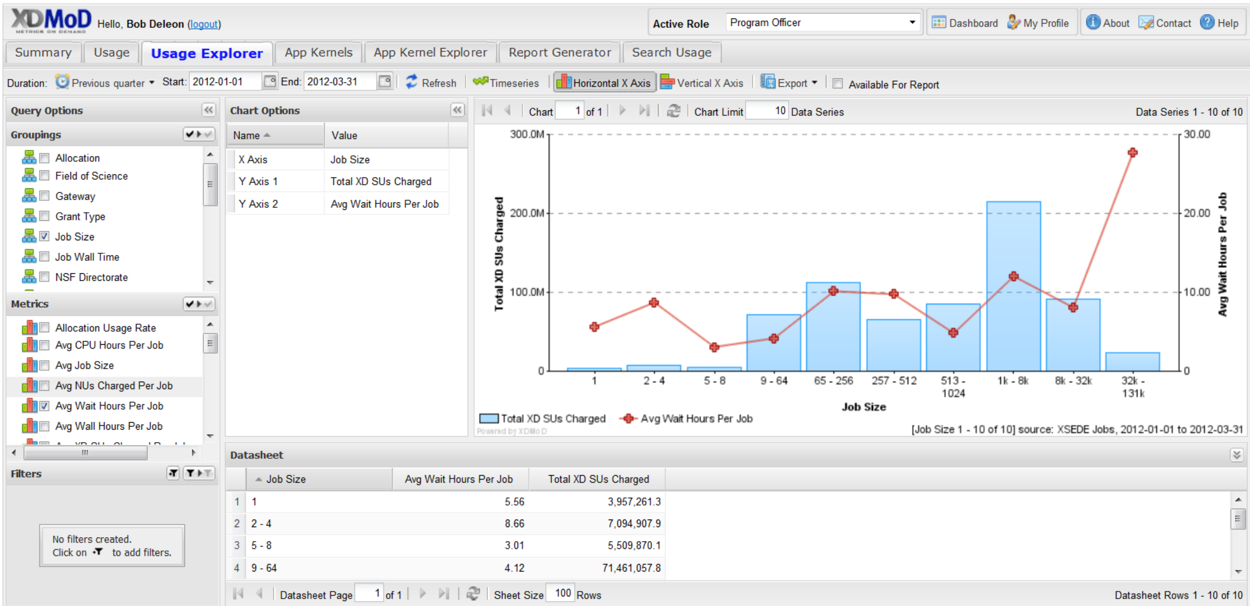


Figure 2. The XDMoD Usage Explorer. A plot of the total XSEDE usage for the first quarter of 2012 is shown on the primary axis (left-hand axis) and the average wait hours per job is shown on the secondary axis (right-hand axis).

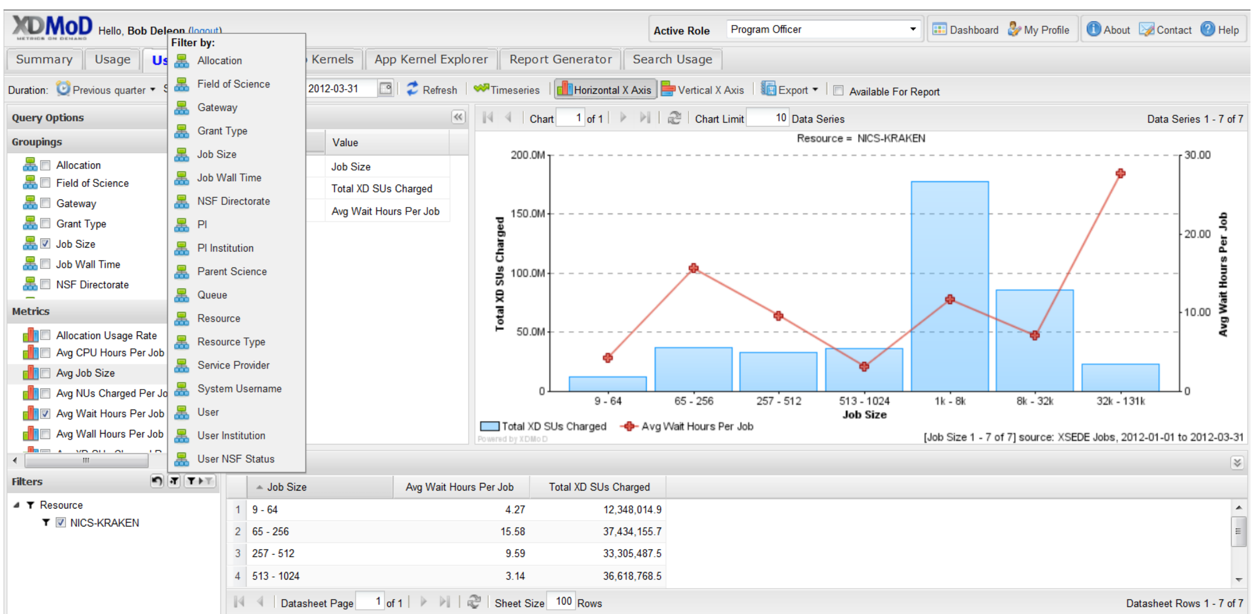


Figure 3. The XDMoD Usage Explorer. A plot of the total usage for the first quarter of 2012 is shown on the primary axis (left-hand axis) and the average wait hours per job is shown on the secondary axis (right-hand axis). The data have been filtered to show only data for NICS-Kraken.

Application Kernel Explorer:

The Applications Kernel Explorer has been similarly improved to make it more powerful and easier to use. Figure 4 shows a screen shot of the XDMoD App Kernel Explorer tab. The resource selected is edge.ccr.buffalo.edu (CCR's large production cluster). The application kernel is HPCC (the HPC challenge benchmark), specifically the subset of FFT floating point performance on 8, 16, 32, 64 and 128 cores. Plots such as this can be made by simply selecting a resource, an application kernel metric and the desired number of cores. As in the Usage Explorer interface, plots can be exported as a PNG or EPS file and the data can be exported as a CSV or XML file.

Figure 4. The Applications Kernel Usage Explorer. Shown is a plot of the 2012 first quarter data on CCR's large production cluster (ccr.edge.buffalo.edu) for the HPCC application kernel FFT floating point performance on 8,16, 32, 64 and 128 cores.

Expanded XDMoD Public View:

The XDMoD public view has been substantially expanded so that non-XSEDE users can get general information about XSEDE usage and view the capabilities of XDMoD itself. No user account is needed to access this view. Currently the XDMoDPublic View allows full access to the USAGE tab that allows the public access to the following (please note since XDMoD 2.0 now incorporates data from sources other than the XD central database, we have introduced the notion of REALMS to more easily identify the various data sources):

- JOBS REALM: 25 job usage metrics broken-down by user, PI, institution, resource, service provider, parent science, NSF directorate, or user NSF status
- ALLOCATIONS REALM: 5 allocation usage metrics broken down by PI, resource, NSF directorate or parent science
- ACCOUNTS REALM: Number of accounts opened, closed or created broken down by resource
- GRANTS REALM: Number of awards and award amount broken down by funding agency and parent science or NSF directorate
- PROPOSALS REALM: Number of funded projects and proposals broken down by NSF directorate or parent science.

It should be emphasized that the public does not have access to the user ids (for security reasons) or the unfunded allocation proposals. In addition, in the very near future, access to the applications kernel data and the usage explorer tool and other tabs will be added.

Figure 5 shows a screen shot of the new public view. The example chosen for Figure 5 is a plot of the total XSEDE usage for the first quarter of 2012. As shown on the chart tree on the left hand side in Figure 5, a large variety of metrics are available even to the public. In addition by clicking on the "Scope" selector box (upper left hand corner) data for individual service providers can be selected as shown in Figure 6 (note that TACC is selected in the Scope selector box).

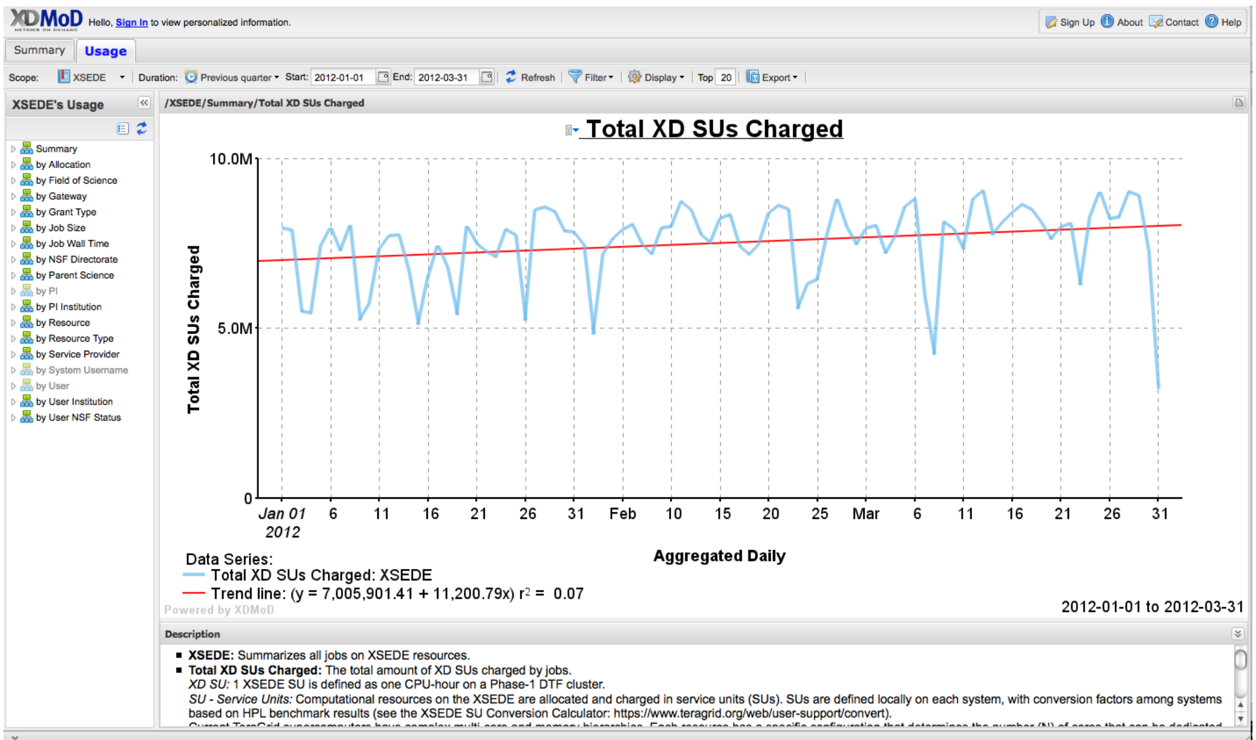


Figure 5. Screen shot from the Public view of XDMoD. Total XSEDE usage for the first quarter of 2012 is shown in the plot.

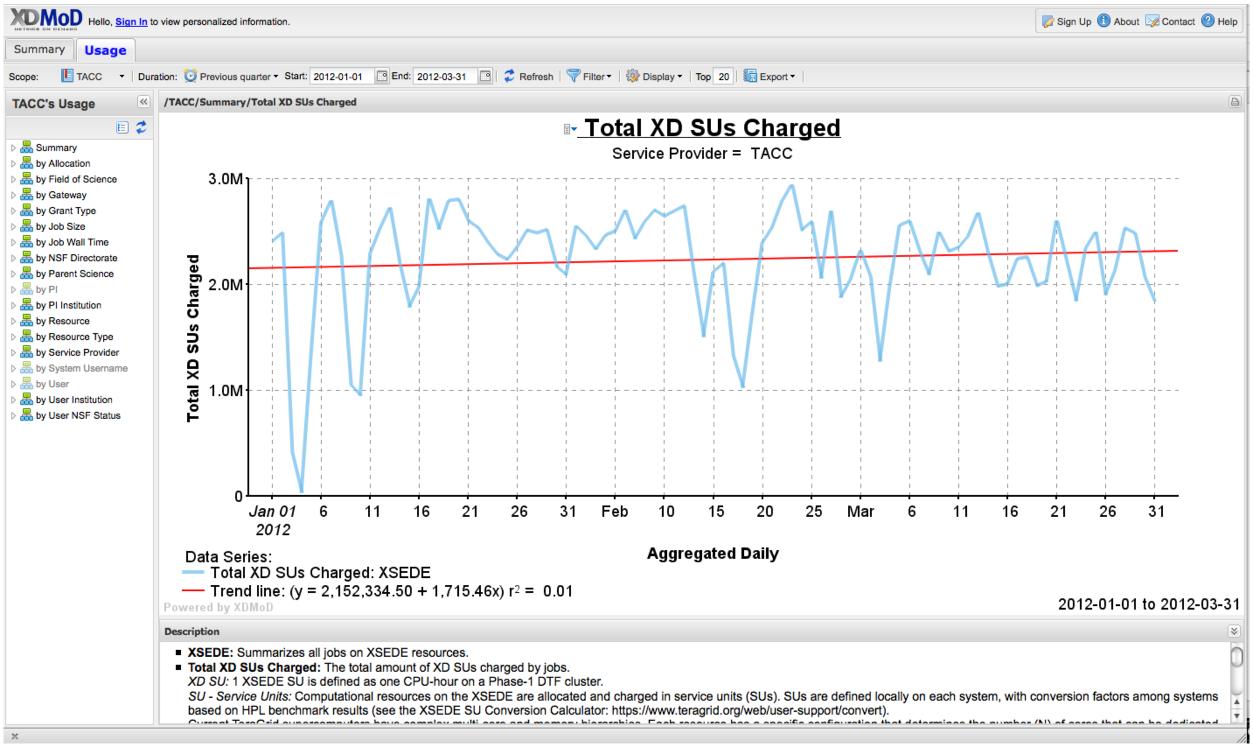


Figure 6. Screen shot from the Public view of XDMoD. Total TACC usage for the first quarter of 2012 is shown in the plot.

Custom Report Builder:

The Custom Report Builder has also undergone substantial improvement. It now has the ability to preview reports and has been modified to make it more powerful and easier to use. For example, as shown below in Figure 7, it is now possible to modify the data range of a chart within the report builder itself. The user can simply update previously selected charts by clicking on the edit chart time frame icon and select a new date range.

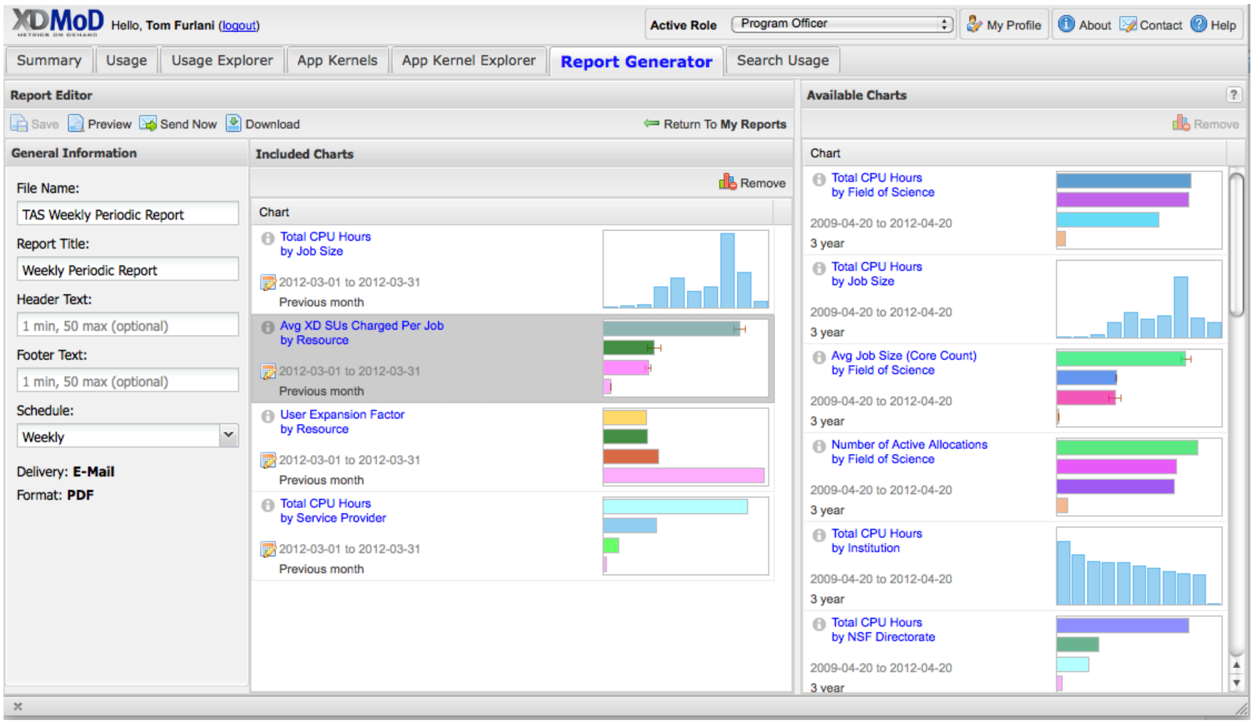


Figure 7. The XDMoD Custom Report Builder. Note that the date range of the plots can now be changed by clicking on the calendar icon for each plot directly in the custom report builder without having to regenerate them outside of the custom report builder interface

9.3.2 *Application Kernel Development:*

TAS allocation for Application Kernel Deployment:

Last quarter, in preparation for wider deployment of the application kernels throughout XSEDE resources, an allocation request (Jones – PI) was made by the TAS team for Blacklight, Forge, Kraken, Lonestar4, Ranger and Trestles. The TAS Application Kernel audit process is designed to run from a user's perspective and, accordingly, an XSEDE user allocation is required. During the past quarter this allocation request was approved. Work is currently in progress on deploying the application kernels on these machines. We anticipate having the application kernels running on a number of XSEDE resources by the end of the next reporting period.

Early Successes of the Application Kernels:

As the cases below illustrate, early implementation of application kernels have already proven invaluable – identifying underperforming and sporadically failing infrastructure that would have likely gone unnoticed, resulting in wasted CPU cycles on machines that are already oversubscribed as well as frustrated end users.

Figure 8 shows the execution of an application kernel during May 2011 based on NWChem, a widely used quantum chemistry program, that is a major application on the large production cluster at the Center for Computational Research at the University at Buffalo, SUNY (a chemistry faculty member is a contributor to NWChem source code development). While the behavior for 8 cores is as expected, calculations on 16, 32 and 64 cores showed widely sporadic behavior, with

some jobs failing out right and others taking as much as seven times longer to run. The source of degradation in performance was eventually traced to a software bug in the I/O stack of a commercial parallel file system, which was subsequently fixed by the vendor in the form of a software patch that is now part of their standard operating system release. It is important to note that this error was likely going on unnoticed by the administrators and user community for sometime and was only uncovered as a result of implementation of a suite of application kernels at CCR that automatically test the environment from the point of view of an NWChem user.

As a further indication of the utility of application kernels, consider Figure 9, which shows a performance increase of a factor of two in the MPI Tile IO benchmark (one of the I/O benchmarks in the current application kernel suite) after a system wide library upgrade from Intel MPI 4.0 to Intel MPI 4.0.1, which supports Panasas file system MPI-IO file hints. Since CCR employs a Panasas file system for its scratch space, this particular application kernel alerted center staff to rebuild scientific applications that can utilize MPI-IO file hints (specifically those for concurrent writes) to improve performance.

Figure 10 shows the unanticipated results brought to light by a periodically running application kernel based on the popular NAMD molecular dynamics package. The application kernel detected an approximate 25% degradation in the NAMD's baseline performance that was the unanticipated result of a routine system-wide upgrade of the application version (note change indicator symbols indicate when the source code for NAMD was upgraded). Once again, without application kernels periodically surveying this space, this change would have likely gone unnoticed.

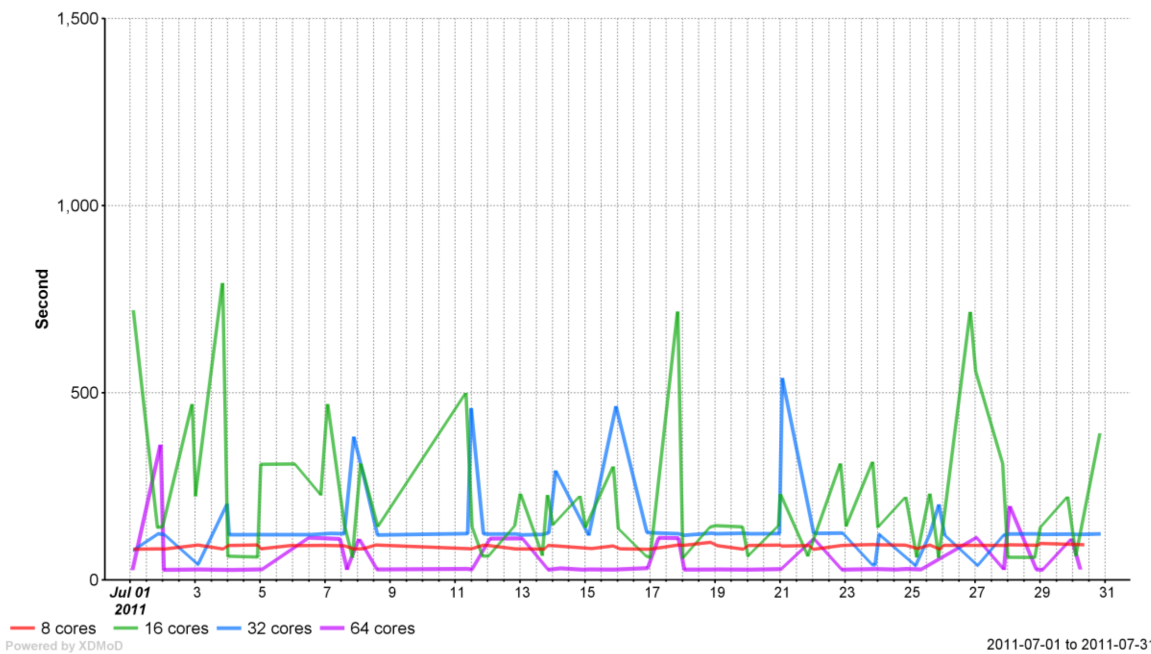


Figure 8. Plot of execution time of an NWChem based application kernel on 8, 16, 32, and 64 processors over a one month time period on CCR's production cluster. The execution time for each core count should be consistent. Behavior on 8 cores is as expected but 16, 32 and 64 core calculations show sporadic performance degradation. Source of degradation in performance was eventually traced to a bug in the I/O stack of a commercial parallel file system software, which was subsequently fixed by the vendor.

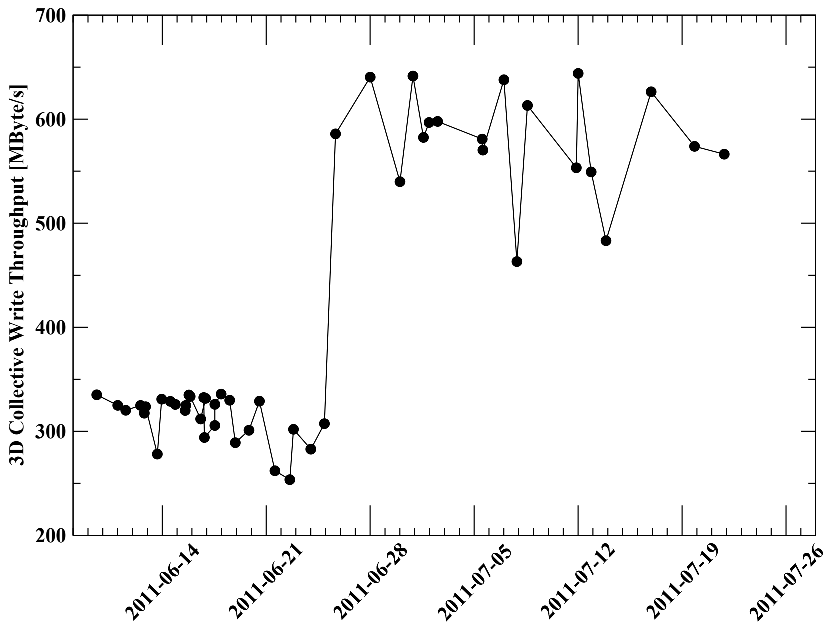


Figure 9. Plot of MPI-I/O concurrent write performance of MPI Tile IO I/O-based application kernel over time. Application kernel detects I/O performance increase of a factor of 2 for MPI Tile IO in library upgrade from Intel MPI 4.0 to Intel MPI 4.0.1 which supports Panasas file system MPI-IO file hints, in this case for concurrent write mode.

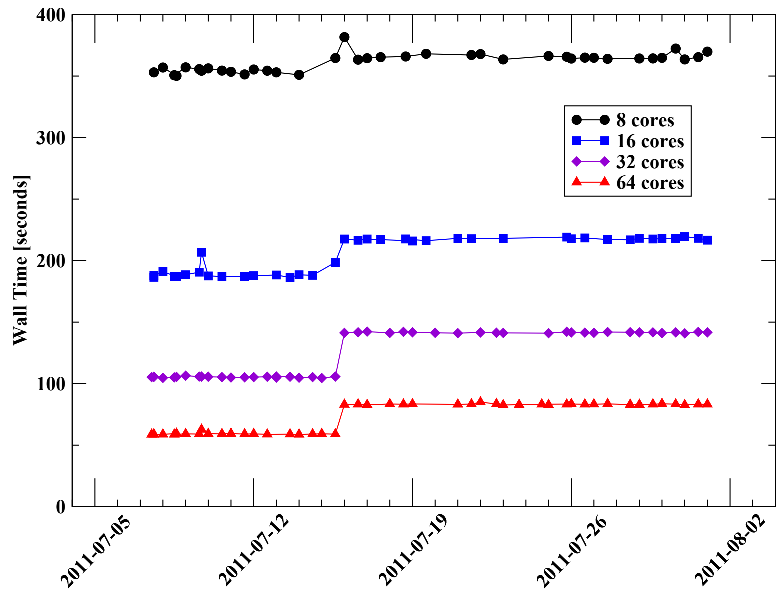


Figure 10. Plot of execution time of NAMD based application kernel on 8, 16, 32, and 64 cores over a one month time period. Jump in execution time at 2011-07-16 indicates when NAMD application underwent routine system-wide upgrade in the application version. Note the degradation in performance after the upgrade.

9.3.3 Incorporation of POPS data into XDMoD:

As part of the process of providing data for advanced cyberinfrastructure planning, a great deal of allocation information was mined from the POPS database. This effort was lead by Indiana University (Gregor von Laszewski and Fugang Wang) who developed software to access a portion of the POPS database. We are presently in the process of establishing the mechanisms to access all of the POPS database and automatically incorporate select portions data directly into the XDMoD data warehouse. This will provide access to information about the allocations process through the XDMoD portal.

9.3.4 Role-based Access to XDMoD:

It was decided to implement five distinct roles: NSF Program Officer, Center Director, Principal Investigator, User and Public. For the 1.6 release, all five roles are fully supported. It is anticipated that end-user roles of Public, User and PI will be supported by TAS and made available through the XSEDE user portal. The details of the interaction between the XDMoD portal and the XSEDE User Portal and their various functions in displaying the XDMoD data are currently being developed.

9.3.5 XDMoD Load Testing:

In preparation for opening XDMoD to all XSEDE users, a series of load testing measurements were made effectively scaling the number of concurrent users of the system from 1 to 200. We present a summary of the results in Figure 11 as well as more detailed results for the larger 100 and 200 user tests. Definitions of terms relating to the load test are as follows.

- "Concurrent users" is the number of users concurrently logged into XDMoD and using the tool. This doesn't necessarily mean that they are all doing the exact same thing at the same moment, but are actively utilizing the portal and exploring data.
- "Hit time" is the time from the point that a request is made by the client (e.g., web browser) until it receives a response from the server. This can be for a REST call, image, etc.
- An "agenda" is the set of pre-defined user actions (a mouse click, data entry, clicking a button or chart, etc.) that was created for the test to represent how we think a particular user role will use the portal. These are defined below in order of increasing data scope (and database load)
    - User - examining their group's usage with some comparison to XSEDE as a whole.
    - Center Director (CD) - examining their center's usage with some comparison to XSEDE as a whole.
    - Program Officer (PO) - examining trends in XSEDE over time as well as center and user data.
- A "test plan" is a mix of agendas representing a set of concurrent users.

Each test plan was a mixture of agendas representing the 3 roles defined above as follows:
- 3 Users (baseline) - 1 PO, 1 CD, 1 User agenda.
- 25 Users - 6 PO, 6 CD, 13 User agendas. All users began the test at the same time.
- 50 Users - 12 PO, 12 CD, 26 User agendas. All users began the test at the same time.
- 100 Users - 15 PO, 25 CD, 60 User agendas. The plan started with 5 PO, 12 CD, 25 User agendas and ramped to the full amount
- 200 Users - 30 PO, 50 CD, 120 Users. The plan started with 5 PO, 12 CD, 25 User agendas and ramped to the full amount.

The reason for the linear ramp-up for the 100 and 200 user load tests is that the server could not handle the full load of all users executing queries all at the same time (this essentially bypasses the database cache because the queries cannot complete before they are executed again). Note that an individual action in an agenda may generate many calls to the server. For example, the first time that a user visits the splash page this results in 107 requests sent to the server (REST calls, image requests, JavaScript file loads, etc.). Our data shows that in the first 20 seconds of the 200 User test 3457 individual requests were made to the server but over the course of the entire test this averaged out to 75 requests being serviced per second. Each test plan was run for 25 minutes and completed as many iterations as possible through its set of agendas during that time.

The test plans were executed by remote servers (remote to CCR) running through pre-recorded agendas and included a random 2-5 second wait between actions. Test data was cross-referenced with the MySQL slow query log as well as the output from the iostat and sar commands running during the tests to monitor CPU and I/O load. For reference, the completion time for each agenda in the baseline test (load of 1 of each agenda) is
- User - 3 minutes
- CD - 5.6 minutes
- PO - 4.7 minutes.

Overall test plan completion time is shown in Figure 11. This chart shows how long on average it took to complete a single instance of a test plan for each of the server loads. The 25, 50, and 100 user tests all show similar trends with an initial spike in the time needed to service each request

(hit time) and a tapering off of the hit time. The increased number of users results in a moderately decreased response time as expected. The 200 user test, however, shows a marked increase in the average completion times mostly due to the large queries being performed by the PO test. The large upward jumps in the graph at 11 min and 23 min are explained in Figure 13.

The Test Plan Completion Time vs. Load for 100 Users (Figure 12) and for 200 Users (Figure 13) show the number of agendas in each role ramping up over time (dashed lines) and the average time to complete each agenda (solid lines). The average time to completion for users remains fairly constant but the CD and PO roles see an initial spike in completion time and then a shorter time when the successive tests begin re-using their large cached queries from the database. Note that it takes roughly 10 minutes to complete an iteration through the PO test plan under the 100 user load. The 200 user test is clearly stressing the hardware, with the average time to complete the user tests increasing from 3.25 minutes to 4.7 minutes (a 45% increase). The first user test takes approximately 3 minutes to complete with 11 minutes for the first CD test and a whopping 23 minutes for the first PO test to complete. The PO role was only able to complete 20 iterations through the test plan vs 39 in the 100-user test. Note that the first completion of each role test corresponds to the jump in test completion time from Figure 11. We've found that the slowdown in the PO tests is largely due to the querying of datasets across all XSEDE over 3-5 years. This causes the database to use temporary tables on disk resulting in much disk contention.

Figure 14 shows the average hit time for the 200 User test. This shows the average time it takes for a request to be serviced for each of the roles during the test. Note that once the PO tests start making their large multi-year queries comparing data from all of XSEDE jobs, the service time skyrockets and begins to affect the other roles. Again, the main issue seems to be the temporary tables to service the large database joins. Also note that it's not likely that we will have this number of PO users (> 15) active at the same time in the near future.

In summary, the present XDMoD hardware and software can accommodate a reasonable mixture of up to 200 simultaneous users. The load testing also suggested strategies to improve XDMoD performance, especially for queries involving large data sets, that will be implemented in the future.

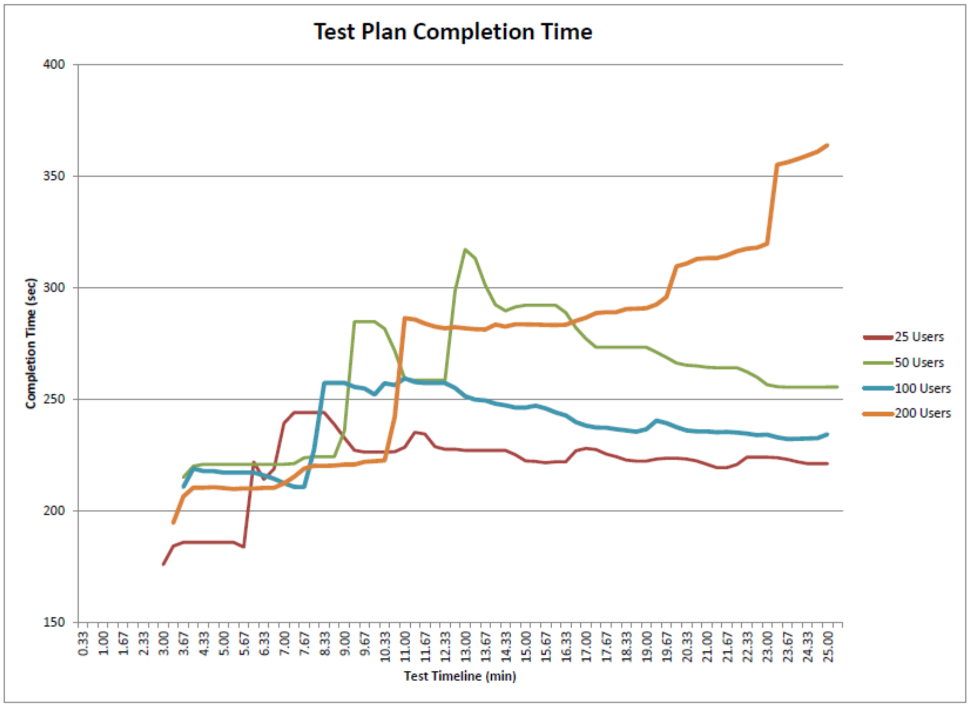


Figure 11. XDMoD load test summary results for all test plans.

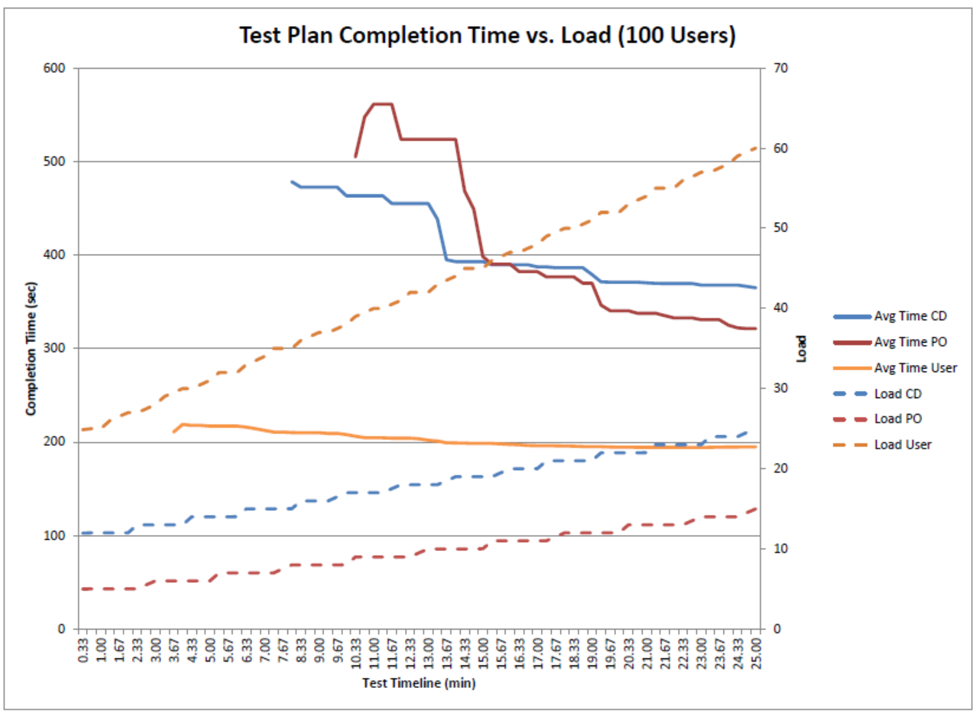


Figure 12. XDMoD 100 USER load test result showing the average agenda completion time versus user load (as described in the text).

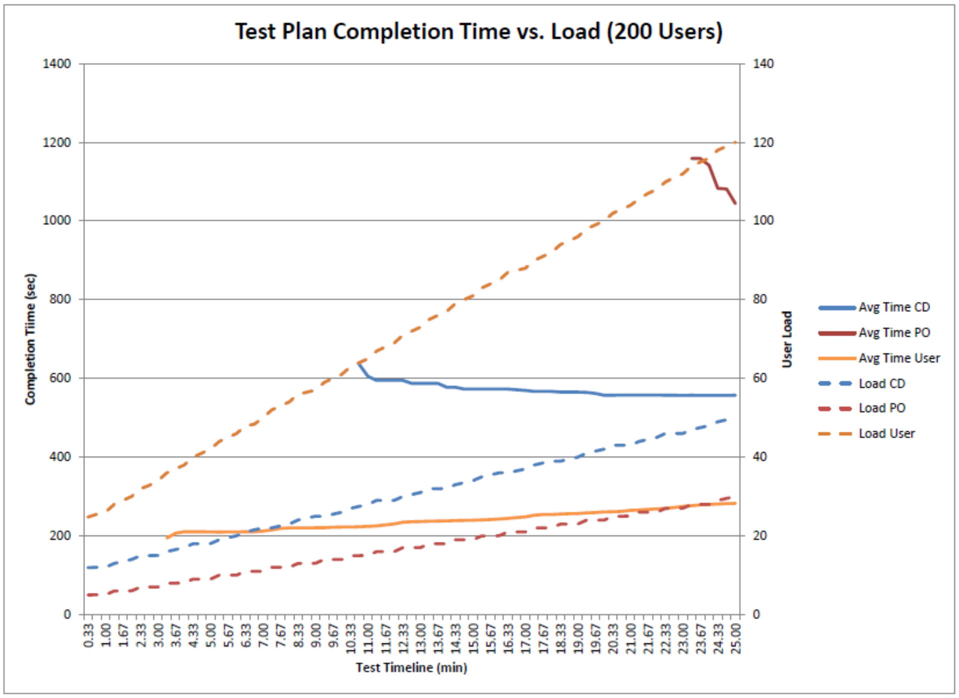


Figure 13. XDMoD 200 USER load test result showing the average agenda completion time versus user load (as described in the text).

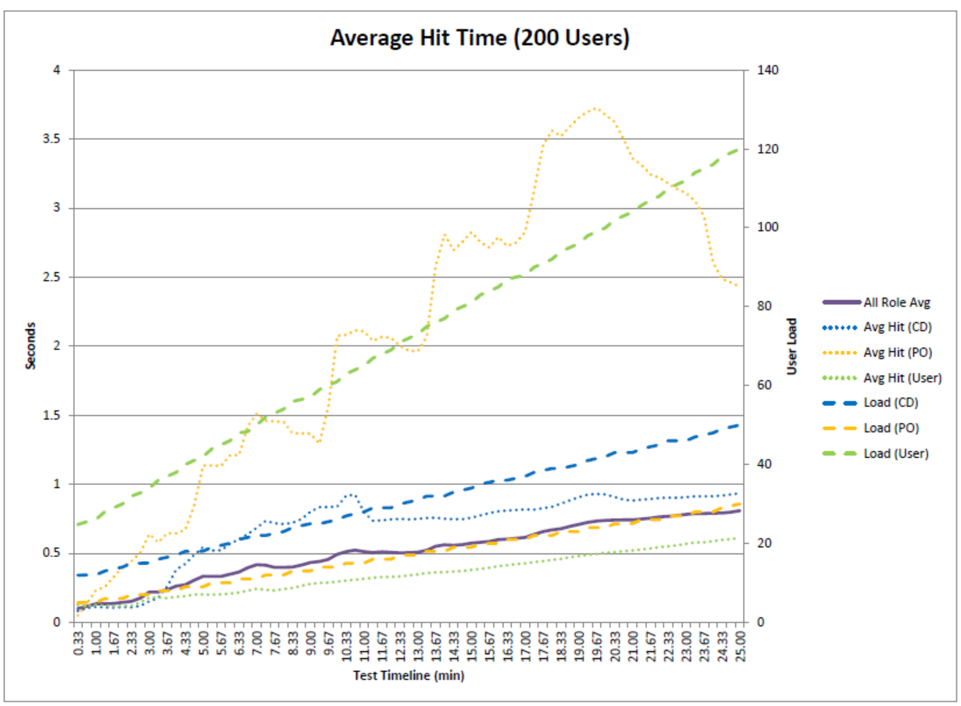


Figure 14. Average hit time versus user load for the 200 USER load test showing the effect of large (intensive) Program Officer queries.

9.3.6 XDMoD Usage:

XDMoD 1.6 (https://xdmod.ccr.buffalo.edu), which was released during the first quarter of 2012, is the first version providing access to all XSEDE users with an active XSEDE allocation as well as a public view with substantial functionality. Current XDMoD usage during the present reporting period is shown in Figure 15. While the usage is substantial, we anticipate increased usage in the future as more users become aware of its availability and capability.

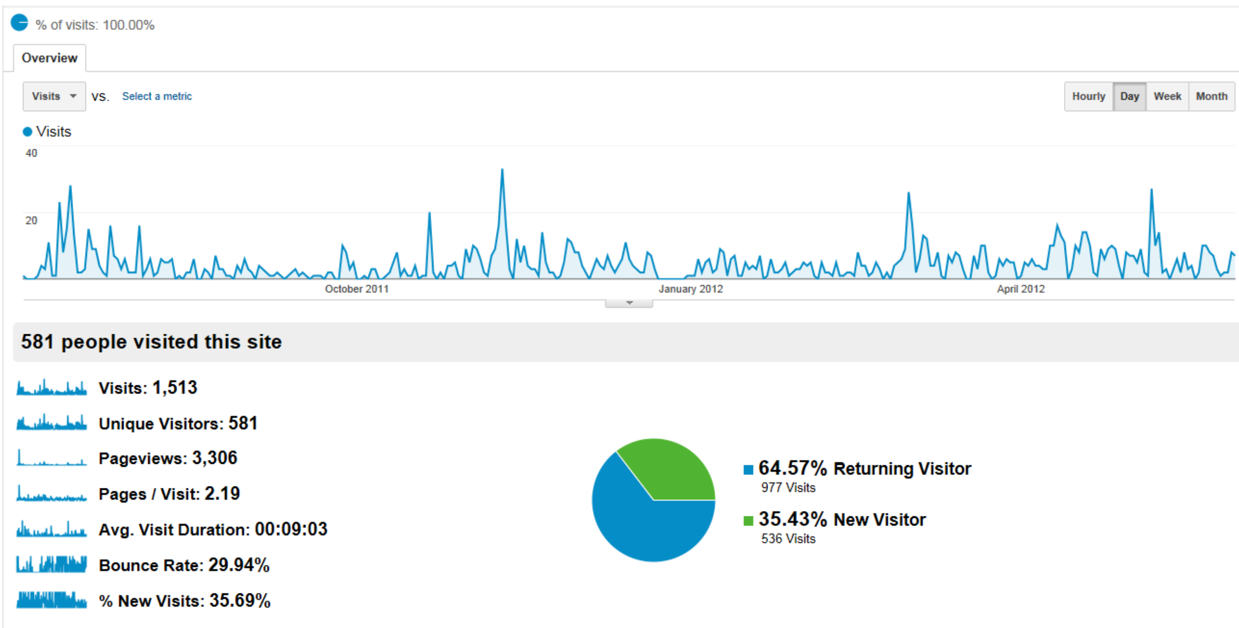


Figure 15. Google analytics overview of XDMoD usage for July 1, 2011 to May 31, 2012.

9.3.7 TeraGrid Central Database (TGcDB) Data Recording Recommendations:

In general, the TGCDB records are very good, however in the course of mirroring the TGcDB and constructing early versions of XDMoD we discovered that there are some areas where data is apparently missing and other areas where there is room for added or improved record storage. Previously we noted inconsistency in batch job usage reporting. Two notable examples are that job exit status is not reported and not all computational resources report usage by processor core as well as by node. Indeed, reporting is not always consistent for jobs on an individual resource. Other data missing or not routinely collected has been noted. Data is inconsistently reported across resources: the number of processors, number of nodes, number of processors per node, HPL Rmax and respective number of cores used in the benchmark and queues. Data is missing for some jobs: priority, memory usage, and a parent and child field of science.

Based upon TAS's examination of the existing records in the TGcDB, we have carefully formulated a series of recommendations for improved record-keeping. We continually refine these recommendations and the latest version is given below in Appendix 1.

9.3.8 XDMoD Reports for all XSEDE service providers:

TAS, through XDMoD, now provides all utilization plots for the XSEDE quarterly and annual reports for each service provider. The full set of plots for each resource is generated automatically using the custom report builder capability of XDMoD.

9.3.9 University of Indiana Subcontract:

IU tasks include: (a) Assist in the development of the XDMoD (XD Metrics on Demand) portal. (b) Collaboratively explore the POPS allocation database with the goal of supporting the incorporation of the POPS data into XDMoD (c) Participate in the development and deployment of application test kernels. (d) Interface with the Futuregrid and TIS (Technology Insertion Service) projects.

Incorporation of POPS data into XDMoD:

As part of the process of providing data for the ACI planning, a great deal of allocation information was mined from the POPS database, for example see Figure 16. This effort was led by Indiana University (Gregor von Laszewski and Fugang Wang) who developed software to access a portion of the POPS data base. We are presently in the process of establishing the mechanisms to incorporate access to all of the POPS data within XDMoD. This will provide access to all of the allocation data through the XDMoD portal.

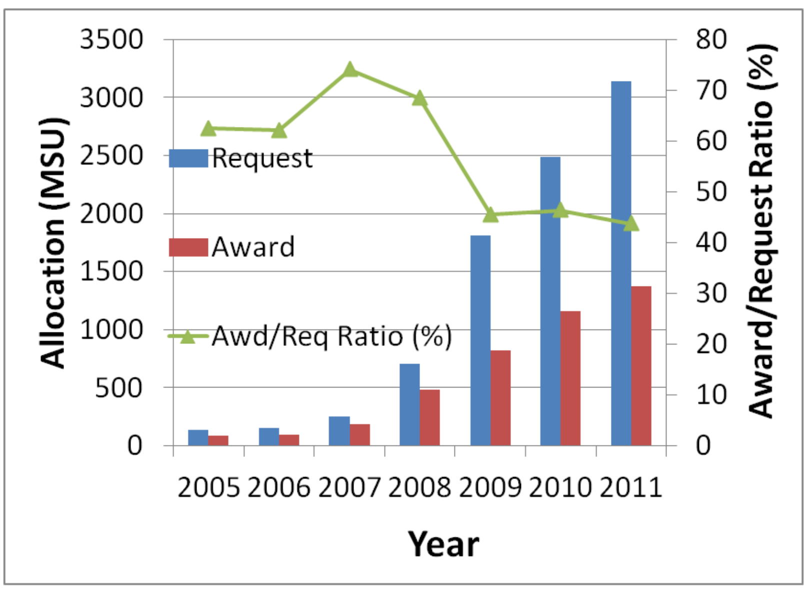


Figure 16. Historical allocation data mined from the POPS data base..

9.3.10 *University of Michigan Subcontract:*

With the change in personnel (loss of Ann Zimmerman's services) there was a period of inactivity on this subcontract. The usability assessment was done locally by UB personnel. At the very end of the reporting period, UM hired a new person to handle these duties (Mr. Jonathan Brier). Currently he has started the process of familiarizing himself with the XDMoD software.

9.4 Coordination with External Projects and Agencies

- Through Indiana University, we have conducted several initial outreach activities to identify opportunities for collaboration and leveraging. We have continued to work closely with the NSF funded FutureGrid to leverage their efforts on Single Sign-On and the integration with InCommon and to incorporate the XDMoD technology into FutureGrid. We will integrate these efforts with our own requirements and collaboratively explore this path. In addition we have also contacted the NSF funded OGCE project that was just recently funded in order to evaluate potential collaboration by leveraging gadgets and their containers. These are ongoing activities and will continue for the foreseeable future.
- We have been working collaboratively with Jim Browne and TACC to further develop TACC_Stats and to integrate it with XDMoD. This project is in the preliminary stages but steady progress is being made. Ultimately in the long term, some version of TACC_Stats will be run on the XSEDE resources and this rich data set will be available to XDMoD users.
- Preliminary planning is underway for developing a new program to adopt the CCR MoD technology to Open Science Grid.
- TAS worked closely with NSF in support of OCI's planning for advanced cyberinfrastructure (ACI) activities. One of the most important aspects of the TAS program and XDMoD in particular is providing the capability of supplying historical usage data which can serve as the basis for cyberinfrastructure planning. Based in part on this interaction, Barry Schneider produced the following paper: B.I. Schneider, *A Data History of TeraGrid/XSEDE Usage: Defining a Strategy for Advanced CyberInfrastructure (ACI)*, Office of Cyberinfrastructure, National Science Foundation, Arlington, Virginia 22230, Dated: March 11, 2012)

9.5 Meetings, Events Publications and Presentations

9.5.1 Meetings:

Weekly meetings of the TAS Working Group are held every Wednesday afternoon. CCR staff members attend in person with University of Michigan and Indiana University subcontractors attending by web and audio conference. The TAS working group includes:

- **University at Buffalo:** Dr. Thomas R. Furlani, PI, Dr. Matthew D. Jones (Co-PI and Technical Project Manager), Mr. Steven Gallo (lead software engineer), Mr. Andrew Bruno (scientific programmer), Dr. Charng-Da Lu (computational scientist), Mr. Ryan Gentner (scientific programmer), Mr. Amin Ghadersohi (scientific programmer), Dr. Abani Patra and Dr. Robert L. DeLeon (Project Manager)
- **Indiana University:** Dr. Gregor von Laszewski and Dr. Fugang Wang
- **University of Michigan:** Dr. Thomas Finholt, Mr. Jonathan Brier.

9.5.2 TG11 Activities:

TAS team members attended TG11 in Salt Lake City, Utah (July 2011).

TAS TG11 Attendees: Dr. Matthew Jones, Mr. Steven Gallo, Mr. Ryan Gentner, Mr. Amin Ghadersohi, Dr. Abani Patra, Dr. Charng-Da Lu, Dr. Robert L. DeLeon and Dr. Gregor von Laszewski (IU).

TAS TG11 accomplishments include: 1) interacting with the XD portal developers about the interface between the XD user portal and XDMoD, 2) meeting with TACC personnel to discuss adding additional capabilities to XDMoD such as systems log data, 3) meeting with SDSC personnel about adding metrics to XDMoD and developing Application Kernels, and 4) meeting with NSF sponsors to discuss XDMoD progress and future development.

TG11 presentations made by TAS:

- "Performance Metrics and Auditing Framework for High Performance Computer Systems" Thomas R. Furlani, **Matthew D. Jones**, Steven M. Gallo, Andrew E. Bruno, Charng-Da Lu, Amion Ghadersohi, Ryan J. Gentner, Abani Patra, Robert L. DeLeon, Gregor von Laszewski, Lizhe Wang and Ann Zimmerman.
- "Automatically Mining Program Build Information via Signature Matching" **Charng-Da Lu**, Matthew D. Jones and Thomas R. Furlani.

9.5.3 SC11 Activities:

TAS team members attended SC11 at Seattle WA (November 2011).

TAS SC11 Attendees: Dr. Thomas Furlani, Dr. Matthew Jones, Mr. Steven Gallo, Mr. Ryan Gentner, Mr. Amin Ghadersohi, Dr. Abani Patra, Dr. Charng-Da Lu, Dr. Robert L. DeLeon, Dr. Gregor von Laszewski (IU) and Dr. Fugang Wang.

TAS SC11 activities included:

1) Formal release of XDMoD 1.5 to authorized users. As of SC11 (Q4 2011) this list was still limited to selected program officers, center directors and XSEDE managers and staff members. As discussed above during Q1 of 2012 XDMoD was opened up to all XSEDE users.

2) Thomas Furlani, Matthew Jones and Steve Gallo, supported by the rest of the TAS team, made a Birds-of-a-feather presentation featuring an XDMoD demonstration and a following question and answer session. The BOF was well attended and led to a lively discussion.

9.5.4 TAS Technical Advisory Committee Meeting at SC11:

The TAS team presented a progress report to the TAS Technical Advisory Committee chaired by Abani Patra. TAS Technical Advisory Committee Contact Information:

1. Dr. Jerzy Bernholc (Atomistic QM, Physics, NC State)
   Email: bernholc@ncsu.edu; http://chips.ncsu.edu/~bernholc/

2. Dr. P.K. Yeung (CFD, Georgia Tech)
   Email: pk.yeung@ae.gatech.edu; http://soliton.ae.gatech.edu/people/pyeung/index.html
3. Dr. Martin Berzins (Chair Computer Science at Utah)
   Email: mb@cs.utah.edu ; http://www.cs.utah.edu/personnel/mb/
4. Dr. Craig Stewart (Executive Director, Pervasive Technology Institute, Indiana),
   Email: stewart@iu.edu; http://www.indiana.edu/~ovpit/bios/cstewart.html
5. Dr. Richard Brower (Lattice QCD, Boston University)
   Email: brower@bu.edu;; http://physics.bu.edu/~brower
6. Dr. Sara Graves (Director, Information Technology and Systems Center; Univ of Alabama, Huntsville), Email: agraves@itsc.uah.edu;; http://www.itsc.uah.edu
7. Dr. Abani Patra (Committee Chair, University at Buffalo), Email: abani@buffalo.edu

9.5.5 Publications:

The TAS team submitted the following paper to "Concurrency and Computation: Practice and Experience":

Performance Metrics and Auditing Framework Using Applications Kernels for High Performance Computer Systems

Thomas R. Furlani*, Matthew D. Jones*, Steven M. Gallo*, Andrew E. Bruno*, Charng-Da Lu*, Amin Ghadersohi*, Ryan J. Gentner*, Abani Patra*, Robert L. DeLeon*, Gregor von Laszewski+, Fugang Wang+ and Ann Zimmerman#

*Center for Computational Research, University at Buffalo, State University of New York, 701 Ellicott Street, Buffalo, NY 14203, +Pervasive Technology Institute, Indiana University, 2719 East 10th Street, Bloomington, IN 47408, #School of Information, University of Michigan, 105 S. State Street, Ann Arbor, MI 48109

It has been accepted in final form for publication and the on-line pre-publication article has been posted to the web (DOI: 10.1002/cpe.2871).

9.6 TAS Appendices

Appendix 1. TAS XDcDB Data Recording Recommendations

Based upon TAS's examination of the existing records in the XDcDB, we have carefully formulated a series of recommendations for improved record-keeping. These recommendations include improved job metrics, resource metrics and user metrics.

Job Metrics

The following are not being CONSISTENTLY *collected but are recommended for each job:*

- Number and type of cores charged (including GPUs)
- Exit status/Return code
- Disk usage (local)
- Memory usage
- Application usage, e.g., Automatic Library Tracking Database implemented by ORNL (c.f., Hadry, Fahey, and Jones, TG10 presentation, link)
- SU/NU consumed
- (future) Energy usage per system and per job (availability very limited until processors have counters for power utilization – 2012/2013?)
- ability to flag jobs as development or production (mostly done by queue now?)
- Eligible queue time – time from when the job is eligible to run until it starts executing.
- Policy expansion factor – (requested wall time + eligible queue time)/ (requested wall time)

Resource Metrics

The following resource data are either not being recorded or are being recorded inconsistently. We recommend both implementing new record keeping on these metrics and correcting historical records.

- Number of nodes
- Number of available nodes
- Number of processors per node
- Processor type and speed
- Processor memory (shared or dedicated)
- Number of GPU/Accelerator nodes
- Number of GPU/Accelerator processors per node
- GPU/Accelerator Processor type and speed
- GPU/Accelerator Processor memory
- Rmax and Rpeak values
- Local SU computation algorithm for each resource
- (future) Resource energy usage (idle and Rmax)

For storage resources

- (local disk usage is part of job reporting above)

- Shared filesystems description and retention policy
- Capacity
- Shared filesystems usage
- Keep information for storage jobs separately from compute jobs so that information collected can be specific to storage

Gateway Usage

- Consistent reporting of gateways (ADD)

User Metrics

- During the allocation process, list PI's other funding sources for the work that the allocation will be used for (POPS db)
- percent credit for parent sciences for multi-disciplinary projects (at allocation process?)

Database Mirroring

- Allow mirroring of the data from the primary to more than one secondary site

TGCDB **Schema Documentation**

- Maintain a schema diagram/description of key tables and relationships
    - acct.accounts, acct.allocations, acct.allocation_adjustments, acct.allocation_on_resources, acct.fields_of_science, acct.jobs, acct.organizations, acct.people, acct.principal_investigators, acct.requests, acct.resources, acct.system_accounts, acct.transactions
- Document formulas, triggers, and other logical elements. For example, NU/SU normalization and calculation process.

10 TAIS/Technology Insertion 1.8

10.1 Highlights

Technology Insertion continues to develop its processes for discovering, capturing, and evaluating technologies that are candidates to become an integral part of the XSEDE cyberinfrastructure.

This quarter saw efforts in the Technology Identification and Tracking and, the Technology Evaluation Process groups come to successful conclusions. The other groups continue to refine their efforts and start new directions in support of the TIS mission.

This year saw effort in all four subgroups in this direction. User needs are being canvased by the Usage, Satisfaction, and Impact Evaluation group. The XTED is being maintained and developed to meet the changing needs of the project. Finally, TIS evaluations were completed with the support of the Evaluation Laboratory, which is creating evaluation system resources to support ever broader technology evaluation products.

10.2 XD – TIS

XTED version 0.7 was released providing support for managing user requests. Evaluations have been started on workflow management systems.

10.2.1 Technology Identification and Tracking

Version 0.7 of XTED was released into production. Version 0.7 provides features supporting User Request feedback management. Development efforts during this quarter are focused on a minor release (v 0.7.1) which, in addition to several bug fixes, will involve the refactoring of code related to technology review tracking. This refactoring is needed to support tracking multiple review activities per related technology record, for instance reviews specific to a version of the underlying technology or only conducted in testing a particular feature of that technology. Version 0.7.1 is slated for QA testing and production deployment in Q3 2012.

Annual Summary for Technology Identification and Tracking

Development efforts related to the XTED database UI interface and backend management services involved an iterative development process with each iteration focusing on a subset of new features, existing feature enhancements and bug fixes. Features were selected by working with the rest of the TIS teams to identify additional functionality in support of their efforts. The primary new functionality developed during this period is iterated below.

Version 0.7 Use Cases

- Create User Request - Allows an Authenticated User to request a particular capability. This is a very generic 'catch-all' for soliciting feedback from the community and covers things such as documenting user forum feedback.
- Manage User Request - TIS Staff manages a User Request, ex. updating content.

- View Requests - The primary listing of all 'Request' related records viewable by TIS Staff

Version 0.6 Use Cases

- Create Technology Record Stub: TIS Staff can create a record stub for a Technology that is not already in XTED. They cannot edit the general details which must be provided by Primary Contact and/or Creator roles. This is required to enable the ability to track reviews on technologies not already entered into XTED.
- Manage Technology Review: TIS Staff manage the review status, reviewer and content for a given Technology.
- View Technology Reviews: The primary listing of all 'Reviewed' technologies in XTED. Additionally, TIS Staff can view a list of all Technology records with a status of 'Selected for Review' or 'Under Review'.

10.2.2 Technology Evaluation Laboratory

The Technology Evaluation Laboratory (TEL) is composed of a combination of reserved resources housed at TACC, the FutureGrid systems, and Service Provider (SP) systems. These systems represent the HPC computing environment found in XD and will facilitate the Technology Evaluation Process (TEP) conducted by the TIS team. The test system housed at TACC has been configured with 6- Dell 710 nodes and can be accessed via ssh at the public address: tis.futuregrid.tacc.utexas.edu.

Annual Summary for Technology Evaluation Laboratory

In the last quarter progress continues evaluating OpenStack to be used for virtualization in the TEL dedicated hardware to assist the Technology Evaluation Process (TEP). There was significant effort in an attempt to setup tests for the Genesis Federated File System (GFFS) however it was determined that GFFS was not ready for evaluation. The TEL team changed focus to Pegasus, a work flow manager, which was chosen for the next evaluation. We worked closely with the FutureGrid team to produce a set of Pegasus Virtual Machines (VMs) which could then be used for testing by the TEP group.

In the last year covering July 1, 2011 through June 30, 2012 significant effort has been put forth to support the Technology Evaluation Process (TEP) team. There is an ongoing effort to produce documentation necessary for the TEP team to perform evaluations on dedicated TEL hardware, XSEDE SP systems, and FutureGrid systems. Progress on documentation will continue as the process is refined to allow more individuals from the TEP team to perform evaluations. In addition to documentation, several administrative tasks were necessary to keep the dedicated nodes up to date with security patches for the latest virtualization technologies. First the dedicated nodes were integrated into FutureGrid's Alamo, which is also housed at TACC, to ease user account management and to leverage the existing NFS shared home file system for consistency across all nodes. After the integration we were able to upgrade the base operating system to

CentOS6.2 to support the latest in virtualization technologies including OpenStack. With regard to virtualization technologies, the team has extended high levels of effort to create and manage several base virtual images for both Nimbus and OpenStack to be utilized by the TEP team. The coming year will bring new challenges as virtualization technologies mature and documentation needs to be refined.

10.2.3 Technology Evaluation Process

Summary for Q2: We worked on the evaluations for Pegasus Workflow Management System and Unicore for Workflow management. The team faced a personnel challenge as one person from the National Institute for Computational Sciences (NICS) left the group.

The Fourth evaluation, Pegasus Workflow Management System, was in progress

Peter Enstrom is the lead engineer for the Pegasus Evaluation. As expected it has taken longer to develop the new requirements, use cases and test plans for Workflow Management systems because Pegasus is the most complicated application we have evaluated to date.

*The Fifth evaluation, Unicore for Workflow Management, was in progress. W*hile we have previous experience with this application, we are short an FTE as previously mentioned from NICS and reliant on the reuse of test cases from the Pegasus evaluation.

Goals

Our goal for the next quarter is to finish the fourth and fifth evaluations and to start the sixth and seven evaluations. Our replacement at NICS should also be identified.

Annual Summary for Technology Evaluation Process

We have completed 3 evaluations during the past year. We've added 3 more test workstations for various hardware/OS types and are using FutureGrid for larger scale testing. The team faced a personnel challenge as one person from NICS left the group, and their replacement came in and left during the year too. We expect to get a replacement for the next year and we intend to increase the number of evaluations we complete during the year.

10.2.4 Technology Insertion Support

The YNT team continues to work with the TIS evaluation teams to develop procedures and instrumentation in the evaluation process to support documenting deployment object components including validation tests and usage examples. Additional useful metadata from the evaluation process, such as time, expertise and effort required to perform each evaluation are being collected during evaluations to inform prospective implementers regarding deployment effort and expertise required. The YNT team continues to review existing deployment models in XSEDE and other CI projects, for guidance in establishing appropriate YNT procedures.

To facilitate transition of recommended and supported technologies in XSEDE, the YNT team continues to observe current deployment testing activities of XSEDE Software Development and Integration (SD&I) processes associated with XSEDE Configuration

Items for execution management services (EMS), global federated file system (GFFS), and Globus Online data movement. During this quarter, the focus of SD&I on these configuration items has been on acceptance testing by XSEDE Operations and Security groups prior to production deployment. The experience and information gathered will guide preparation of deployment materials and documentation in ways that can readily be utilized by XSEDE SDI Configuration Item teams, as well as XSEDE Operations integration test teams.

Annual summary for Technology Insertion Support

The Technology Insertion Support (YNT) team facilitates transition of supported and recommended technologies from evaluation to deployment by XSEDE service providers and users. YNT works with the evaluation teams to ensure that deployment, design, installation, configuration, maintenance, and usage methods for supported and recommended technologies are sufficiently documented, together with validation tests and usage examples.

YNT publications are intended for use by systems administration staff at XSEDE service providers to guide deployment of technologies for use with their resources. For technologies intended for end-user installation and use, YNT publications are intended for use by systems administration staff supporting end-users. Documentation for usage examples will be prepared for use by end-users directly.

During this project year, the YNT team worked with the TIS evaluation teams to develop procedures and instrumentation in the evaluation process. This included documenting deployment object components such as validation tests and usage examples. Additional useful metadata from the evaluation process, such as time, expertise and effort required to perform each evaluation were observed during evaluations to inform prospective implementers regarding deployment effort and expertise required. The YNT team reviewed deployment models in XSEDE and other CI projects, for guidance in establishing appropriate YNT procedures.

To facilitate transition of recommended and supported technologies in XSEDE, the YNT team observed current XSEDE Software Development and Integration (SD&I) processes associated with XSEDE Configuration Items for execution management services (EMS), global federated file system (GFFS), and Globus Online data movement. These processes include preparation of software services, software clients and documentation, testing of prototype deployments based on expected usage scenarios within XSEDE, security reviews of software services, and acceptance testing by SD&I Operations staff. The experience and information gathered helps to guide preparation of TIS technology deployment materials and documentation in ways that can readily be utilized by XSEDE SD&I Configuration Item teams, as well as XSEDE Operations integration test teams. This will support closer integration of YNT and XSEDE SD&I activities as we move forward into PY 2012-2013.

10.2.5 Usage, Satisfaction, and Impact Evaluation

Led by the TIS user advocate, the Usage, Satisfaction and Impact Evaluation team is

charged with engaging the user community in the evaluation and testing process, by various methods including focus group teleconferences aimed at finding out what functional requirements are not being met, and with feeding these requirements into the TIS identification and tracking, evaluation and testing activities.

Over the past year, which marked the transition from working with TeraGrid to working with XD, we have primarily been leveraging user engagement opportunities provided by the Campus Champions, other TEOS, and ECSS Novel and Innovative Projects programs. This has resulted in two technologies being proposed and approved for evaluation (Pegasus WMS and MEDICI). We have requested that XTED be expanded to also allow for the entry of user requests: this was enabled in XTED version 0.7. This allowed us to enter two user requests: for improved file transfer tools, and for MEDICI.

We work with the YNT team to develop an evaluation priority metric to guide selection and to establish priorities among technology evaluation candidates (see 10.2.4 above), stressing the need for the user requirements to be given appropriate weight in this metric. Now that XTED contains both technology nominations and user requests, this process is being implemented.

Over the past few months, we have become very active in planning the alignment of TIS activities with those of all other XD services. We now participate in the monthly "forum" calls organized by the User Engagement team, and are monitoring user requests sent via the help@xsede.org and feedback@xsede.org addresses for items that may be relevant to TIS. We have also established a collaborative process with the Systems Engineering team.

Annual Summary for Usage, Satisfaction, and Impact Evaluation

Led by the TIS user advocate, the Usage, Satisfaction and Impact Evaluation team is charged with engaging the user community in the evaluation and testing process, by various methods including focus group teleconferences aimed at finding out what functional requirements are not being met, and with feeding these requirements into the TIS identification and tracking, evaluation and testing activities.

Over the past year, which marked the transition from working with TeraGrid to working with XD, we have primarily been leveraging user engagement opportunities provided by the Campus Champions, other TEOS, and ECSS Novel and Innovative Projects programs. This has resulted in two technologies being proposed and approved for evaluation (Pegasus WMS and MEDICI). We have requested that XTED be expanded to also allow for the entry of user requests: this was enabled in XTED version 0.7. This allowed us to enter two user requests: for improved file transfer tools, and for MEDICI.

We work with the YNT team to develop an evaluation priority metric to guide selection and to establish priorities among technology evaluation candidates (see 10.2.4 above), stressing the need for the user requirements to be given appropriate weight in this metric. Now that XTED contains both technology nominations and user requests, this process is

being implemented.

We participate in the monthly “forum” calls organized by the User Engagement team, and are monitoring user requests sent via the help@xsede.org and feedback@xsede.org addresses for items that may be relevant to TIS. We have also established a collaborative process with the Systems Engineering team.

11 ExTENCI

XSEDE and Open Science Grid have been working together to support science that can make use of both cyberinfrastructure platforms.

11.1 Overview

The Extending Science Through Enhanced National Cyberinfrastructure (ExTENCI) Project is a joint Open Science Grid (OSG) and TeraGrid project, funded by OCI. The PIs are Paul Avery (U. Florida) and Ralph Roskies (PSC). The planned two year project began in August 2010 and has received a no-cost extension through July 2013.

The goal of ExTENCI is to develop and provide production quality enhancements to the national cyberinfrastructure that will enable specific science applications to more easily use both OSG and XSEDE/TeraGrid or broaden access to a capability to both XSEDE/TeraGrid and OSG users.

ExTENCI has four primary areas of work, each of which is discussed below.

11.1.1 ExTENCI – Distributed File System (Lustre-WAN) (Paul Avery and Ralph Roskies/J Ray Scott)

Distributed File System updated software to the latest version of Lustre (2.1) and base OS (CentOS 5.7 and Scientific Linux 6.1). FSU and FIU were incorporated into the Lustre network and a remote Lustre Object Storage Server (OST) was added to make the first instance of a distributed OST. The CMS, ATLAS, and LQCD applications were tested and the CernVM File System was ported to the PSC distribution stack. The supported VM images now include all major VM systems (XEN, VirtualBox, VMWare, and KVM). Performance testing and benchmarking were completed across the network. Four papers on this work were presented at conferences.

11.1.2 ExTENCI – Virtual Machines (Carol Song, Sebastian Goasguen)

Virtual Machines has created ‘Elastic IP’ and ‘Security Group’ services for leasing of public IP addresses that are in production in Clemson’s OneCloud and has been tested by FutureGrid and DESY in Germany. A Cloud Dashboard has been developed at Purdue that enables users to start, interact with, and stop VMs via a browser. A study was completed and presented on the utility of using Cloud resources for NanoHUB. Work has begun on deploying VMs to FutureGrid’s Eucalyptus, Nimbus, and OpenStack services. Three papers on this work were presented at conferences.

11.1.3 ExTENCI – Workflow & Client Tools (Michael Wilde)

Workflow & Client Tools continued work in earthquake hazard prediction (from the SCEC project) and branched out to add several new science applications in addition to the protein structure work in the project’s first year. For SCEC, we now have a single script to run the SCEC Cybershake processing on both OSG and XSEDE and have developed a new data management interface the data handling/management capabilities of Swift. The SCEC Broadband code has been installed, problems resolved, and results verified. A version of parallel Broadband code in VMs was run on Futuregrid as a precursor to running the “hanging wall” benchmark in one day. Focus of the Swift workflow engine work shifted this year to supporting science applications in protein science, theoretical chemistry, earth systems science, and sociology of science. These applications resulted in two published papers and used over 1.1 million CPU hours on OSG and XSEDE. Two additional papers were presented at conferences and three ExTENCI status updates were presented at the Open Grid Forum quarterly meetings.

11.1.4 ExTENCI – Job Submission Paradigms (Shantenu Jha)

Job Submission Paradigms has completed, hardened and deployed the SAGA-Condor adaptor and tested it on OSG. The project has integrated the SAGA-Condor adaptor within the BigJob Pilot-Job mechanism and integrated the Condor enabled BigJob within the DARE framework of the Cactus Gateway. Support of Cactus-based Numerical Relativity and Computational Biology application groups has begun. Five papers were presented at conferences.

12 XD Service Provider Reports

2012 Annual Report: August 1, 2011, through July 31, 2012

XD Service Provider Forum Leadership

Carol Song, Chair	Purdue University (PU)
David Y. Hancock, Vice Chair	Indiana University (IU)

XD Service Provider Principal Investigators

Sean Ahern (UT-NICS)	U Tennessee – National Institute for Computational Science
Jay Boisseau	Texas Advanced Computing Center (TACC)
Geoffrey Fox	Indiana University FutureGrid
Kelly Gaither	Texas Advanced Computing Center (TACC)
David Y. Hancock	Indiana University (IU)
Michael Levine	Pittsburgh Supercomputing Center (PSC)
Miron Livny	University of Wisconsin, Madison (Open Science Grid)
Richard Moore	San Diego Supercomputer Center (SDSC)
Michael Norman	San Diego Supercomputer Center (SDSC)
Carol Song	Purdue University (PU)
Craig Stewart	Indiana University (IU)
John Towns	National Center for Supercomputing Applications (NCSA)
Jeffery Vetter	Georgia Tech
Thomas Zacharia (UT-NICS)	U Tennessee – National Institute for Computational Science

12.1 Overview

This section summarizes the main accomplishments of the Service Provider Forum (SPF) during the first project year.

1. Defined the charter, bylaws, and process to formally organize the SP Forum

The foremost task of the SP Forum was to define and formally organize the forum.

An interim Service Provider Forum was formed to develop the SP Forum charter document and to initiate the process of formally establishing the forum. The membership of the Interim Forum was selected based on the anticipation that they would be Level 1 or Level 2 SPs (see below) when the SPF was formally instantiated, and that they had experience or important points of view in the type of governance represented by XSEDE and the SPF. The members of the Interim Service Provider's Forum were:

Name	Institution	Resource/Service
Carol Song, *Interim Chair*	Purdue	Steele/Condor/Wispy
Sean Ahern	Univ of Tennessee-Knoxville/NICS	RDAV/Nautilus
Jay Boisseau	Univ of Texas-Austin/TACC	Ranger/Spur/Ranch
Jay Boisseau	Univ of Texas-Austin/TACC	Lonestar
Mark Fahey	Univ of Tennessee-Knoxville/NICS	Kraken
Geoffrey Fox	Indiana Univ	FutureGrid
Kelly Gaither	Univ of Texas-Austin/TACC	Longhorn
Craig Stewart	Indiana Univ	Quarry/HPSS/Data Capactiro/RockHopper
Mike Levine	PSC	Blacklight
Miron Livny	Univ of Wisconsin	OSG
Honggao Liu	LSU	QueenBee
Rich Loft	NCAR	Frost
Richard Moore	UCSD/SDSC	Trestles
Michael Norman	UCSD/SDSC	Gordon
John Towns	Univ of Illinois/NCSA	Forge/Ember/MSS
Jeffrey Vetter	Georgia Tech	Keeneland

The forum members had extensive discussions regarding the nature of SPF, i.e., whether it would mainly serve as a platform to interact with the XSEDE project or it would be open to a broader community of service providers beyond the XSEDE project. The forum ultimately decided to establish a forum that is open to any service provider that sees value in being part of this community. "*While the Service Provider Forum is chartered to fulfill certain XSEDE-specific requirements and goals and while SPs may have Cooperate Agreements (CA) with the NSF that mandate their interaction with XSEDE as members of the XSEDE Federation, the forum will be open to any Service Provider that sees value in exchanging experiences, sharing best practices and in leveraging the common tools and capabilities offered by XSEDE and thus become an active element in the national Cyberinfrastructure ecosystem*." (SP Forum Charter). This vision of inclusiveness has the potential to positively impact cyberinfrastructure development on the campuses and nationwide.

The Charter document defines the forum's governance and bylaws, and the process for a service provider to join the SP Forum. It is available as a public document at https://www.xsede.org/web/guest/project-documents. Other documents the SPF has approved in the past are the *Service Provider Forum Membership Application form*, and the *Application*

template for Service Providers to join the XSEDE Federation. These documents were ratified by unanimous vote of all forum members in March 2012.

2. Completed the organization of the forum by electing SPF officials and representatives to XSEDE Advisory Board (XAB)

The SPF has formally organized and elected its officials according to the process defined in the SP Forum Charter in April 2012, with the following result:

Position	Name	Institution
SPF Chair	Carol Song	Purdue University
SPF Vice Chair	David Hancock	Indiana University
Representative to XAB	Dick Glassbrook	Georgia Tech
Representative to XAB	Miron Livny	University of Wisconsin, Madison

The current interim SPF members will follow the process for service providers to formally join the SP Forum as defined in the SPF Charter document.

3. Worked with the XSEDE project to define and formalize the storage allocations policy and process

In response to the increasing need for data storage on the XSEDE compute resources and the increasing diversity for data services, e.g., data sharing, XSEDE began the process of defining storage allocations procedures and presenting guidelines and helpful information about obtaining storage resources allocations. Traditionally, two primary storage tiers are provided: direct-attached to a compute system, and archival systems with higher capacity but lower transfer rates, available to all of a site's compute platforms. Direct attached storage is partitioned as scratch storage available to all users but purged regularly, and persistent storage with user quotas, whereas archival storage was usually provided without quotas. Under XSEDE, storage is a supplemental component of most resources, enabling computation and visualization services. Some current XSEDE users have requested to make simulation data available to the broader research community and it is anticipated that this usage mode may become one of the primary roles of storage resources.

The SPF worked with the XSEDE project, especially the User Services group, on initial concepts regarding the tiers of the storage resources and how they should be allocated based on their interactions with users. SPs provided feedback on the proposed implementation plan, and worked on the storage allocation plan through discussions, edits and comments. The SPF also surveyed the SPs about the storage systems that will be allocated through XSEDE (survey result can be found at the SPF wiki under "Other documents"). The survey result is a collection of system characteristics and SP plans regarding a few critical implementation issues, with the following questions:

- Provider
- System
- System Type
- Allocation independent of compute project?
- Will this be an XSEDE resource?
- Total Capacity for XSEDE
- Date of retirement (known or estimated)
- Max storage space w/ automatic approval
- Duration of storage with automatic approval
- Max storage reviewed and approved by SP

- Duration of storage approved by SP
- Max storage space for XRAC request
- Duration of storage for XRAC request
- Retention (duration) beyond allocation expiration date
- When will this system be ready for storage allocation to XSEDE users?
- Does this system currently support project quota management? If not, what is estimated time/effort to implement the support?
- Data migration methods with other SP storage resources

4. Accepted the first new forum member, Rutgers University, as a Level 3 Service Provider.

The Rutgers Discovery Informatics Institute (RDI^2) has applied to join the SP Forum as a Level 3 Service Provider. RDI^2 is an advanced computation institute that provides academic and industry researchers with the resources, skills, and expert support necessary to leverage advanced computation technology for computational and data-enabled science and engineering. The institute has an IBM Blue Gene®/P supercomputer, "Excalibur", installed in March 2012, and additional racks of IBM's Blue Gene®/Q, a Linux cluster, visualization resources to be acquired in 2012, all located in an expanded facility on the university's Livingston Campus in Piscataway, New Jersey. Its initial interaction with XSEDE will be through co-branding as stated in the Rutgers' application. The SPF voted to approve Rutgers membership in August 2012. Rutgers has become the newest member and the first Level 3 SP of the forum.

5. Provide input and feedback to the XSEDE project

On an on-going basis, the SPF provides input and feedback to the XSEDE project in various forms. For example, the SPF provided comments, sometimes edits, to a number of XSEDE documents, including the XSEDE award letter to NSF Program Directors, the baseline software and security documents. The SPF also provided candidate nominations for the XSEDE Advisory Board (XAB) and the User Advisory Committee (UAC). The SPF provided input regarding the values of being an XSEDE service provider to the XSEDE annual review. Dave Hancock represented the SPF at the annual review in June 2012. The SPF representative (Miron Livny) participated in the XAB meeting via the phone in April 2012.

6. Other accomplishments

- Defined quarterly reporting format and content so that all SPs use a common template for the XSEDE quarterly reports
- Created the SP Forum wiki for meeting agenda and minutes, document and information sharing (https://www.xsede.org/group/sp-forum/wiki).
- Weekly or biweekly conference calls, with annual face-to-face meeting held during the XSEDE12 conference in Chicago, July 2012.

Looking forward, the SP forum will complete the work on the storage allocation process and implementation. The forum will work with XSEDE leads in various areas to better understand the project plans and the potential impact. The SPF members will work with the STCI SUPREMM project team to study usage and benefits on the XSEDE resources.

The following sections include the annual executive summaries of the Level 1 Service Providers in the Forum.

12.2 FutureGrid – SP Forum Annual Report Summary

FutureGrid is a national-scale Grid, Cloud and HPC computing test-bed service of modest size that includes a number of computational resources at five distributed locations. FutureGrid systems total 4400 cores and provide totally reconfigurable generic nodes with some with special hardware or software. The FutureGrid network is unique and can lend itself to a multitude of experiments specifically for evaluating middleware technologies and experiment management services. This network can be dedicated to conduct experiments in isolation, using a network impairment device for introducing network latency, jitter, loss, and errors to network traffic within FutureGrid. All network links within FutureGrid are dedicated (10GbE lines for all but to Florida, which is 1GbE), except the link to TACC. The significant number of distinct systems within FutureGrid provides a heterogeneous distributed architecture and are connected by high-bandwidth network links supporting distributed system research. One important feature to note is that some systems can be dynamically provisioned, e.g. these systems can be reconfigured when needed by special software described below, with proper access control by users and administrators.

We had an active year with 105 submitted projects covering our major usage areas: Computer Science and Middleware(59%), Computer Systems Evaluation (29%), Domain Science Applications (26%) and Training Education and Outreach (10%) with some projects having multiple tags and so totals are >100%. Our registered users at end of reporting period was over 1150. In this reporting year, we added significant new support and outreach staff to address identified needs. This cut back our plans to expand hardware resources although we also added the small Bravo and Delta clusters supporting large memory, large disk and GPU applications.

We greatly expanded our support of educational applications as the flexible interactive usage mode of FutureGrid is well suited to education. This was highlighted in an XSEDE all hands meeting AHM paper and BOF. One special highlight was our Science Cloud Summer School with over 170 participants at 10 sites which started July 30 2012. This year also saw an expansion in the use of FutureGrid for testing of XSEDE software both as a test operational Grid and as an interoperability endpoint for SAGA, Genesis and Unicore. International collaborations included the European Middleware Initiative EMI whose software is deployed on FutureGrid with an online tutorial. The cloud projects focused on new experimental applications and new technologies. Here MapReduce is very important as highlighted in substantial interest shown in an XSEDE AHM BOF in this area. FutureGrid offers both Hadoop and Twister as available Platforms as a Service MapReduce systems.

FutureGrid made major progress in the “Cloud Infrastructure as a Service” area with the supported deployment of latest releases of Nimbus, OpenStack and Eucalyptus. We were first academic deployment of the commercial Eucalyptus 3 and this release resolved the major difficulties with the previous release. More importantly we now have a full operational ability called FutureGrid RAIN to dynamically deploy both cloud and bare metal environments on demand and used it operationally for first time at the end of this reporting period. This software is still dependent on some components like the IBM xCAT and Moab that we want to replace (augment) with more general open source systems. Further we are still working on some performance issues. However this demonstrates a new critical feature that “Computing Testbeds as a service” like FutureGrid must address. Our RAIN system integrates an image repository that can store images as templates so they can be deployed on bare metal or different commercial or academic virtual machine management environments. This provides an interoperability capability that will allow applications tested on FutureGrid to be moved to other clouds such as Amazon.

FutureGrid has interacted with XSEDE on integrating accounting approaches, EOT (for example Summer School) and software testing.

12.3 Georgia Tech – SP Forum Annual Report Summary

12.3.1 Executive Summary

Keeneland has participated in the XSEDE Service Provider Forum, but did not begin to provide production services until July 1, 2012 with the Keeneland Initial Delivery System (KIDS). KIDS was upgraded from M2070 GPU technology to M2090 GPUs to provide a performance boost for users and current technology to continue software and application development. The Keeneland Full Scale System (KFS) was installed in July 2012 and is expected to become the production resource in late September or early October 2012.

KIDS – 120-node cluster composed of HP SL390 nodes, each with two Westmere CPUs, three NVIDIA M2090 GPUs, 24 GB host memory, and a Mellanox QDR IB interconnect through a QLogic QDR IB enterprise switch.

KFS – 264-node cluster composed of HP SL250 nodes, each with two Sandy Bridge CPUs, three NVIDIA M2090 GPUs, 32 GB host memory, and a Mellanox FDR IB interconnect through a Mellanox FDR IB enterprise switch.

12.3.2 Science Highlights

The Keeneland Initial Delivery System (KIDS) only became available as an XSEDE production resource in July 2012. Prior to that time it was used primarily for software and application development. At a high level the research areas utilizing KIDS include computer science and computational research, astronomical sciences, atmospheric sciences, behavioral and neural sciences, biological and critical systems, chemistry, design and manufacturing systems, Earth sciences, materials research, mathematical sciences, mechanical and structural systems, molecular biosciences, physics, cross-disciplinary activities, and education/training. Both the number of users and the breadth of science fields is well beyond what was initially envisioned for KIDS. The primary users of KIDS are involved in molecular biosciences, physics, and chemistry.

The Keeneland Project hired a former post-doc from Jack Dongarra's group to provide advanced application support. The first user project was very successful. The BEAST/BEAGLE Phylogenetics code was upgraded to use 3 GPUs per node to use MPI to scale the workload. The user can run the BEAST/BEAGLE application using the same command they used before. That command can hide the fact that we are starting MPI jobs on separate nodes. The input file from the user will also stay the same. Jobs that would have required days to months to complete now run in minutes to hours.

12.3.3 Education, Outreach, and Training Activities

Video from the February 2012 workshop at Georgia Tech were placed online for access by new Keeneland users and the larger computational community. No demographics on access are available.

12.4 Indiana University – SP Forum Annual Report Summary

12.4.1 Executive Summary

During XSEDE Program Year 1, Indiana University Pervasive Technology Institute – Research Technologies made a transition from being a TeraGrid Resource Provider to being a Service Provider with one Level 2 system (Quarry), one system proposed as a Level 3 system (Rockhopper), one system being phased out of XSEDE availability (Data Capacitor), and one system that may or may not yet be available on a continuing basis for allocation via XEDE. IU's involvement in XSEDE providing the Quarry Virtual Machine service and access to Penguin-On-Demand via the Rockhopper cluster are consistent with IU's general history in TeraGrid. Our expertise makes us well able to implement novel and new systems.

Quarry is an IBM Intel-based cluster is a continuation and expansion of services available under TeraGrid. IU was the first Resource Provider within TeraGrid to offer a VM service. Quarry has become the premier hosting site for VMs that deliver science gateways for XSEDE, as well as serving as a VM hosting site for many operational XSEDE activities. Quarry is very actively used as a host for Virtual Machines and Science Gateways. Some of the most significant science gateways supported on Quarry include:

ParamChem: Dr. Sudhakar Pamidighantam, National Center for Supercomputer Applications
Galaxy: Dr. Louise Laurent, the Scripps Research Institute
Dark Energy Survey: Prof. August Evrard, University of Michigan
Einstein Genome Gateway: Dr. Joseph Hargitai, Einstein Medical School
IU and Cornell are leading within XSEDE in experimenting with "____ as a Service" models coordinated with XSEDE. Rockhopper is available on a "fee for service" basis, providing a user-funded complement to XSEDE services. Rockhopper (rockhopper.uits.iu.edu) is Penguin Computing's Penguin-On-Demand (POD) supercomputing cloud appliance hosted by Indiana University. Rockhopper is a collaborative effort between Penguin Computing, IU, the University of Virginia, the University of California Berkeley, and the University of Michigan to provide supercomputing cloud services in a secure US facility.

IU has also played a leadership role in the creation of the Service Providers Forum (SPF), with David Hancock as the inaugural vice chair of the SPF.

12.4.2 Indiana University Level 2 Service Provider Systems

Level 2 - Quarry (Virtual Machines) - The Quarry Gateway Web Services Hosting resource at Indiana University consists of multiple Intel-based HP systems geographically distributed for failover in Indianapolis and Bloomington, IN.

Level 3 (Pending) - Rockhopper - Rockhopper is a supercomputing "cluster on demand" services in a secure US facility. Rockhopper is a 4.4 TFLOPS system based on AMD processors.

Being phased out from XSEDE allocations: Data Capacitor - The Data Capacitor is a high speed/high bandwidth Lustre storage system for research computing that serves all IU campuses.

Fate relative to XSEDE yet unclear – Scholarly Data Archive (HPSS) - HPSS comprises 23 IBM x3650 M2 servers, two TS3500 tape libraries, a total of 48 Jaguar4 tape drives, 500TB disk cache, and 15PB of tape. Data are replicated between the Indianapolis and Bloomington campuses. While TeraGrid allocations will be provisioned until at least the end of 2012, currently HPSS is not accepting new XSEDE allocations. (The possibility of new allocations of one particular type remains under discussion).

12.5 NCSA – SP Forum Annual Report Summary

This past year NCSA operated three unique HPC resources: *Ember*, *Lincoln*, and *Forge*; and also the only allocated archival storage system, *MSS*. *Ember* consisted of 4 SGI Altix UV 1000 systems each with 384 cores and 2 TB of memory with Intel Xeon Nehalem-EX (6-core) processors. This system was discontinued from XSEDE access at the end of September 2011. *Lincoln* consisted of 192 compute nodes (Dell PowerEdge 1950 dual-socket nodes with quad-core Intel Harpertown 2.33GHz processors and 16GB of memory) and 96 NVIDIA Tesla S1070 accelerator units. Each Tesla unit provides 345.6 gigaflops of double-precision performance and 16GB of memory. This system was retired at the end of September 2011. *Forge* consists of 36 Dell PowerEdge C6145 quad-socket nodes with dual 8-core AMD Magny-Cours 6136 processors and 64 GB of memory. Each node supports either 6 or 8 NVIDIA Fermi M2070 GPUs. *MSS*, NCSA's hierarchical archival storage system is available for permanent storage of data. Access is via FTP- and SSH-based transfer clients, including GridFTP clients. NCSA's *MSS* now holds more than six petabytes of data and has the capacity to archive ten petabytes of data. These resources are embedded in a multi-10Gbps local network, complemented by multi-10Gbps external connectivity, to support data transfers between resources and to/from external endpoints.

This suite of resources has enabled ground-breaking research across a number of domains and represents NCSA's ongoing commitment to investigate and support novel and innovative resources and resource usage modalities. *Ember* was the primary shared memory system in the TeraGrid and XSEDE portfolio until supplanted by PSC's *Blacklight*. *Ember* enabled groundbreaking work in estimating the potential for breakage in steel beams, studying the migration of massive planets in proto-planetary disks, and exploring sources of hydrogen for tomorrow's fuel cells. *Ember* was retired from NSF use at the end of September 2011.

Lincoln was the first production GPU-computing cluster offered to the TeraGrid/XSEDE community, and was replaced by *Forge* in Fall 2011 to provide researchers with the current state-of-the-art NVIDIA GPU. *Lincoln* and *Forge* have enabled the development of new accelerated applications, as well as providing production resources to support applications that had already been enhanced to take advantage of GPUs. In particular, we focused strongly on community applications, such as Amber and NAMD, early on. The molecular dynamics community was an early adopter and made use of *Lincoln* and *Forge* to model surfactants of one class to improve products like detergents and shampoo, but also another class of surfactants as a way of controlling the delivery of drugs in the body and improving their impact.

A much broader range of application areas have been effective in harnessing this important emerging computing platform. For example, *Forge* was used to better understand the fluid dynamics of the interaction between the vortex wake coming off the rotor blades and how those vortex structures affect the lift behavior and drag behavior of the rotor blades for rotorcraft design. Of particular note, a major step forward in enabling whole-cell simulations of bacteria. "This is the first time that we're modeling entire cells with the complete contents of the cellular cytoplasm represented," says Illinois postdoctoral researcher and lead author Elijah Roberts. "We're looking at the influence of the whole cellular architecture instead of modeling just a portion of the cell, as people have done previously." NCSA's *Forge* GPU computing system is retiring September 30, 2012.

12.6 NICS – SP Forum Annual Report Summary

12.6.1 Executive Summary

In 2009, the National Institute for Computational Sciences (NICS) delivered the first academic petaflop computer to the NSF community—a Cray XT5 called Kraken (OCI-0711134). By the end of 2010 two Cray systems at NICS, Kraken and the smaller Cray XT4 Athena, were primary providers of computer time to the TeraGrid, delivering more than 70% of all NSF compute cycles. In the first XSEDE project year, the Cray XT5 continued to add impressive milestones to a fabled history. Not only did **Kraken deliver its 2 billionth CPU hour** to open science in November 2011, but the computer also sustained a **utilization of 94%** for the year and was **available to users 98%** of the time with several periods of over 30 days of continuous availability.

On July 25, 2011 NICS decommissioned the Athena system (a 166 TF Cray XT4). In its last year of service, Athena was available for use 99% of the time and had 93% system utilization. For the month of production that fell within this XSEDE project year, Athena was available 99.8% of the time and experienced 92% utilization.

Under the OCI-0906324 award, the SGI Altix called Nautilus and the Remote Data and Visualization (RDAV) center has served to aid in the significant challenge of transforming large-scale data into knowledge and insight by providing scientists with well-engineered and well-supported remote visualization, analysis, and scientific workflow technologies. For PY1, Nautilus was available 92.6% of the time with an average utilization of 47.8%. Nautilus-enabled science, operational statistics, collaborations, outreach, and training have been detailed in the RDAV annual report that was submitted on June 1st, 2012.

This year ended with 671 active projects containing 3,469 users with active accounts on NICS resources. Users of NICS resources generated 17 known publications this year with an additional 21 known works accepted for publication. NICS resources delivered 59% of all computational hours consumed by the NSF community in the first XSEDE project year.

Through partnerships with Oak Ridge National Laboratory, the Georgia Institute of Technology's Keeneland project, the National Center for Supercomputing Applications' Blue Waters project, and the Joint Institute for Computational Sciences' Application Acceleration Center of Excellence; NICS continued in this year to specify and evaluate leading-edge and next generation technologies while increasing XSEDE's expertise and value to the community in preparation for these technologies and associated software. At the same time, NICS disseminates knowledge and expertise through participation in the EPSCoR Track 2 project. This collaboration bridges campuses by building cyberinfrastructure (CI) linked, community specific knowledge environments that embody the desktop to XSEDE ecosystem.

There were 2,838 help tickets opened this year and support staff at NICS resolved 2,737 tickets over the same period with a mean-time-to-resolution of just over 12 hours. Staff at NICS provided 7 training events in this project year and participated in 11 E & O events. These events collectively reached over 600 individuals of which over 200 were from underrepresented populations. Staff at NICS also published 9 articles, had an additional 14 accepted for publication, and submitted 13 more for consideration. NICS staff also gave 16 presentations and prepared 6 posters on HPC related topics during the year. These and other activities have been detailed in the NICS annual report that was submitted on June 29th, 2012.

12.6.2 Resource Descriptions

NICS had three NSF funded computational resources in production during the fifth project year: Kraken, Athena, and Nautilus. These systems shared a Network File System (NFS) that contains user directories, project directories and software directories. An optional one-time password tokens provide secure access to both the computational and storage resources at NICS.

12.6.2.1 Kraken

Kraken is a Cray XT5 consisting of 9,408 compute nodes, each containing two 6-core AMD Istanbul Opteron processors and 16 GB of on-node memory. The result is 112,896 compute cores that deliver 1.17 PF at peak performance with 147 TB of memory. Communications take place over the Cray SeaStar2+ interconnect. A parallel Lustre file system provides 3.3 PB (raw) of short-term data storage.

12.6.2.2 Athena

Athena was a Cray XT4 with 4512 compute nodes interconnected with SeaStar, a 3D torus. Each compute node had one four-core AMD Opteron for a total of 18,048 cores. All nodes had 4 Gbytes of memory: 1 Gbyte of memory per core. A parallel Lustre file system provided 100 TB (raw) of short-term data storage.

12.6.2.3 Nautilus

Nautilus, an SGI Altix UV 1000 system, is the centerpiece of NICS Remote Data and Visualization (RDAV) Center that is also located at ORNL. It has 1024 cores (Intel Nehalem EX processors), 4 TB of global shared memory, and 8 GPUs in a single system image yielding 8.2 TF at peak performance. A parallel Lustre file system provides 427 TB (raw) of short-term data storage.

12.6.2.4 HPSS Archival Storage

The High Performance Storage System (HPSS), developed and operated by ORNL, is capable of archiving hundreds of petabytes of data and can be accessed by all major leadership computing platforms. Incoming data is written to disk and later migrated to tape for long term archiving. This hierarchical infrastructure provides high-performance data transfers while leveraging cost effective tape technologies. Robotic tape libraries provide tape storage. The center has four SL8500 tape libraries holding up to 10,000 cartridges each. The libraries house a total of 24 T10K-A tape drives (500 GB cartridges, uncompressed), 60 T-10K-B tape drives (1 terabyte cartridges, uncompressed), and 20 T10K-C tape drives (5 terabyte cartridges, uncompressed). Each T10K-A and T10K-B drive has a bandwidth of 120 MB/s. Each T10K-C tape drive has a bandwidth of 240 MB/s. Disk storage is provided by DDN storage arrays with nearly a petabyte of capacity and over 12 GB/s of bandwidth. This infrastructure has allowed the archival system to scale to meet increasingly demanding capacity and bandwidth requirements with nearly 10 PB of NICS data stored as of May 31, 2012.

12.7 Pittsburgh Supercomputing Center – SP Forum Annual Report Summary

The Pittsburgh Supercomputing Center (PSC) operates and supports, *Blacklight*, a powerful and unique resource for the national research community. *Blacklight*, an SGI Altix UV 1000 acquired with the assistance of an NSF grant and operated as an XSEDE resource, is the world's largest shared-memory system, providing two partitions of 16TB each. *Blacklight*, as well as other PSC systems, is supported by a central file system, archival storage and extensive LAN, MAN and WAN infrastructure. PSC has completed transition from a disk-tape archive system to a new scalable, disk-only file repository that provides much faster access to files. Its initial deployment has four petabytes. With operational funding from NIH, PSC also operates *Anton*, a special-purpose computer for molecular dynamics which is used by many NSF-supported researchers.

PSC resources enabled significant progress in many areas; e.g., genomics, protein folding, climate science, machine learning, natural language processing, anesthesiology, AIDS, hearing and deafness, Chagas disease, and pandemics. Users are finding *Blacklight's* large shared memory and other characteristics such as the NUMAlink interconnect to be very valuable. For example, Niel Henriksen of Tom Cheatham's group at the University of Utah has been running jobs that use more than 2 terabytes of memory on *Blacklight*. He said, "I'm really impressed with how fast these calculations can run on *Blacklight*. Henriksen says that the same calculations on *Kraken* take an order of magnitude longer and they swamp the disk resources. Jobs that use more than 2 terabytes have become commonplace.

PSC staff worked with many users to move applications to *Blacklight* and to optimize their performance. Special efforts were required to move users from NSCA's Altix UV system, *Ember*, when it was withdrawn as an XSEDE resource. PSC created the Memory Advantage Program (MAP) which provides individualized user support for porting applications to exploit *Blacklight's* unique large shared-memory capability.

PSC Systems staff worked with SGI to improve UV performance in many areas including the scalability and performance of memory-resident file systems. In networking, PSC staff worked with the NSF-supported Galaxy Genomics Portal at Penn State to enable connectivity between Galaxy and XSEDE via PSC's Three Rivers Optical Exchange so that Galaxy can use XSEDE as a backend.

PSC engaged in a broad range of Training, Education and Outreach activities, which included enlisting new communities into HPC, several major STEM education programs, and HPC training workshops. PSC supported a number of undergraduate interns and one high school senior.

PSC collaborates with two groups at Harvard University on the wiring diagram of the mouse visual cortex. PSC also worked with the Broad Institute, one of the largest genome sequencing centers in the world, to make two cutting-edge genome assembly codes available on *Blacklight*. Other collaborations include the NSF OCI petascale cosmology, earthquake and weather forecasting projects, and the Gates Foundation Vaccine Modeling Initiative.

Scientific co-director, Ralph Roskies, was appointed to the Board of Regents of the National Library of Medicine and other PSC staff members serve on national and local advisory boards.

12.8 Purdue University – SP Forum Annual Report Summary

Purdue provide services of an HPC cluster (Steele), a high-throughput computing resource (the Purdue Condor pool), and a cloud resource (Wispy) to the XSEDE user community. From Aug. 2011 to July 2012, Purdue XSEDE resources have supported 194 PIs with 203 allocations from 112 institutions.

The Steele cluster continues to be busy and highly utilized by the XSEDE users in the past 12 months. It has maintained 98% uptime throughout the reporting period. Steele continues to see higher allocation requests than its available cycles each quarter, and the trend is even higher demand in the September allocation cycle (as high as 62M SUs in the most recent round). XSEDE users access Steele through the XSEDE queues, and in addition to the portion of the cluster allocated for NSF users, XSEDE users have access to the entire cluster through its standby queues with a wall clock limit of 4 hours for the jobs. In this manner, XSEDE users typically consumes as high as three times of the cycles allocated for XSEDE on a monthly basis.

The Purdue's Condor pool continues to grow in size as Purdue's community cluster program expands. All 5 Purdue community clusters run Condor. Purdue's Condor pool continues to serve both XSEDE and OSG user communities, in addition to the campuses that are part of the DiaGrid initiative. Among the users, Thomas Robitaille, formerly a Spitzer postdoctoral fellow at the Harvard-Smithsonian Center for Astrophysics now a researcher at the Max Planck Institute for Astronomy is developing a new model grid to study galactic star formation. This new model will dramatically improve the parameter space coverage, the quality of the models and therefore the science that can be derived. This project has consumed more than 25M NUs in the Condor pool in the past 12 months. The total XSEDE usage of Condor was approx. 785K wall hours (7.5M CPU hours), while the pool delivered a total of 29M CPU hours to scientific computation and education projects in the past 12 months.

Wispy is an experiment cloud system in XSEDE. Its usage includes those from the joint TeraGrid/OSG ExTENCI project in the development, testing of virtualization technologies for science users. Wispy has recently been upgrade in July 2012 to new hardware based on Intel's Sandy Bridge series of processors. The SP staff has revamped the authentication on Wispy, simplifying the process for XSEDE users and providing clear documentation for Wispy access (https://www.xsede.org/web/guest/purdue-wispy). A Cloud Dashboard has been developed. FutureGrid is interested in using it for its users. Purdue team is working with the FutureGrid team to share the Dashboard tool.

In addition to providing computation and data resources, Purdue SP continues to support science gateways that utilize XSEDE resources and are being used for both research and education. As an example, a CI-TEAM project (WaterHUB) developed SWATShare, powered by the XSEDE HPC resource and HUBzero technology, and used in hydrology classrooms where students were not only able to simulate real-world scale watershed and calibrate the models, but also able to utilize water data from the National Climatic Data Center, the U.S. Geological Survey and other sources, which was impossible in a classroom before. The team presented their findings at the XSEDE12 conference, July 2012.

In support of interoperable cyberinfrastructures, Purdue continues to bridge OSG and XSEDE, e.g., the OSG high-throughput HPC (HTHPC) user jobs flow seamlessly to XSEDE Steele. A total of 315,837 OSG MPI jobs ran on Steele and consumed 233.5K processor-core hours on Steele. During the same period, various OSG VO serial jobs, totally more than 4 millions, consumed a total of approximately 12 million processor hours in Purdue's Condor pool in the past 12 months.

Purdue SP staff also supports and contributes to various XSEDE EOT activities and campus information technology community. For four years in a row (2009-2012), Purdue's Rosen Center for Advanced Computing has won the Campus Technology's Innovators awards to its DiaGrid (Condor pool), Community Cluster program, HUBzero and Purdue Studio of mobile applications for teaching and learning, including the resources that contribute to the national cyberinfrastructure TeraGrid/XSEDE.

12.9 San Diego Supercomputer Center – Service Provider Annual Report Summary

SDSC has successfully deployed the innovative *Gordon* system in early 2012, the first NSF resource dedicated specifically to data-intensive computing, and has continued to effectively serve the gateway and modest-scale user community with the high-productivity *Trestles* cluster.

Gordon is an Appro-integrated cluster that was co-designed by SDSC and Appro to address the challenges of data intensive computing, with a focus on *moving data efficiently through the memory hierarchy*. The beginning of the reporting period focused on risk mitigation and testing, using early Gordon hardware (e.g. I/O nodes) and *Dash*, the risk-reduction system for *Gordon*. The *Gordon* compute nodes were delivered to SDSC in November 2011, shortly after Intel released the Sandy Bridge processors. After passing a rigorous set of acceptance tests and a 20-day reliability test, *Gordon* was officially accepted on February 24, 2012 and has now been added to the XSEDE resource portfolio as the first system designed specifically for data intensive computing. Since entering production status, *Gordon* has been stable and performing very well.

The interest in *Gordon* is high, and we are encouraging allocation requests that make use of its unique architecture features such as flash memory and large memory vSMP "supernodes." In addition, several users have been allocated dedicated use of *Gordon's* I/O nodes; this is a new prototype allocations model for data mining and database projects.

Trestles is now in its second year of production and is successfully supporting the modest-scale/gateway user community, with a focus on user productivity and fast turnaround. The system is well over-requested in the allocation process. The system utilization has recently ramped up from the targeted 70% level to now well over 90%; queue wait times and expansion factors, while still good relative to most XSEDE systems, are increasing beyond desired levels. We are analyzing usage data and user job patterns to optimize utilization while maintaining low expansion factors (results presented at XSEDE'12).

SDSC completed deployment of the *Data Oasis* Lustre parallel file system, which supports *Gordon*, *Trestles* and SDSC's *Triton*, with 4PB capacity and demonstrated 100 GB/s aggregate bandwidth. *Gordo*n users have access to a ~2PB/50GBps scratch file system, while *Trestles* users access a 400TB/10GBps scratch system. We recently deployed 400 TB of *Data Oasis* "project storage" for allocated storage accessible by both *Gordon* and *Trestles* users; an additional 2 PB of project storage will be deployed next year. This represents a new paradigm of intermediate-term allocated storage, with the hope of significantly reducing long-term archival requirements.

In addition to *Gordon* and *Trestles*, SDSC provided other XSEDE production resources during the reporting period. The *Dash* system was used effectively as an allocated resource for porting of applications and initial production science in a *Gordon*-like architecture; it was decommissioned as an XSEDE resource after *Gordon* was deployed. The *GPFS-WAN* storage system served as an allocated resource by users and also used by XSEDE for evaluation of GPFS for its future wide-area storage designs; this resource was decommissioned in June 2012 after all its users were assisted in migrating necessary data to other storage resources. Legacy archival data from TeraGrid/XSEDE-allocated users was retained in the HPSS and SAM-QFS tape archives; all

users have been notified and assisted in migrating necessary data, and these systems are being decommissioned in the summer 2012.

SDSC conducted a large number of effective education, outreach and training activities that engaged users, students and new user communities with conferences, summer workshops and education programs.

The resources delivered by SDSC to XSEDE users have resulted in a number of important scientific advances, including:

- The successful CIPRES gateway for analysis of phylogenetic data sets has utilized *Trestles* (and recently *Gordon*) for its entire ~10M SU allocation this past year. This single gateway uses <1% of XSEDE compute resources, but now accounts for 29% of all XSEDE users. In addition to the broad adoption of this gateway, the scientific impact is enormous with gateway users producing nearly 200 scientific publications this past year. Please see "SDSC's CIPRES Science Gateway Clarifies Branches in Evolution's 'Tree of Life'" at http://www.sdsc.edu/News%20Items/PR070312_cipres.html .
- Harvard researchers use *Trestles* and now *Gordon* to perform computations for their Clean Energy Project (CEP), with the goal of creating the next generation of organic solar cells as an inexpensive and efficient source of energy. Please see "SDSC's Trestles Supercomputer Speeds Clean Energy Research" at http://www.sdsc.edu/News%20Items/PR042312_solar.html for the full story.

Researchers from the University of California campuses at Irvine and San Diego have found a new approach to the creation of customized therapies for virulent flu strains that resist current antiviral drugs. The researchers utilized simulations on *Trestles* to predict how pocket structures on the surface of influenza proteins promoting viral replication can be identified as these proteins evolve, allowing for possible pharmaceutical exploitation. The findings, published in Nature Communications, could aid development of new drugs that exploit so-called flu protein 'pockets.'

12.10 TACC – Service Provide Annual Report Summary

12.10.1 Ranger Executive Summary

This reporting period marks the fourth full year of operations of the Ranger resource. This system remains in high demand, and continues to exceed metrics for performance and stability. Within the reporting period the system has been used by 1,920 users to run 478,420 jobs consuming 483.5 million service units (SUs). A total of 1,088 principal investigators used Ranger within the reporting period in support of 1,306 research projects with 88.5% of the SUs being utilized by XSEDE PIs. Researchers from institutions within the state of Texas also continued to use Ranger via the pool of allocation designated for that community. Details of this usage, as well as highlights of some of the scientific achievements will be provided in the annual report (submitted separately). The outstanding reliability of Ranger has resulted in great demand among the open science community. For the current quarter, researchers requested more than 100M SUs on Lonestar.

Overall, the Lonestar project continues to be a success. Users continue to flock to the resource in large numbers, and the effort to continue to expand the system with external funding continues.

12.10.2 Lonestar Executive Summary

This reporting period marks the first full year of operations of the Lonestar-4 resource. This system remains in high demand, and continues to be the most over requested system in the XSEDE ecosystem. During the past year, the system has continued to exceed metrics for performance and stability, delivering 2-4x the performance of Ranger on most user applications. Within the reporting period the system has been used by more than 1,600 users to run nearly 425,000 jobs consuming 184 million service units (SUs). Of this total, more than 46% of delivered cycles went to the XSEDE community, well over the committed 40% of the system promised in the original grant agreement (XSEDE users received nearly 13M more SUs than TACC promised in the period). Details of this usage, as well as highlights of some of the scientific achievements are provide in the annual report. The outstanding performance of Lonestar has put it in great demand among the open science community. For the current quarter, researchers requested more than 100M SUs on Lonestar. Lonestar is the first non-Track 2 system in XSEDE to have more than 100M SUs requested in a quarter, and with requests exceeding available cycles by more than 600%, is the most over requested resource in XSEDE/TeraGrid history.

As detailed in the original proposal, the Lonestar project involved leveraging the funding provided by NSF with funding from other sources to build a much larger and capable system. The proposed 302 TeraFlop system represented the $2.86M project budget, and an additional $6.14M in funds from other partners to deploy the $9M system. Fundraising has exceeded that target, and currently $11.4M has been made available for investment in the system. The original 302TF system has been augmented with additional compute nodes (taking the total to 311TF), large shared memory nodes, and GPU nodes for remote visualization and GPGPU computing to provide a comprehensive environment for scientific computing.

Overall, the Lonestar project continues to be a success. Users continue to flock to the resource in large numbers, and the effort to continue to expand the system with external funding continues.

12.10.3 Longhorn Executive Summary

The third year of the Longhorn XD Vis project has been successful in continuing to recruit new users, developing effective visualizations, developing tools to respond to evolving user needs, and training current and next generation scientists and engineers. Our efforts over the last year have focused on increasing visualization and data analysis

usage on Longhorn. Additionally, we have focused on the enhancement and improvement of the user experience by improving the remote visualization capabilities, particularly through the Longhorn Visualization Portal. Longhorn usage has continued to increase with a diverse portfolio of users wanting Visualization and Data Analysis. There are currently 423 persons running jobs from 607 projects with 513 PIs burning an average of 600K SUs over the third year.

More details can be found in the Longhorn annual report (to be submitted separately.)

13 Indiana University Pervasive Technology Institute – Research Technologies – XSEDE Service Provider Report Q2 Calendar 2012

13.1 Executive Summary

During the current reporting period there were no unscheduled outages of network connectivity from the Indiana University Pervasive Technology Institute – Research Technologies SP resources and XSEDE.

IU has also played a leadership role in the creation of the Service Providers Forum (SPF), with David Hancock as the inaugural vice chair of the SPF.

13.1.1 *Resource Description*

Level 2 - Quarry (Virtual Machines) - The Quarry Gateway Web Services Hosting resource at Indiana University consists of multiple Intel-based HP systems geographically distributed for failover in Indianapolis and Bloomington, IN. Currently there are four HP DL160 front-end systems at each site, all configured with dual quad-core Intel E5603 processors, 24 GB of RAM, and a 10 gigabit Ethernet adapter. There are a total of 48 XSEDE Virtual Machines. The front-end systems host the KVM-based virtual machines. Virtual Machine (VM) block storage is provided by two HP DL180 servers at each site configured with a quad-core Intel X5606 processor, 12 GB of RAM, a 10 gigabit Ethernet adapter, and a RAID controller attached to an HP storage array. Quarry is used solely for hosting Science Gateway and Web Service allocations, or services to support central XSEDE infrastructure. Requests are restricted to members of approved projects that have a web service component.

Level 3 (Pending) - Rockhopper - Rockhopper is a collaborative effort between Penguin Computing, IU, the University of Virginia, the University of California Berkeley, and the University of Michigan to provide supercomputing “cluster on demand” services in a secure US facility. Researchers at US institutions of higher education and federally funded research centers can purchase computing time from Penguin Computing and receive access via high-speed national research networks operated by IU. It takes just minutes to go from holding a credit card in one’s hands and filling out a web form to being computing on Rockhopper (the system itself is owned by Penguin; cycles on Rockhopper are purchased from Penguin). Rockhopper is a 4.4 TFLOPS system based on AMD processors.

Being phased out from XSEDE allocations: Data Capacitor - The Data Capacitor and Data Capacitor WAN are high speed/high bandwidth Lustre storage systems for research computing that serves all IU campuses. Each are comprised of 6 servers (2 MDS, 4 OSS). The total usable storage capacity is 1.1 TB for the Data Capacitor and 339 TB for Data Capacitor WAN.

Fate relative to XSEDE yet unclear – Scholarly Data Archive (HPSS) - HPSS forms the core of IU's Scholarly Data Archive, a tape-based archival data storage service. It comprises 23 IBM x3650 M2 servers, two TS3500 tape libraries, a total of 48 Jaguar4 tape drives, 500TB disk cache, and 15PB of tape. Data are replicated between the Indianapolis and Bloomington campuses. While TeraGrid allocations will be provisioned until at least the end of 2012, currently HPSS is not accepting new XSEDE allocations. (The possibility of new allocations of one particular type remains under discussion).

13.2 Science Highlights

Empirical Force Field Dihedral Parameter Optimization through Airavata Workflows under ParamChem

Researchers at the University of Illinois and University of Maryland (Baltimore) have collaborated with the Indiana University Science Gateway Group to deploy services and workflows to automate the process of parameterization of Dihedral Parameters used in CGenFF (Charmm General Molecular Forcefield(FF)) in an extensible framework using XSEDE resources. The Apache Airavata workflow toolkit is used for registering workflow components, prefabricated automation steps, and workflow execution management. Asynchronous workflow monitoring enables long running optimization procedures.

Molecular Forcefield Parametrization is critical in enabling the simulation of large molecular systems and their interactions as a way of understanding their behavior. Study of biomolecules and their interactions including how they are modulated by pharmaceuticals is essential to rapid drug prototype development and reducing toxicity. Including various models in the process enables testing the validity of the models and corroborating experimental observations. The workflow infrastructure will benefit a large community of modeling scientists who deal with large and novel molecular systems for which transferable parameters can be generated that enable simulation of new systems including chemical cures of cancer and other diseases and to understand how biomolecules work in physiological and pathological states.

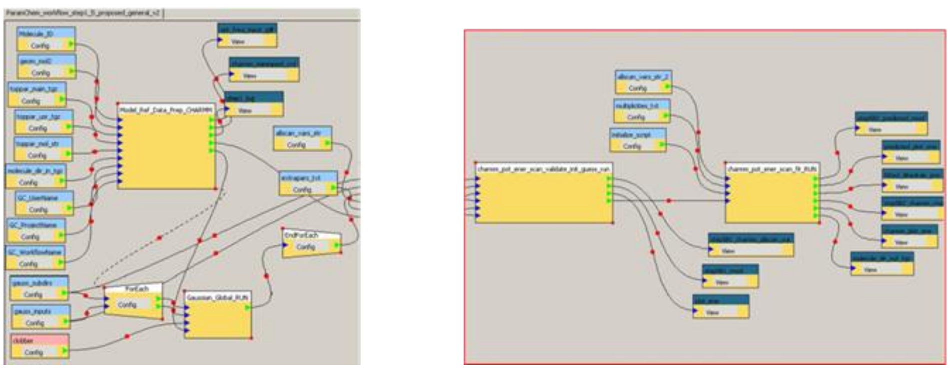


Figure 2. Example workflow to optimize dihedral angles.

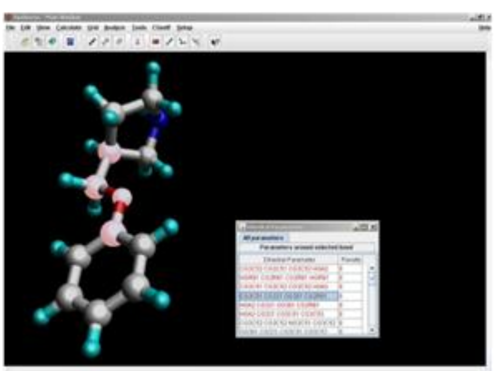


Figure 3. Jamberoo Molecule editor with molecule and table of initial guesses from CGenFF program.

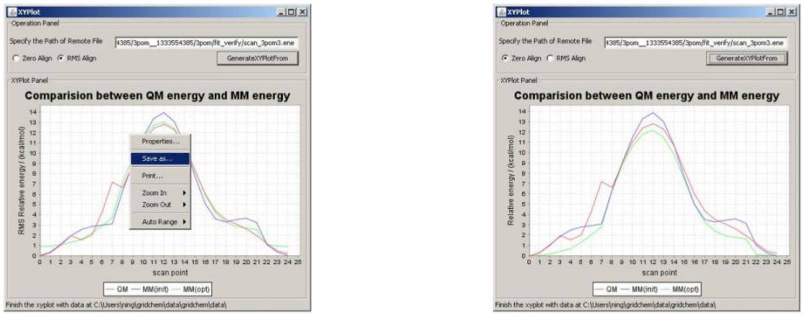


Figure 4. Visualize the parameterization validation results and history file retrieval.

FF parameterization is a tedious process of creating model systems for which reference data need to be obtained by high-level simulations and fitted at the FF level of description for use in long simulations. The parameterization process itself consists of an iterative procedure of optimization of various components of the FF model for each novel molecular moiety. These complex reference data computations require XSEDE supercomputing resources and supporting workflow management tools. This work is supported by grants from National Science Foundation: Grant # 0823198 and #1032742.

Principal Investigators: Dr. Sudhakar Pamidighantam (University of Illinois), Prof. Alex MacKerell (University of Maryland at Baltimore).

13.3 User-facing Activities

(If multiple awards at the SP institution, provide sub-sections for each award and for common infrastructure/services)

13.3.1 System Activities

What activities have happened with respect to the SP resources and supporting infrastructure (filesystems, networking, etc.) that impact users positively or negatively, and if negatively what is being done to address it? Describe new deployments (or decommissioning), significant configuration or performance changes, any significant problems or successes, and key system statistics like uptime, availability and utilization.

13.3.2 Services Activities

The Science Gateways group contributed to the ongoing gateway coordination, development and execution of work plans for requests made through ECSS in gateway related areas. The activities include community engagement and coordination of gateway support personnel across all sites. The group continues to lead the gateway use case requirements sessions for all funded gateway ECSS staffers. Gateway use cases will be finalized during Q3. Two gateway tutorials were accepted for the XSEDE12 conference: *"Developing Science Gateways using Apache Airavata API"* and *"Hands-on Tutorial for Building Science Gateway Applications on Cyberinfrastructure".*

13.4 Security

No Security issues to report.

13.5 Education, Outreach, and Training Activities

Campus Bridging

The Campus Bridging team engaged in three major activities during this quarter: definition of use cases and quality attributes with the Architecture and Design team; discussion of campus bridging with XSEDE and preparation of a paper for the XSEDE12 conference; and continued work on pilot projects for testing the Global Federated Files System (GFFS) software on campuses. The GFFS Pilot program remains in a "wait state", awaiting version of software that have passed an operational readiness review and made available for testing.

PTI Ready, Set, Robots Robot Workshop 2012

PTI conducted its sixth year of the Robot Workshop. This year eight PTI staff worked with 52 future engineers and scientists programming LEGO Mindstorms™ robots. In anticipation of the NASA Curiosity landing later in the summer, the workshop adopted a Mars rover theme. Students ranging from seventh grade to high school juniors programmed robots to simulate a Mars sample return rover mission. The test arena contained several swatches of colored tape that represented different mineral deposits of interest to scientists back on earth. The challenge was to detect a minimum number of the minerals and then dock with the return vehicle by homing in on an infrared signal.

The students worked in teams of three or four. Each team was paired with a PTI staff member as a mentor. The tasks were broken down into sub-tasks with mini-evaluations for each sub-task. The mini-evaluations gave the teams a chance to explain their algorithms they developed and used as well as exchange ideas about challenges they all found. The evaluations contained several criteria, but the students excelled in the creativity criteria.

On the second day of the two-day workshop, each team presented and demonstrated its solution to an assembled group of parents, siblings, and PTI staff. The interest in this workshop remains high and PTI continues to seek ways to bring this material to even more future researchers next year.

Figure 5. Ready, Set, Robots participants program their robot.

XSEDE12 Conference

The final program for the XSEDE12 conference includes four tracks: Science, Technology, Education, Outreach and Training, and Software and Software Environments. The committee accepted 65 paper, 19 tutorials and 39 student posters. The conference will run July 17-19 in Chicago, Illinois, with projected attendance between 500-600 total. A full report on the conference will be published after the conclusion of the event.

13.5.1 Presentations

Michael, S. 2012. How to Tune Your Wide Area File System for a 100 Gbps Network. (Presentation) Lustre User Group (Austin, TX, 23-25 April, 2012).

Walgenbach, J. 2012. Secure Identity Mapping for Lustre 2.X. (Presentation) Lustre User Group (Austin, TX, 23-25 April, 2012).

Simms, S.C. 2012. Using the Lustre File System at 100Gb across 2,300 Miles,(Presentation) N-Wave Stakeholder Meeting (Boulder, CO, 22-23 May, 2012).

Simms, S.C. 2012. Secure Identity Management for Lustre 2.X (Presentation), Lustre Roadmap BOF at ISC 2012 (Hamburg, Germany, 19 June, 2012).

Hancock, D.Y. 2012. HPC 101 (Presentation) InCNTRE Internship Program - Summer of Networking (Bloomington, IN, 21 June, 2012).

13.6 SP Collaborations

No collaborations this quarter.

13.7 SP-Specific Activities

XSEDE Networking

During the second quarter of the 2012 calendar year, there were zero (0) unscheduled outages and two (2) scheduled maintenance events that directly affected Indiana University's connection to XSEDE. Maintenance accounted for three hundred one (301) minutes of maintenance downtime. There were one hundred thirty-one thousand forty (131,040) minutes during the second quarter of 2012 calendar year, yielding a 99.77% uptime. During the outages and maintenance events, Indiana University's XSEDE connectivity failed over to our Internet2 and NLR connectivity, giving an actual site-to-site XSEDE uptime of 100%.

XSEDE Annual User Survey

The 2012 XSEDE User Satisfaction Survey was launched on May 1, 2012, after receiving approval from XSEDE leadership and IU IRB, and closed on June 11, 2012. The survey sample population was generated randomly from (1) a complete list of XSEDE users and (2) a list of NSF funded principal investigators from the NSF award database. As of June 11 there were 734 responses representing just over 7% of the total survey population (400+ were current XSEDE users). The final survey report will be available not later than August 31, 2012.

XSEDE Operation Center Failover Site

The IU GlobalNOC continued work on implementation activities in preparation for acting as a failover site for the XSEDE Operations Center. The status of various failover preparations follows:

- Phones – completed XSEDE NOC Greeting on IU phone system
- Email – completed setup of e-mail acceptance for failover on IU side. Pending XSEDE/NCSA setup, policy, procedures and documentation before further implementation.
- Ticketing – IU awaits final decision on new trouble ticket system for XSEDE. Pending XSEDE.

- Monitoring – currently SNAPP monitoring from Pittsburgh. Steven McNally looking at how IU NOC to get access to current monitoring. Pending XSEDE.
- Process and Procedure Documentation – XSEDE evaluating moving all documentation to XSEDE staff wiki. Pending XSEDE.
- Fail-over Documentation – Fail-over documentation is part of process and procedure documentation above – pending move to XSEDE staff wiki. Pending XSEDE.
- Training – Training for IU NOC staff pending finalization of trouble ticket system. Pending XSEDE.

13.8 Publications

13.8.1 Publications

Michael Kluge, Stephen Simms, Thomas William, Robert Henschel, Andy Georgi, Christian Meyer, Matthias S. Mueller, Craig A. Stewart, Wolfgang Wünsch, Wolfgang E. Nagel, Performance and quality of service of data and video movement over a 100 Gbps testbed, Future Generation Computer Systems, Volume 29, Issue 1, January 2013, Pages 230-240, ISSN 0167-39X, 10.1016/j.future.2012.05.028.

Scott Michael, Liang Zhen, Robert Henschel, Stephen Simms, Eric Barton, Matthew Link, A study of lustre networking over a 100 gigabit wide area network with 50 milliseconds of latency, Proceedings of the fifth international workshop on Data-Intensive Distributed Computing, June 18-22, Pages 43-52, ISBN: 978-1-4503-1341-4, 10.1145/2286996.2287005.

Richard Knepper, William Johnson, Scott Michael, Robert Henschel and Matthew Link. Wide-area 100Gb Networking at the SCinet Research Sandbox. Proceedings of the 7th IEEE International Conference on Networking, Architecture, and Storage, Xiamen, China June 28-30, 2012.

13.8.2 Technical reports (not peer reviewed)

Henschel, R., S.C. Simms, D. Hancock, S. Michael, T. Johnson, N. Heald, T. William, D.K. Berry, M. Allen, R. Knepper, M. Davy, M.R. Link and C.A. Stewart. *Technical Report: Report on Lustre use across an experimental 100Gb network spanning 2,175 mi.* Indiana University. 2012. Available from: http://hdl.handle.net/2022/14137

13.9 Metrics

13.9.1 Standard User Assistance Metrics

Time to Resolution	account issues	file systems	grid software	jobs/batch queues	login/access issues	mss/data issues	network issues	software/apps	system issues	other
0-1 hr										
1-24 hr										
1-7 d										1
1-2 wk										
> 2 wk	1	1				2		1		1
Still Open						1			1	1

Table 1. IU ticket resolution times by category from the XSEDE Ticket System.

Summary statistics for XSEDE Knowledge Base:

Total number of document available in the XSEDE KB at the end of the 2nd quarter=497

Number of documents added = 19

Number of documents modified = 79

Total number of documents retrieved = 71,009

Total number of documents (retrieved minus bots) = 57,451

13.9.2 Standard systems metrics

13.9.3 SP-specific Metrics

13.9.3.1 Quarry Virtual Machines

No hardware changes were made to the service during this quarter. A four hour outage window was required for network upgrades; otherwise there were no service outages.

Quarry remains a very actively used system as a host for Virtual Machines and Science Gateways. There are a total of 50 allocated VMs active at present on Quarry, including 21 for Science Gateways, 12 for IU internal gateway efforts, and 17 to support XSEDE core services.

13.9.3.2 Rockhopper

During the past quarter a total of 32 distinct individuals made use of Rockhopper.

13.9.3.3 Data Capacitor

There remain a total of 53 XSEDE allocations, for a total of 2.25 TB, on the Data Capacitor, either held over from TeraGrid or provided under XSEDE upon special request, due to need for particular capabilities provided uniquely by the Data Capacitor.

13.9.3.4 Scholarly Data Archive – HPSS

There remain a total of 23 XSEDE allocations, for a total of 53TB, on the IU Scholarly Data Archive (HPSS) tape archival storage system.

System	# Allocations	# VMs allocated	# TB allocated	# accesses	Storage high water mark	TB uploaded	TB downloaded
Quarry VM	21	50					
Rockhopper				1,383			
Data Capacitor	53		2.25				
Scholarly Data Archive – HPSS	23		53				

Table 2. Indiana University Pervasive Technology Institute – Research Technologies Service Provider system key usage metrics.

System	Overall % uptime	# planned downtimes	Planned downtime duration total (minutes)	# unplanned downtimes	Unplanned downtime duration total (minutes)	Total minutes in reporting period
Quarry VM		1	240			
Rockhopper						
Data Capacitor						
Scholarly Data Archive – HPSS						

Table 3. Indiana University Pervasive Technology Institute – Research Technologies Service Provider system key usage metrics.

14 NCSA - Service Provider Quarterly Report

14.1 Executive Summary

NCSA's *Forge* GPU computing system is retiring September 30, 2012. Users were sent a notice to prepare them for the transition to other systems. NCSA is determining if it can maintain access to the tape archive beyond September 30, 2012.

14.1.1 Resource Description

Forge consists of 36 Dell PowerEdge C6145 quad-socket nodes with dual 8-core AMD Magny-Cours 6136 processors and 64 GB of memory. Each node supports 6 NVIDIA Fermi M2070 GPUs.

NCSA's hierarchical archival storage system is available for permanent storage of data. Access is via FTP- and SSH-based transfer clients, including GridFTP clients. NCSA's mass storage now holds more than six petabytes of data and has the capacity to archive ten petabytes of data.

14.2 Science Highlights

Biophysics: Protein Structure Refinement Using Physical Principles (Justin MacCallum, Stony Brook University)

Protein structure prediction has greatly advanced over the previous 20 years. Current state of the art methods are able to produce a reasonably accurate model for the majority of target sequences. However, while these models are often globally correct, they get many of the local details wrong. Justin MacCallum and his team are developing physics-based methods to further refine these models with the aim of routinely producing models of sub-3 Angstrom accuracy. They use a consensus approach to identify which parts of the structure prediction are likely to be correct. They then use loose restraints to maintain these structural features. They then simulate the protein with replica exchange molecular dynamics on NCSA's Forge. Initial tests indicated success on 100% of test proteins, including two difficult cases where no predictor was successful in the CASP9 refinement competition. They are currently collaborating with experimental colleagues to further validate their tool. MacCallum's application runs about 100 times faster on GPU than CPU clusters, so Forge "really lets us tackle problems we couldn't address before" he notes, adding that what would take years to run on a CPU cluster is reduced to just weeks on Forge.

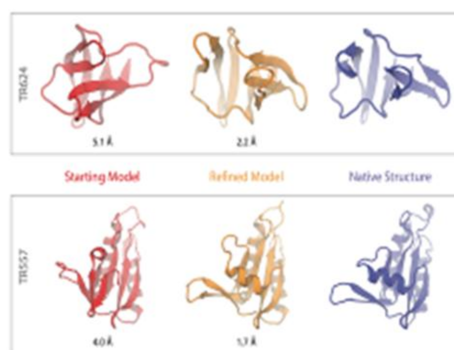


Figure 1 As part of the CASP competition, researchers are given a starting structure, shown in red, and then try to improve, or refine, the model. The MacCallum team's refinement structure, made using the tools they are developing, is shown in yellow. It closely resembles the native (correct) structure as determined by crystallography

Geography and Regional Science: Extending and Sustaining GISolve as a GIScience Gateway Toolkit for Geographic Information Analysis (Shaowen Wang, University of Illinois at Urbana-Champaign)

Viewshed analysis, often supported by Geographic Information Systems (GIS), is widely used in many application domains. A viewshed is created from a digital elevation model by using an algorithm that estimates the difference of elevation from one cell (the viewpoint cell) to the next (the target cell). However, as terrain data continue to become increasingly large and available at high resolutions, data-intensive viewshed analysis poses significant computational challenges. General-purpose computation on GPUs provides a promising means to address such challenges. This work has established a parallel computing approach to data-intensive viewshed analysis of large terrain data using NCSA's Forge. The approach Shaowan Wang and his team uses exploits high-bandwidth memory of GPUs and parallelism of massive spatial data to enable memory-intensive and compute-intensive tasks while CPUs are used to achieve efficient I/O management. Furthermore, a two-level spatial domain decomposition strategy was developed to mitigate a performance bottleneck caused by data transfer in the memory hierarchy of GPU-based architecture. A suite of experiments was designed to evaluate computational performance of the approach. The experiments demonstrated significant performance improvement over a well-known sequential computing method, and enhanced ability of analyzing sizable datasets that the sequential computing method cannot handle. The team's work has been published in the *International Journal of Geographical Information Science* and in the *Proceedings of TeraGrid 2011 Conference: Extreme Digital Discovery*, and presented at the *Association of American Geographers 107th Annual Meeting*. The team also won the Best Research Poster Award at the TeraGrid 2011 conference.

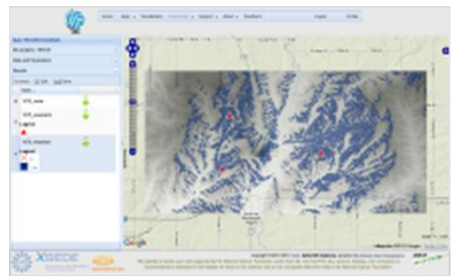


Figure 2 This shows the viewshed app in the CyberGIS Gateway. The viewshed app uses GPGPU for larg-scale and data-intensive viewshed analysis. Gateway provides an online problem-solving environment for community users to conduct cyberinfrastructure-empowered (currently using XSEDE, OSG, and local resources) spatial analysis and enable scientific collaborations.

14.3 User-facing Activities

14.3.1 System Activities

Filesystems and Storage:

Continued support of NCSA's Dell NVIDIA Cluster Forge and the archival system MSS.

Networking:

NCSA continues to manage the DNS services for XSEDE which includes the teragrid.org and xsede.org domains. We delegate subdomains to other SP sites and make DNS adds and deletions as requested by project groups. We setup and manage the dynamic dns system. We also acquire certificates for the various web services being deployed on XSEDE.

Systems:

NCSA's *Forge* GPU computing system is retiring September 30, 2012. Users were sent a notice to prepare them for the transition to other systems.

14.3.2 Services Activities

Continued support of NCSA's Dell NVIDIA Cluster Forge and the archival system MSS.

14.4 Security

The NCSA security team did not have any incidents involving XSEDE users or resources for the second quarter 2012. There were also no changes to the security monitoring during the quarter.

14.5 Education, Outreach, and Training Activities

NCSA provides online training via CI-Tutor. Course usage statistics are captured in the following table:

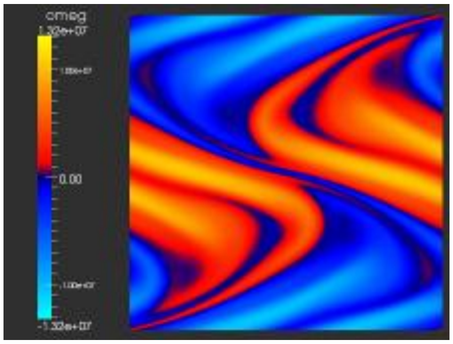
Course Title	**Apr-12**	**May-12**	**Jun-12**	**Total**
Access Grid Tutorials	9	6	0	15
BigSim: Simulating PetaFLOPS Supercomputers	4	4	3	11
Debugging Serial and Parallel Codes	12	7	13	32
Getting Started on theTeraGrid	10	4	6	20
Intermediate MPI	16	15	14	45
Introduction to MPI	92	69	86	247
Introduction to Multi-core Performance	16	11	12	39
Introduction to OpenMP	25	27	34	86
Introduction to Performance Tools	12	7	8	27
Introduction to Visualization	7	7	12	26
Multilevel Parallel Programming	11	20	12	43
Parallel Computing Explained	40	41	39	120
Parallel Numerical Libraries	4	9	7	20
Performance Tuning for Clusters	8	3	9	20
Tuning Applications for High Performance Networks	11	1	4	16
Using the Lustre File System	10	8	15	33
XSEDE Cybersecurity	1	1	2	4
Total	288	240	276	804

14.6 SP Collaborations

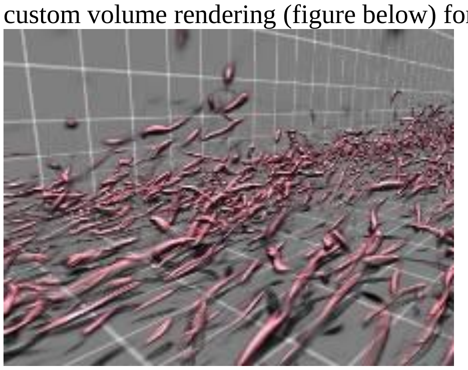
Nothing to report.

14.7 SP-Specific Activities

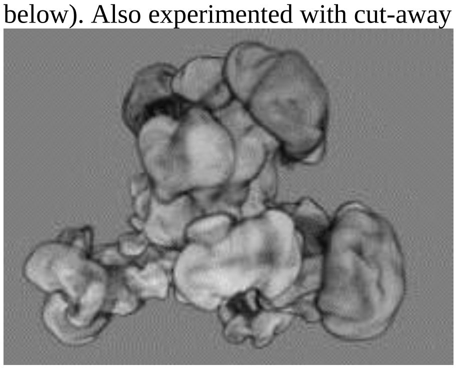
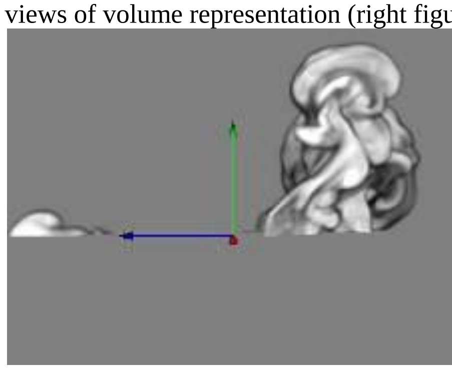
Started work on Asymptotic Scaling of Heat Transport in Infinite Prandtl Number Fluids at Very High Resolution with PI Benson Muite of the University of Michigan. Visualization support is focusing on instrumenting application code with ParaView Coprocessing libraries. This enables the application to avoid writing data which is prohibitive at scale, instead writing images (figure below) directly which also saves from having to do visualization as a post-processing step.



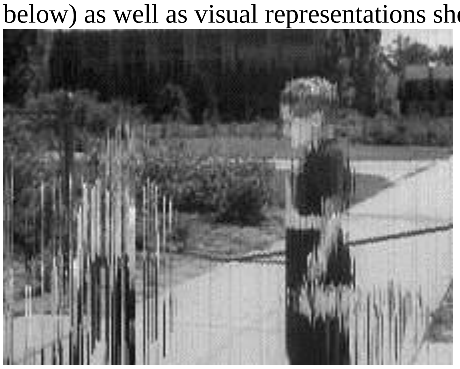
Working with A. Ferrante, University of Washington, on DNS of Spatially Developing Turbulent Boundary Layers. Developed visualizations to verify simulation and HDF5 reader, scene construction and custom volume rendering (figure below) for full production run.



Working with D. Townsley, University of Alabama, on Supernovae simulations. Modified custom volume renderer to represent both density and color using different variables from a single data-set (left figure below). Also experimented with cut-away views of volume representation (right figure below).

Working with V. Kuhn, University of Southern California, on Interactive Large Scale Media Analytics. Experimented with several representations exploring temporal simultaneity of single movie (left figure below) as well as visual representations showing entire movie collections (right figure below).

14.8 Publications

Papers

None to report this quarter.

14.9 Metrics

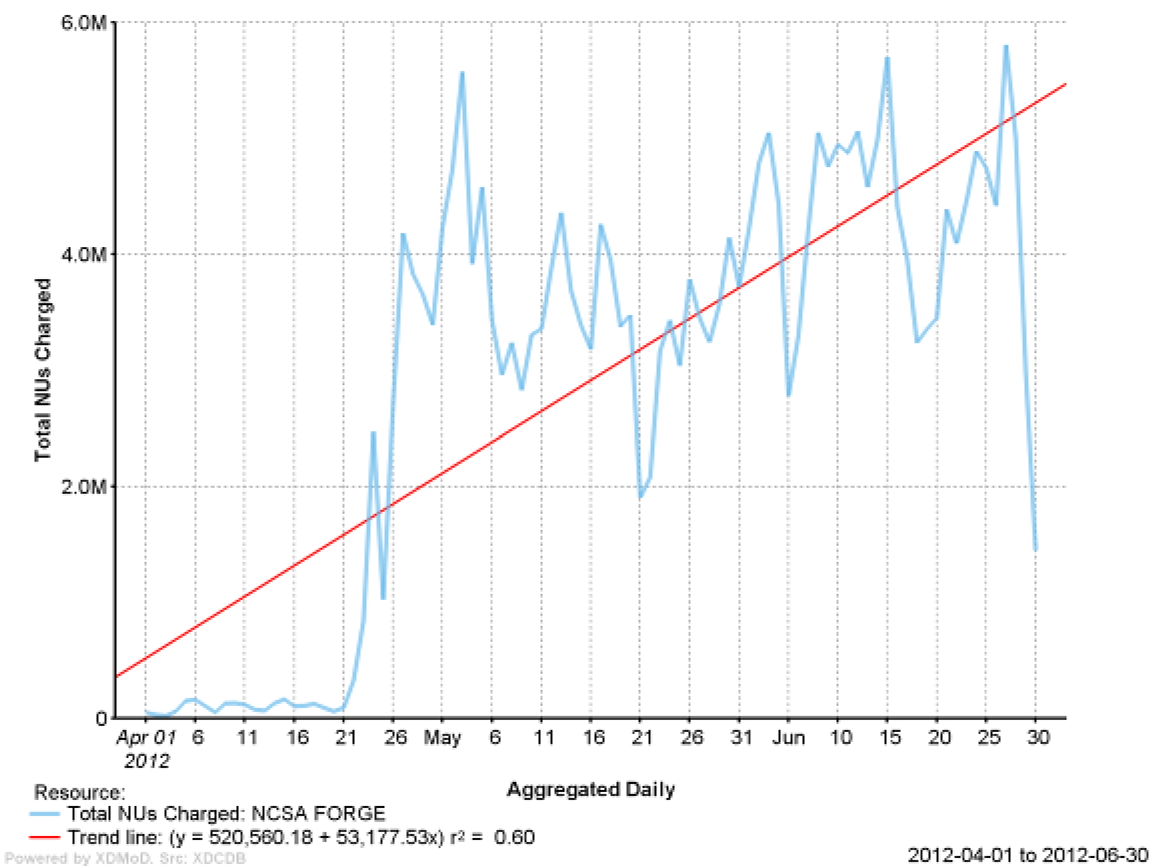
NCSA Ticket resolution times by category from XSEDE Ticket System

Time to Resolution	account issues	file systems	grid software	jobs/batch queues	login/access issues	mss/data issues	network issues	software/apps	system issues	other
0-1 hr									2	
1-24 hr						1			5	
1-7 d				1	1				3	1
1-2 wk		1				1			18	2
> 2 wk				3		6			24	2
Still Open						1		2	10	

Total NUs Charged by Resource

Resource = NCSA-FORGE

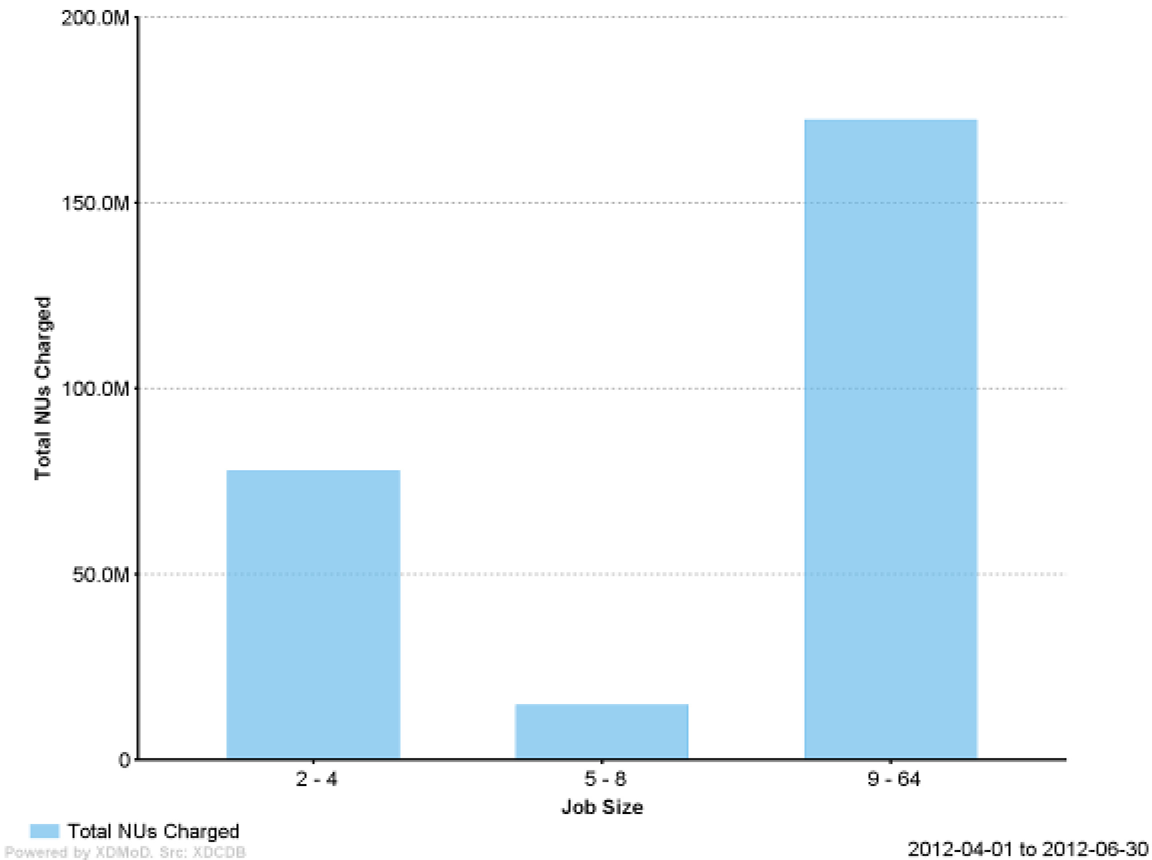
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = NCSA-FORGE

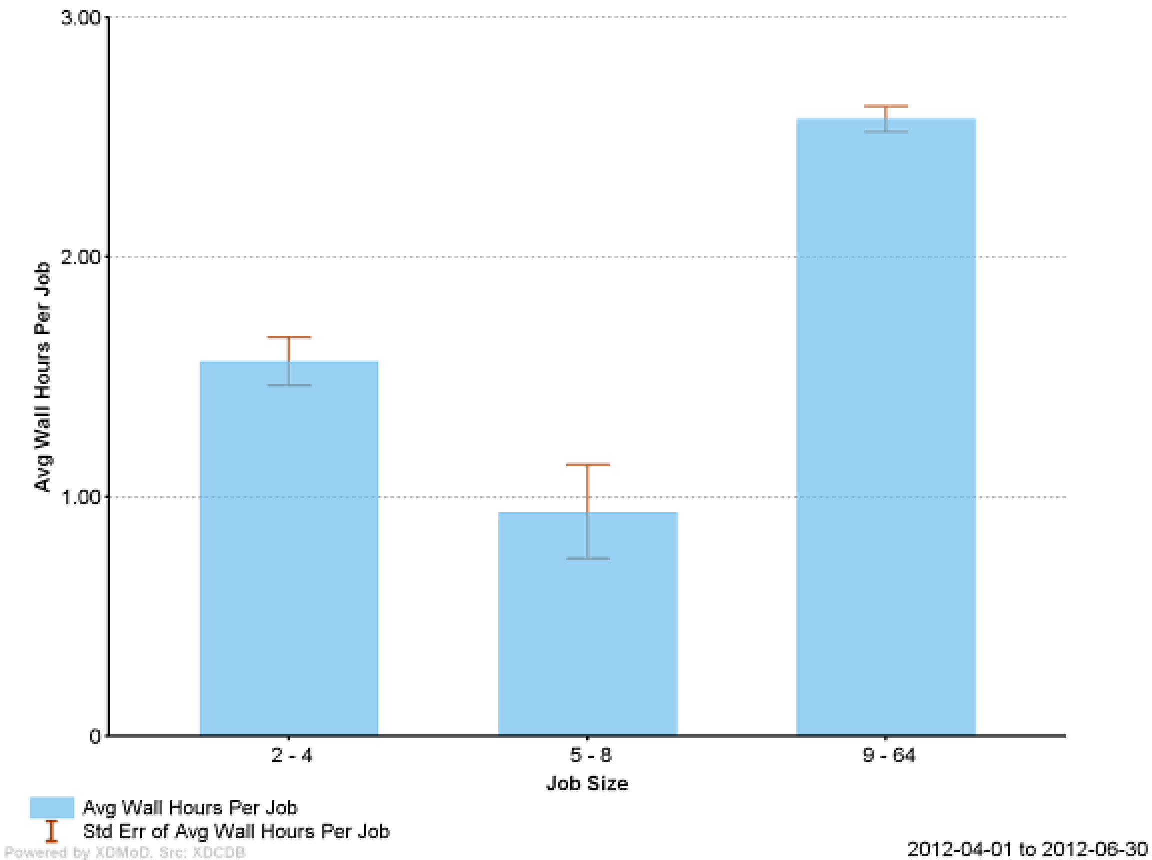
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = NCSA-FORGE

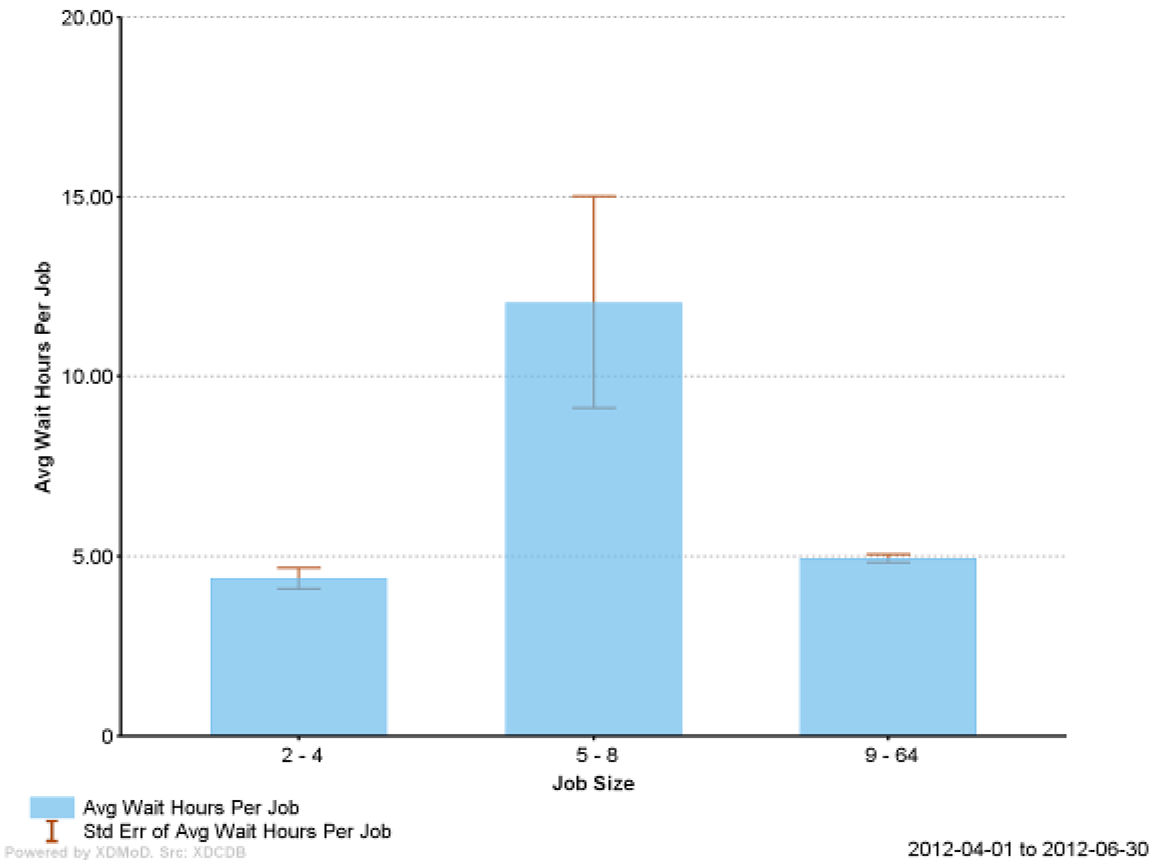
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = NCSA-FORGE

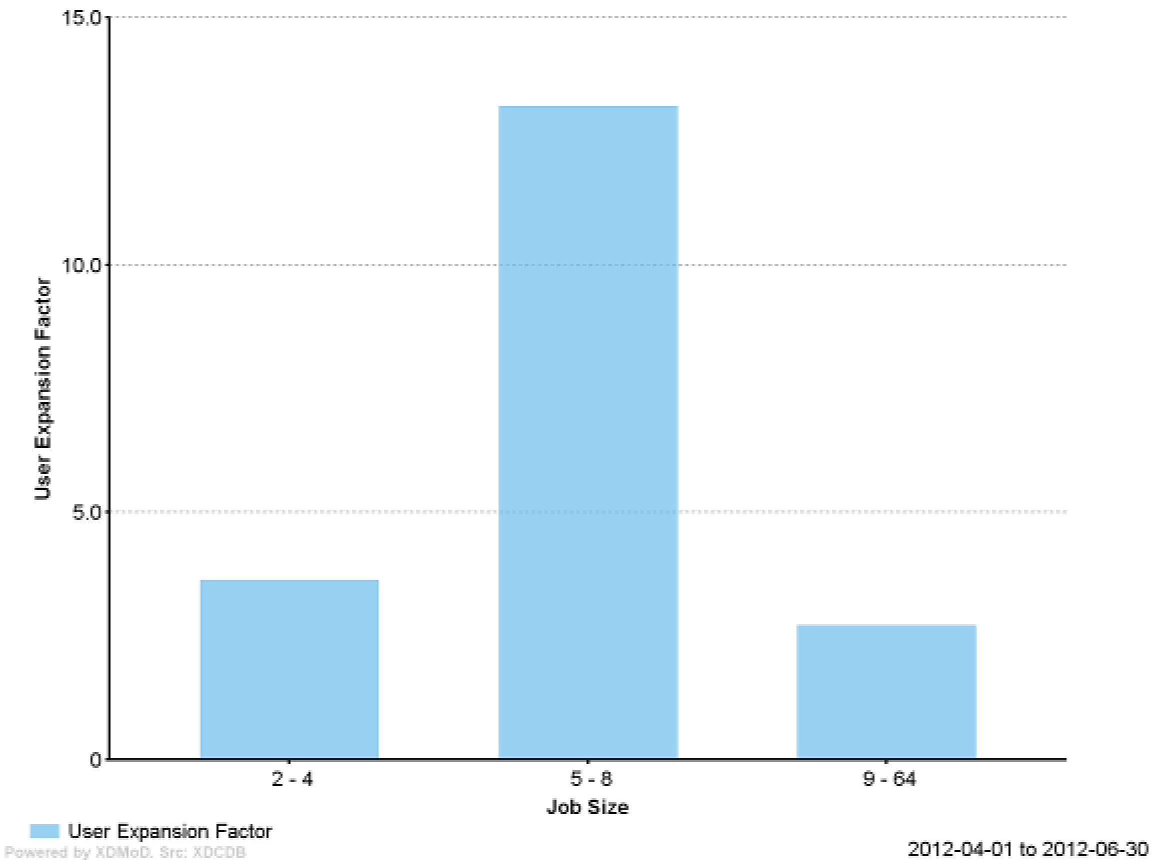
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = NCSA-FORGE

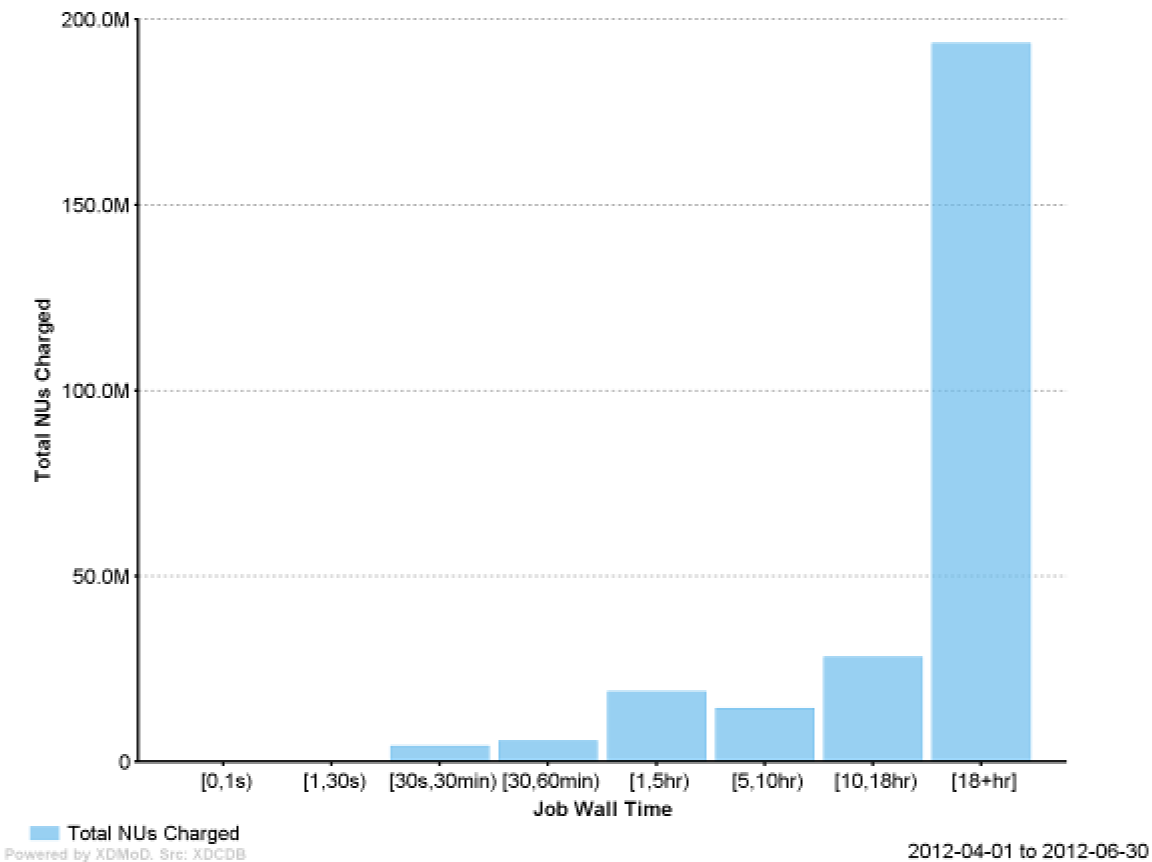
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = NCSA-FORGE

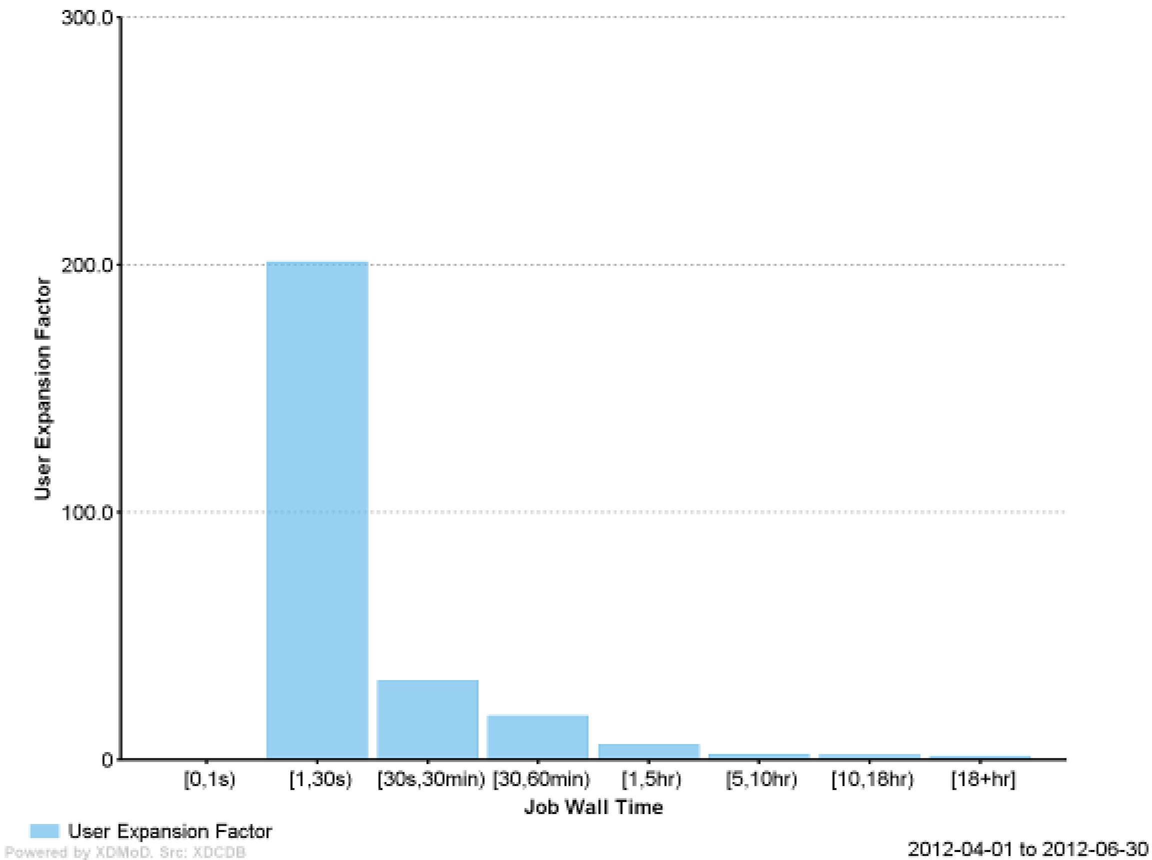
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = NCSA-FORGE

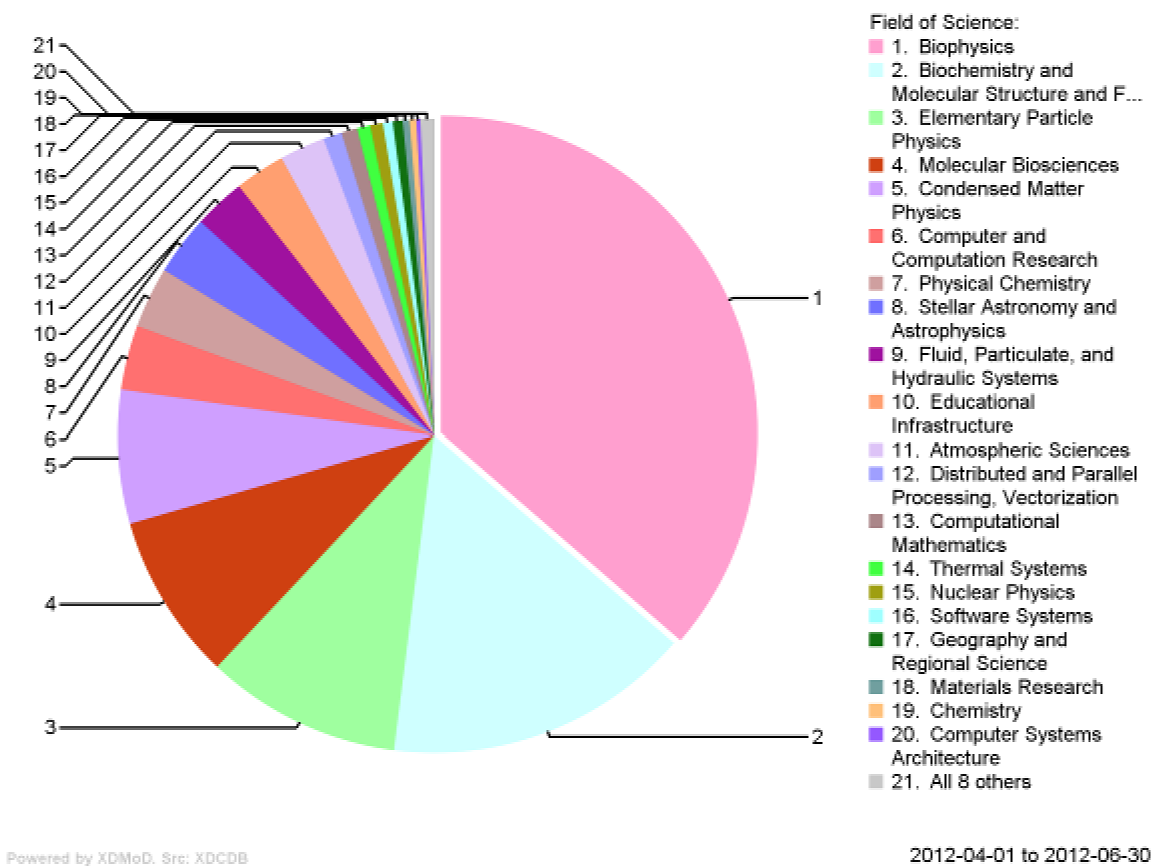
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = NCSA-FORGE

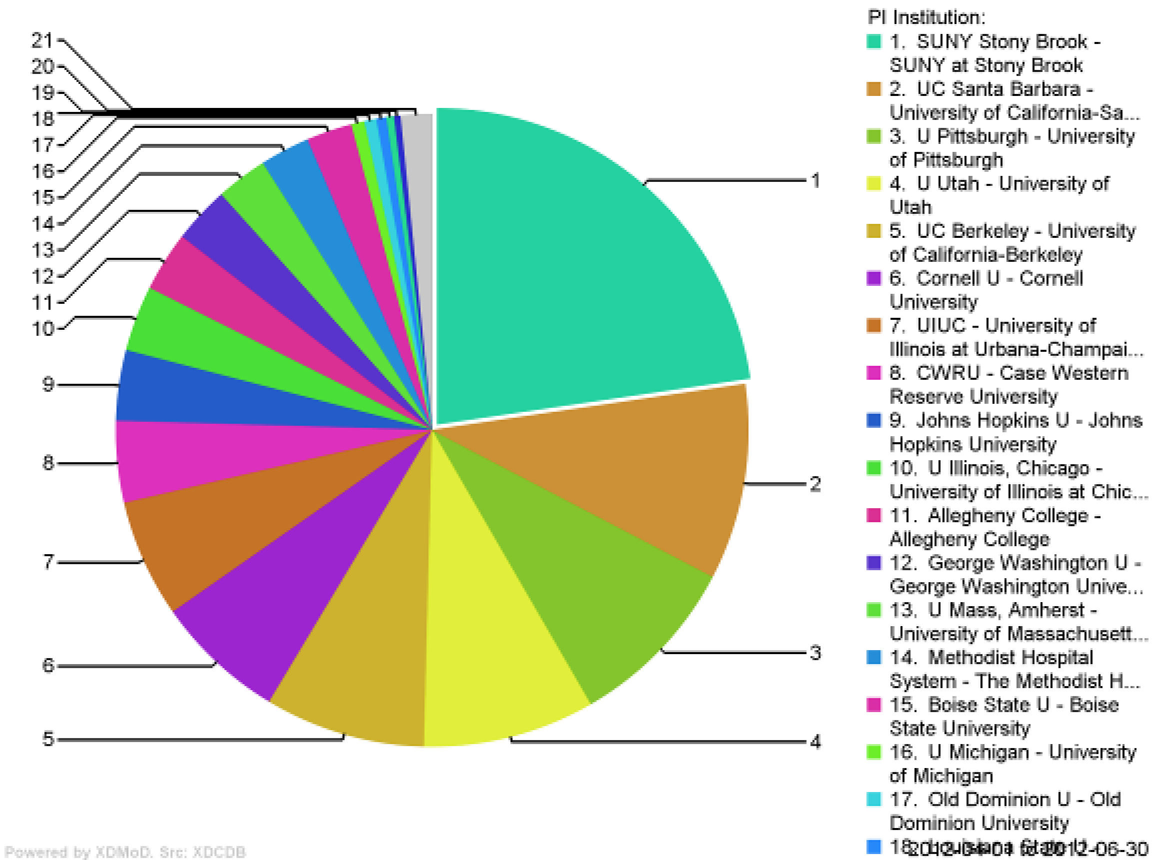
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = NCSA-FORGE

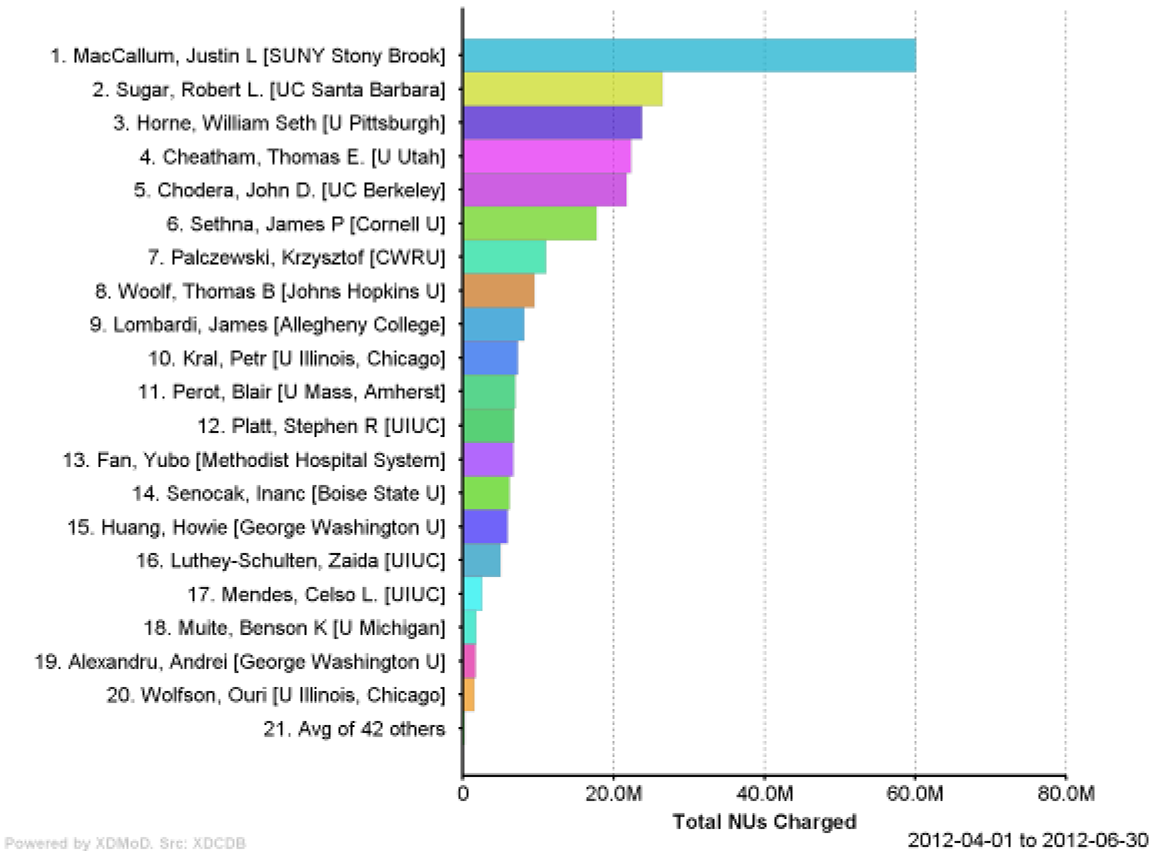
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = NCSA-FORGE

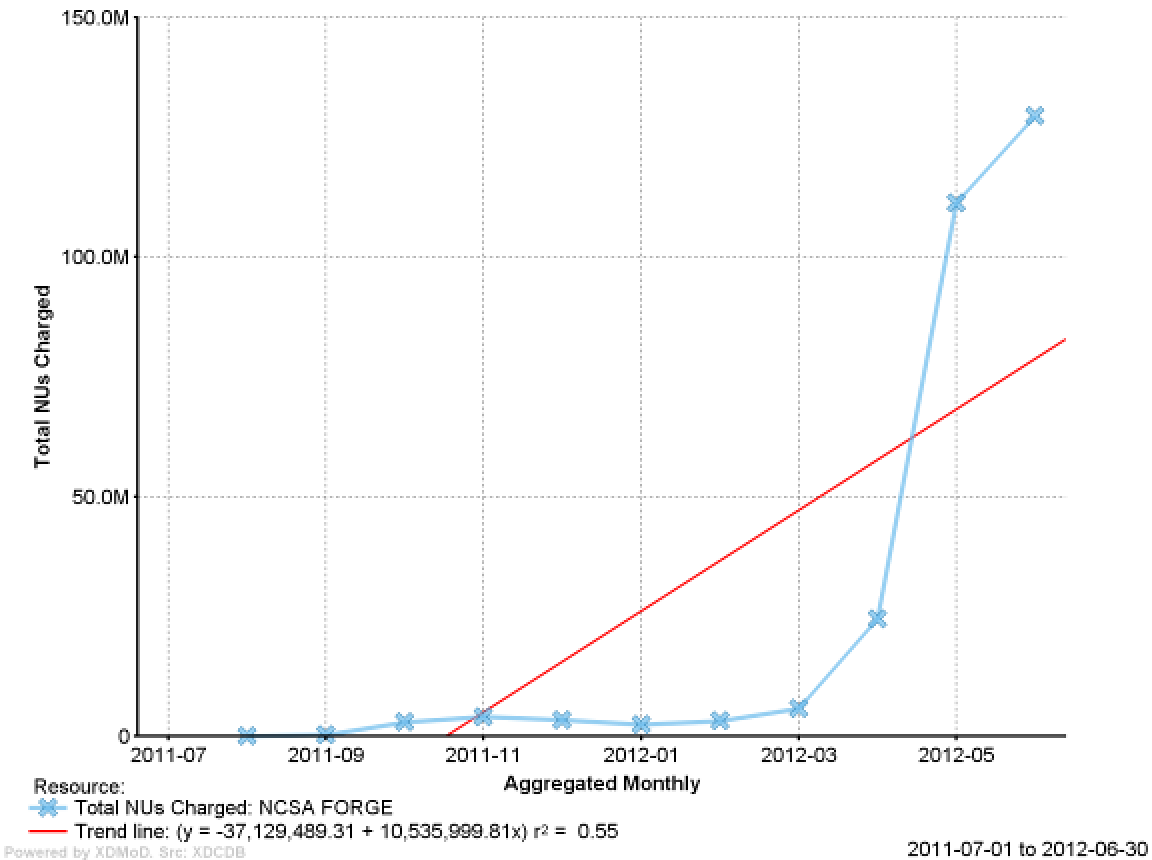
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = NCSA-FORGE

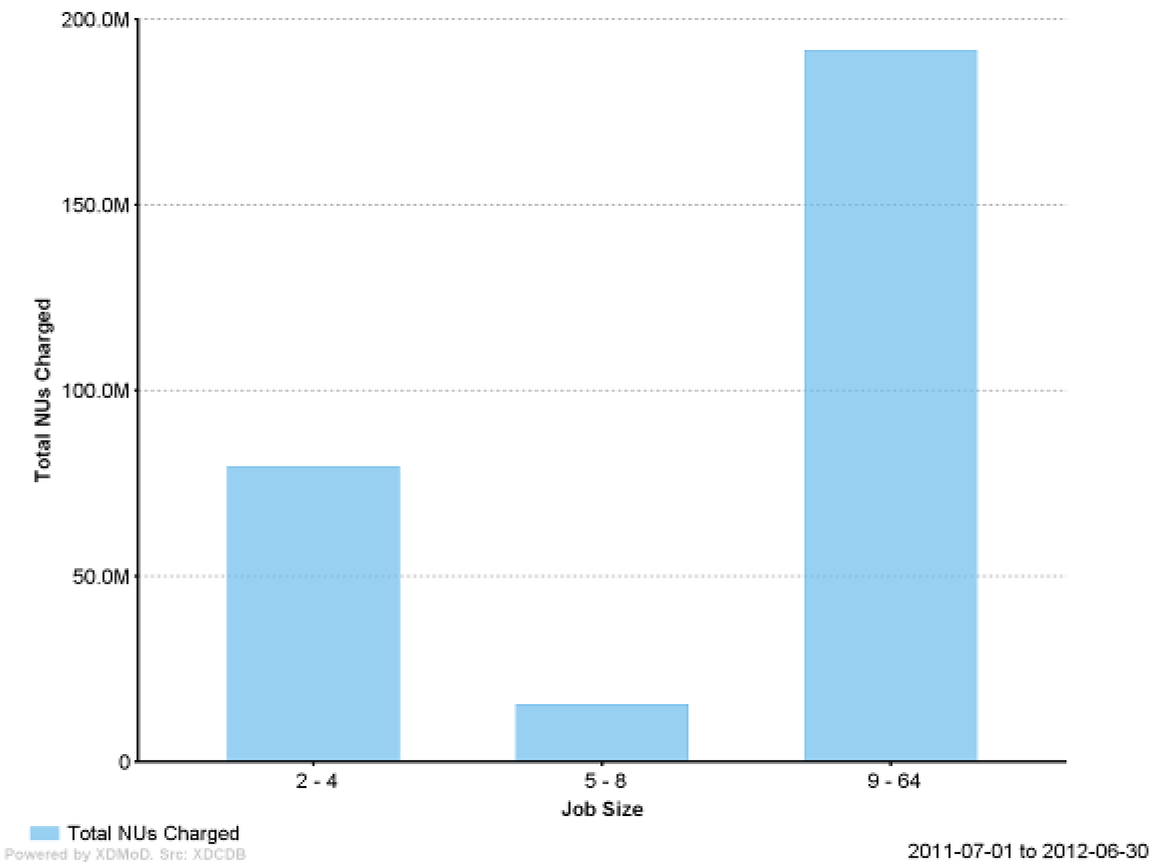
2011-07-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = NCSA-FORGE

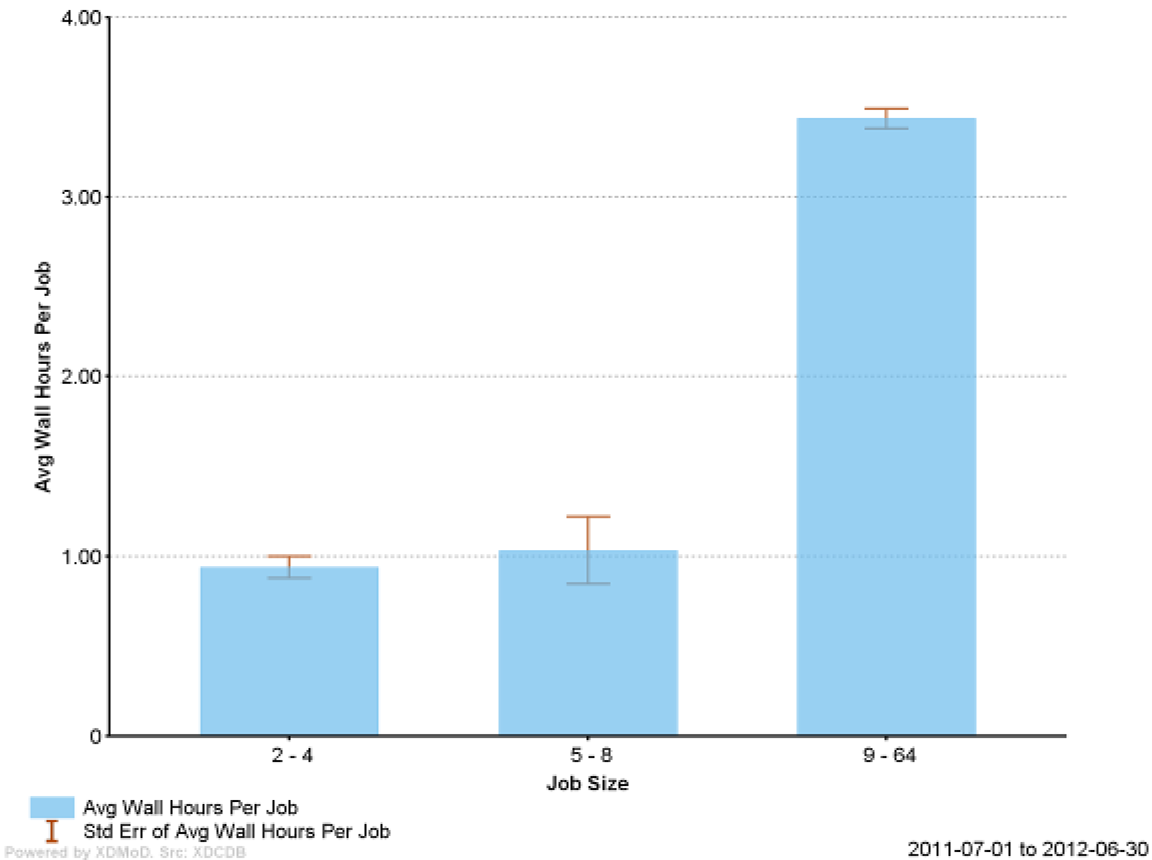
2011-07-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = NCSA-FORGE

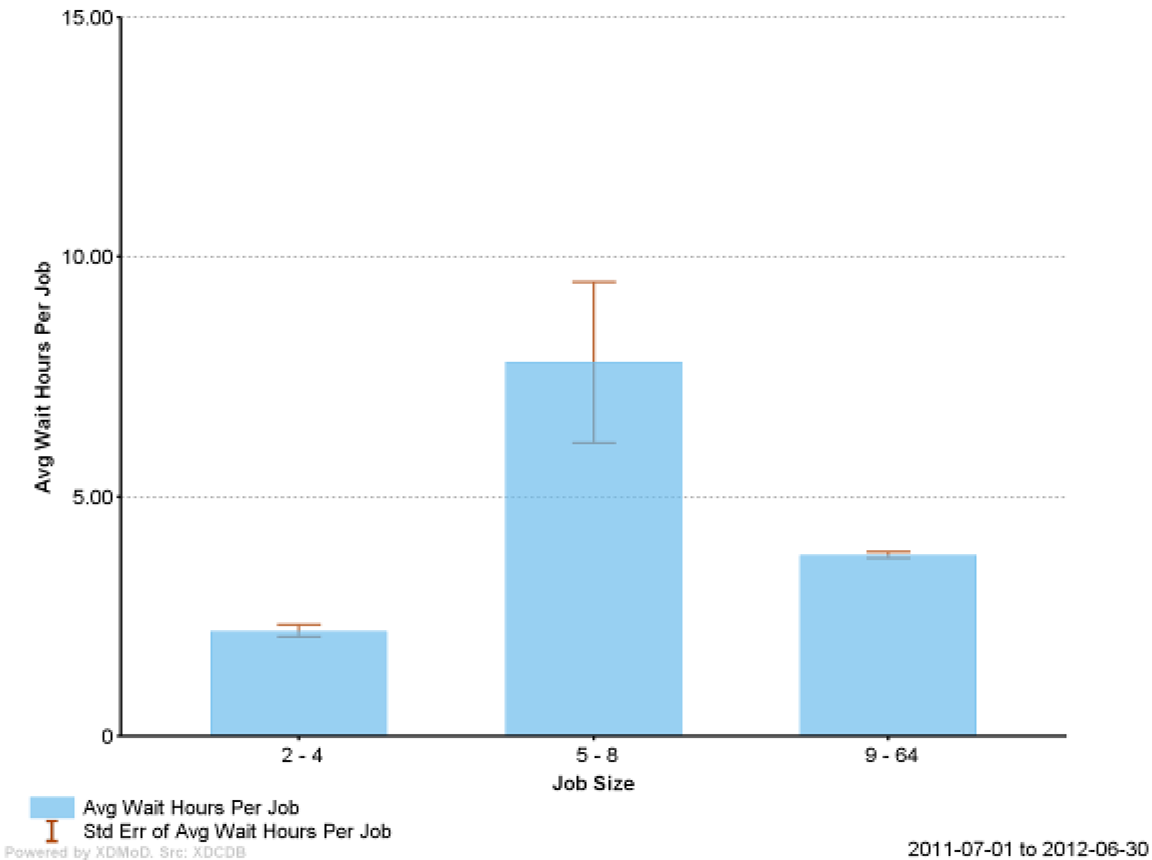
2011-07-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = NCSA-FORGE

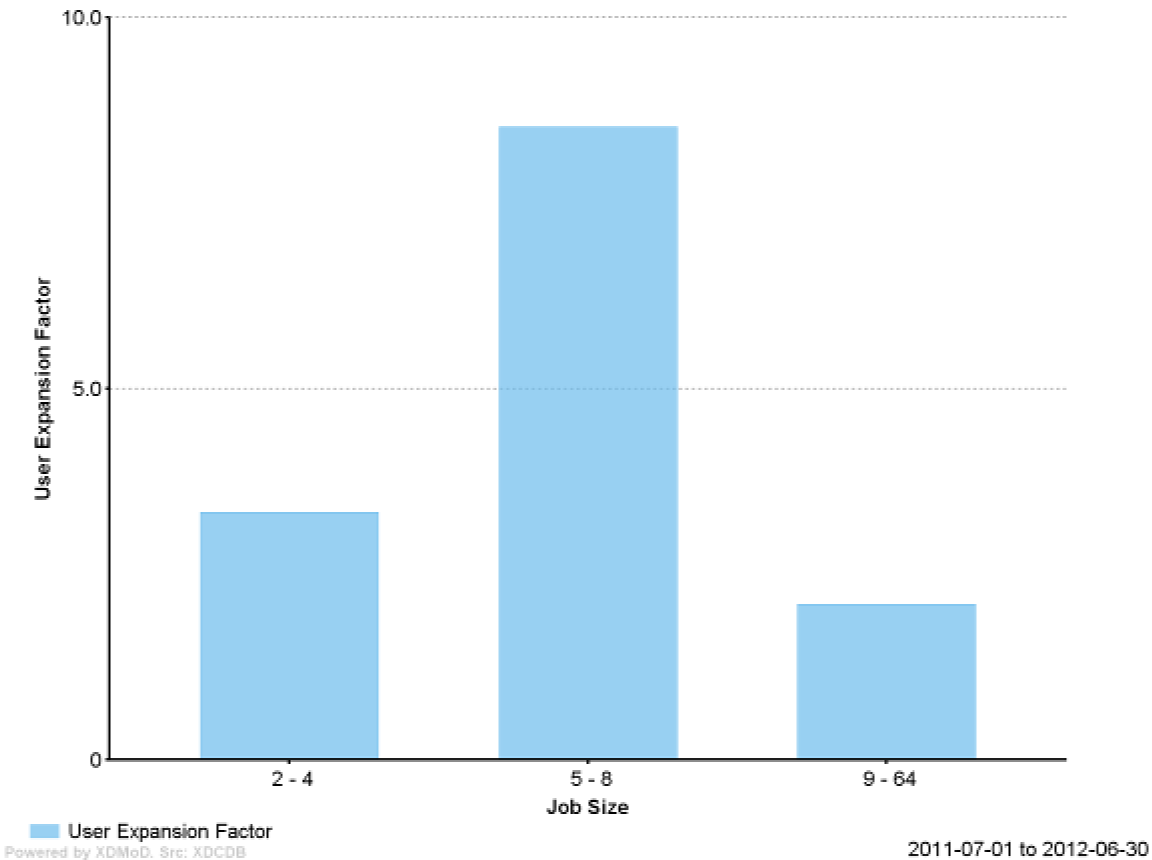
2011-07-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = NCSA-FORGE

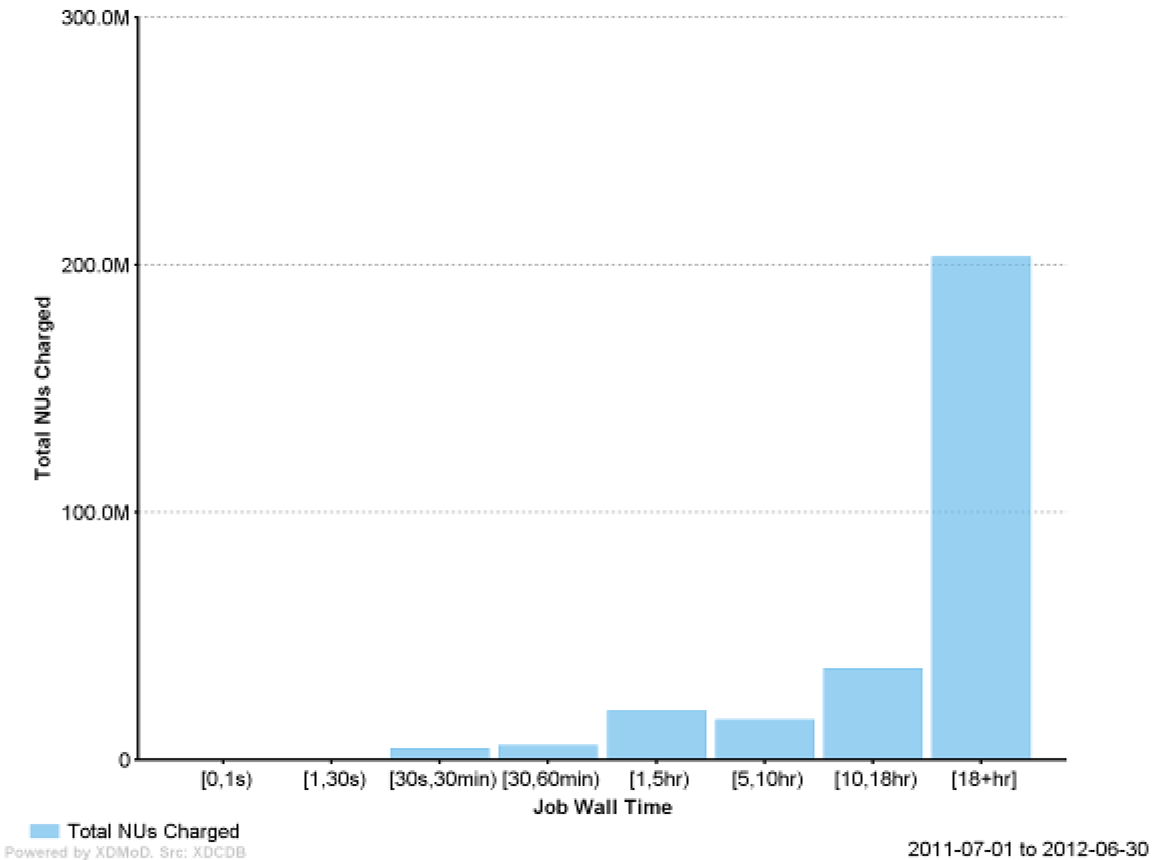
2011-07-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = NCSA-FORGE

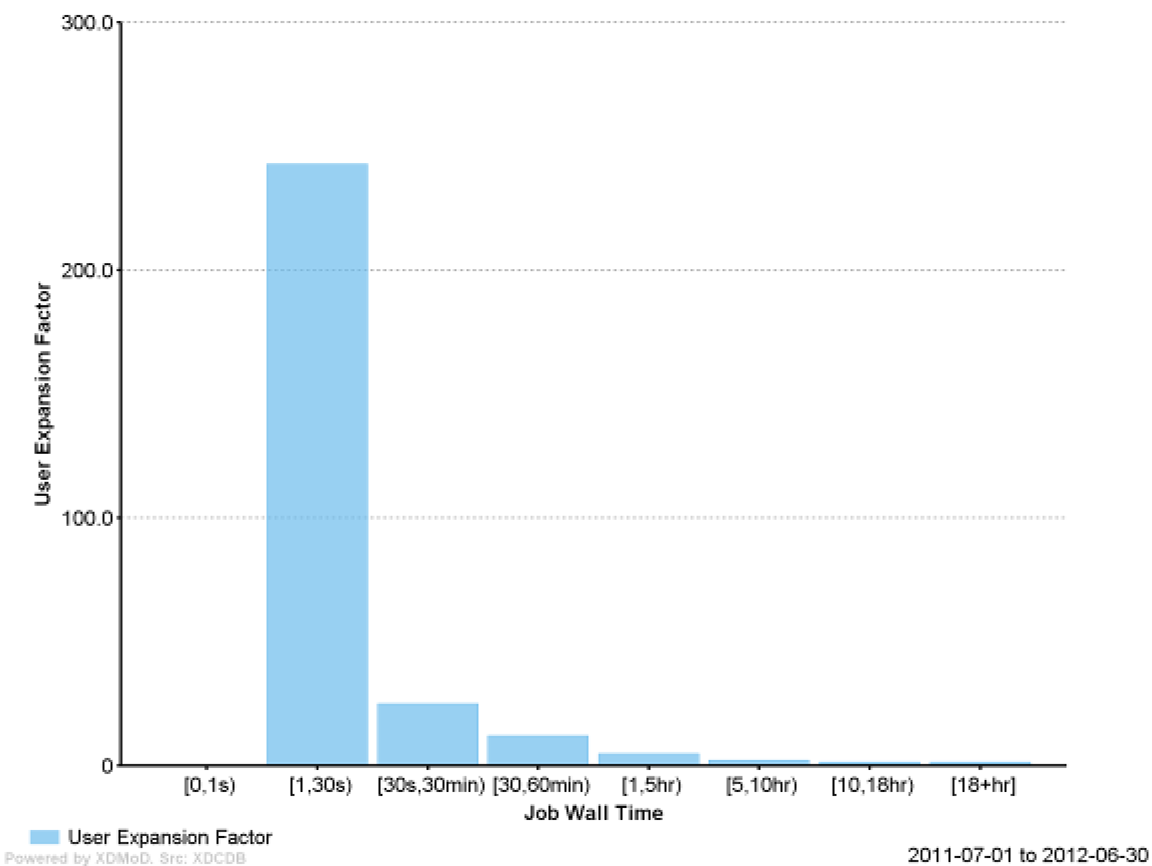
2011-07-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = NCSA-FORGE

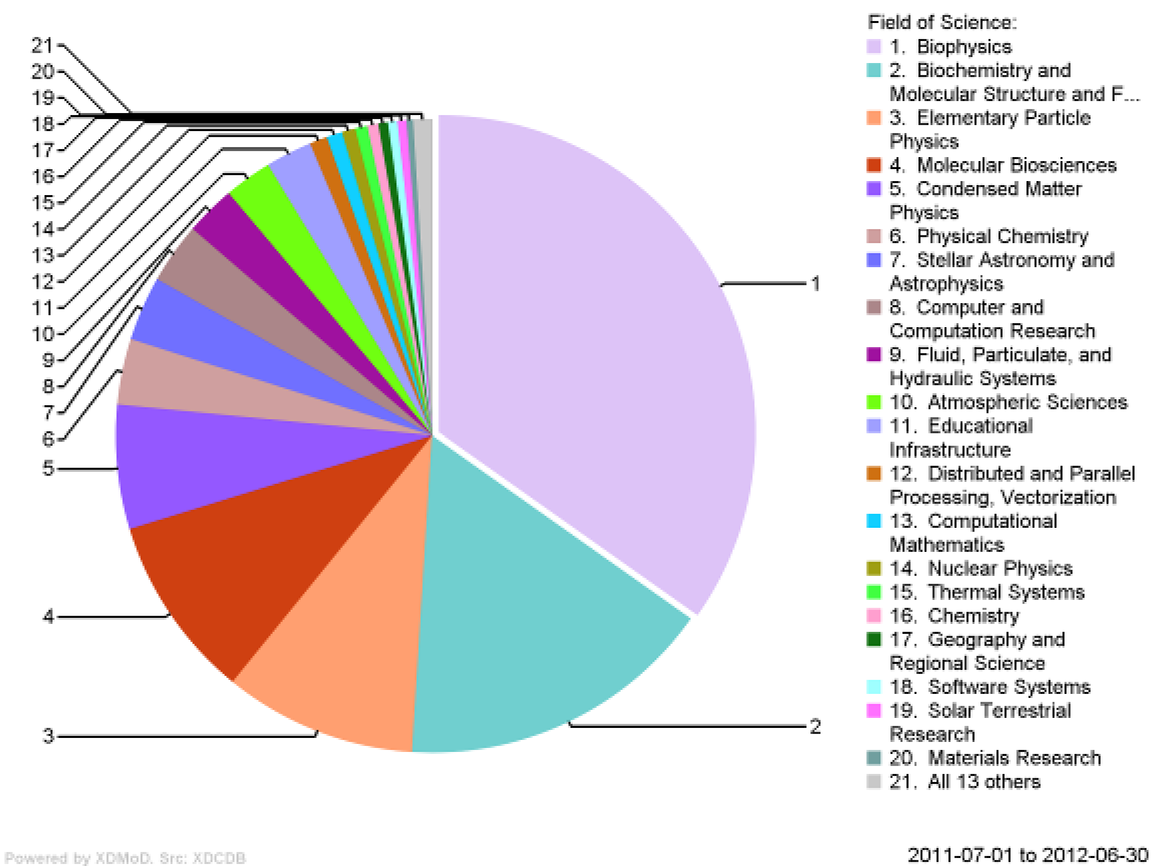
2011-07-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = NCSA-FORGE

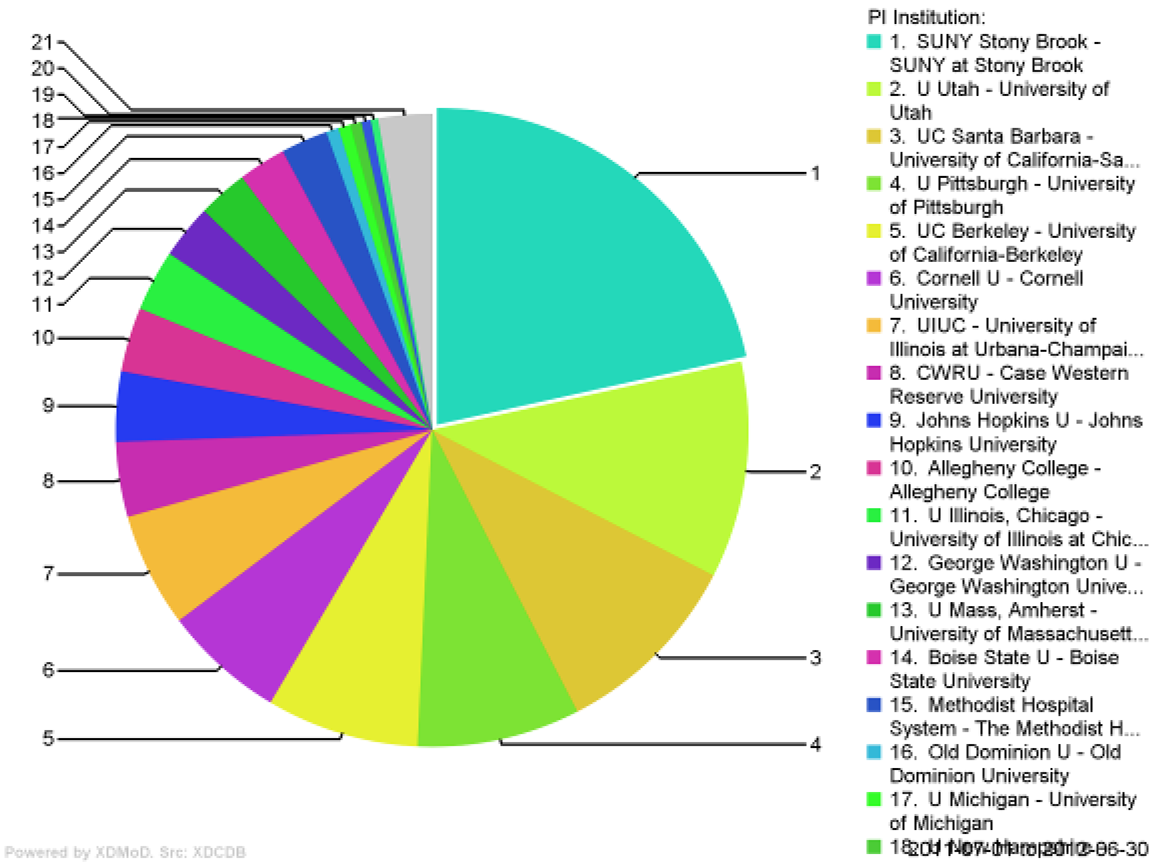
2011-07-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = NCSA-FORGE

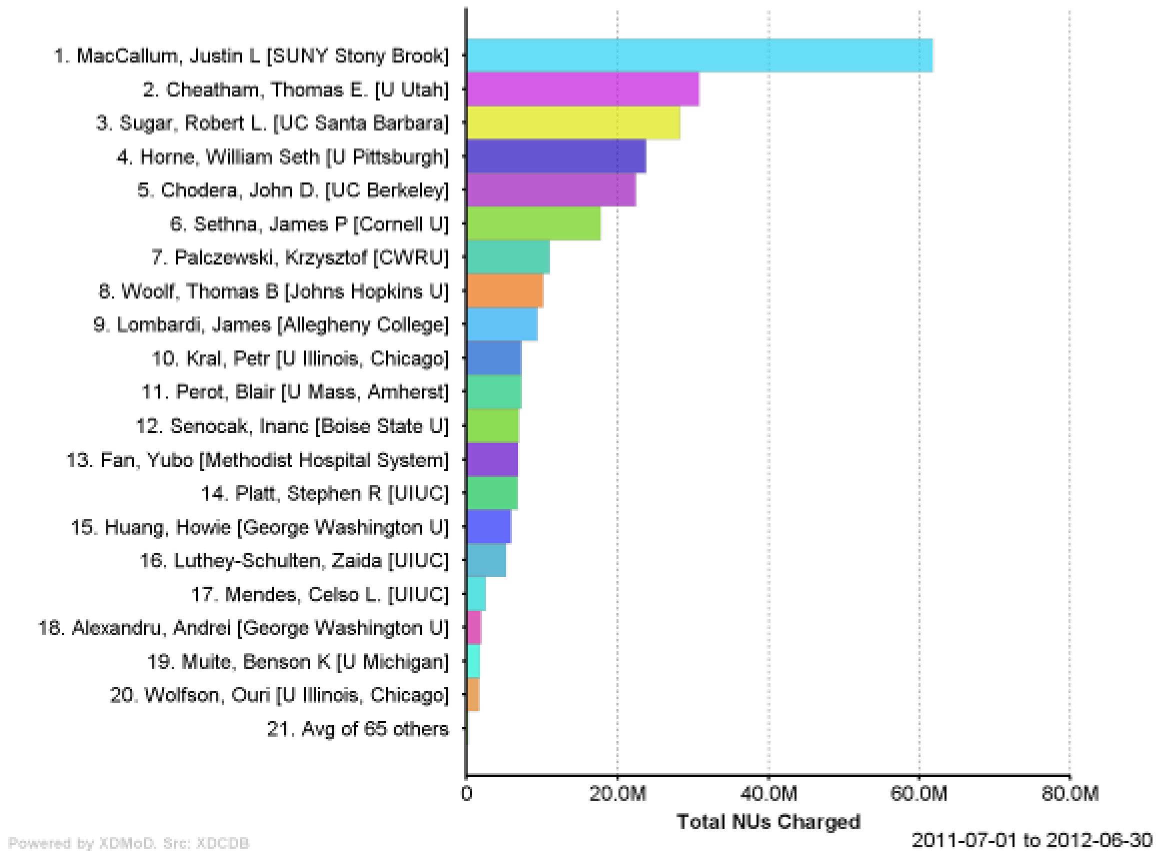
2011-07-01 to 2012-06-30



Total NUs Charged by PI

Resource = NCSA-FORGE

2011-07-01 to 2012-06-30



15 NICS – 2012 Annual Report

15.1 Executive Summary

In 2009, the National Institute for Computational Sciences (NICS) delivered the first academic petaflop computer to the NSF community—a Cray XT5 called Kraken (OCI-0711134). By the end of 2010 two Cray systems at NICS, Kraken and the smaller Cray XT4 Athena, were primary providers of computer time to the TeraGrid, delivering more than 70% of all NSF compute cycles. In the first XSEDE project year, the Cray XT5 continued to add impressive milestones to a fabled history. Not only did **Kraken deliver its 2 billionth CPU hour** to open science in November 2011, but the computer also sustained a **utilization of 94%** (Figure 27) for the year and was **available to users 98%** of the time with several periods of over 30 days of continuous availability.

On July 25, 2011 NICS decommissioned the Athena system (a 166 TF Cray XT4). In its last year of service, Athena was available for use 99% of the time and had 93% system utilization. For the month of production that fell within this XSEDE project year, Athena was available 99.8% of the time and experienced 92% utilization (Figure 27).

Under the OCI-0906324 award, the SGI Altix called Nautilus and the Remote Data and Visualization (RDAV) center has served to broaden the services provided by NICS to the NSF community and increase the potential for breakthrough science. RDAV's purpose is to aid in the significant challenge of transforming large-scale data into knowledge and insight by providing scientists with well-engineered and well-supported remote visualization, analysis, and scientific workflow technologies. For PY1, Nautilus was available 92.6% of the time with an average utilization of 47.8% (Figure 27). Nautilus-enabled science, operational statistics, collaborations, outreach, and training have been detailed in the RDAV annual report that was submitted on June 1st, 2012.

This year ended with 671 active projects containing 3,469 users with active accounts on NICS resources (Figure 24). Users of NICS resources generated 17 known publications this year with an additional 21 known works accepted for publication. NICS resources delivered 59% of all computational hours consumed by the NSF community in the first XSEDE project year (Figure 31).

Through partnerships with Oak Ridge National Laboratory, the Georgia Institute of Technology's Keeneland project, the National Center for Supercomputing Applications' Blue Waters project, and the Joint Institute for Computational Sciences' Application Acceleration Center of Excellence; NICS continued in this year to specify and evaluate leading-edge and next generation technologies while increasing XSEDE's expertise and value to the community in preparation for these technologies and associated software. At the same time, NICS disseminates knowledge and expertise through participation in the EPSCoR Track 2 project. This collaboration bridges campuses by building cyberinfrastructure (CI) linked, community specific knowledge environments that embody the desktop to XSEDE ecosystem.

There were 2,838 help tickets opened this year and support staff at NICS resolved 2,737 of these tickets during the year with a mean-time-to-resolution of just over 12.2 hours (Figure 28, Figure 29, and Figure 30). Staff at NICS provided 7 training events in this project year and participated in 11 E & O events. These events collectively reached over 600 individuals of which over 200 were from underrepresented populations. Staff at NICS also published 9 articles, had an additional 14 accepted for publication, and submitted 13 more for consideration. NICS staff also gave 16 presentations and prepared 6 posters on HPC related topics during the year. These and

other activities have been detailed in the NICS annual report that was submitted on June 29th, 2012.

15.1.1 Resource Descriptions

NICS had three NSF funded computational resources in production during the fifth project year: Kraken, Athena, and Nautilus. These systems shared a Network File System (NFS) that contains user directories, project directories and software directories. An optional one-time password tokens provide secure access to both the computational and storage resources at NICS.

15.1.1.1 Kraken

Kraken is a Cray XT5 consisting of 9,408 compute nodes, each containing two 6-core AMD Istanbul Opteron processors and 16 GB of on-node memory. The result is 112,896 compute cores that deliver 1.17 PF at peak performance with 147 TB of memory. Communications take place over the Cray SeaStar2+ interconnect. A parallel Lustre file system provides 3.3 PB (raw) of short-term data storage.

15.1.1.2 Athena

Athena was a Cray XT4 with 4512 compute nodes interconnected with SeaStar, a 3D torus. Each compute node had one four-core AMD Opteron for a total of 18,048 cores. All nodes had 4 Gbytes of memory: 1 Gbyte of memory per core. A parallel Lustre file system provided 100 TB (raw) of short-term data storage.

15.1.1.3 Nautilus

Nautilus, an SGI Altix UV 1000 system, is the centerpiece of NICS Remote Data and Visualization (RDAV) Center that is also located at ORNL. It has 1024 cores (Intel Nehalem EX processors), 4 TB of global shared memory, and 8 GPUs in a single system image yielding 8.2 TF at peak performance. A parallel Lustre file system provides 427 TB (raw) of short-term data storage.

15.1.1.4 HPSS Archival Storage

The High Performance Storage System (HPSS), developed and operated by ORNL, is capable of archiving hundreds of petabytes of data and can be accessed by all major leadership computing platforms. Incoming data is written to disk and later migrated to tape for long term archiving. This hierarchical infrastructure provides high-performance data transfers while leveraging cost effective tape technologies. Robotic tape libraries provide tape storage. The center has four SL8500 tape libraries holding up to 10,000 cartridges each. The libraries house a total of 24 T10K-A tape drives (500 GB cartridges, uncompressed), 60 T-10K-B tape drives (1 terabyte cartridges, uncompressed), and 20 T10K-C tape drives (5 terabyte cartridges, uncompressed). Each T10K-A and T10K-B drive has a bandwidth of 120 MB/s. Each T10K-C tape drive has a bandwidth of 240 MB/s. Disk storage is provided by DDN storage arrays with nearly a petabyte of capacity and over 12 GB/s of bandwidth. This infrastructure has allowed the archival system to scale to meet increasingly demanding capacity and bandwidth requirements with nearly 10 PB of NICS data stored as of May 31, 2012.

15.2 Metrics

15.2.1 Standard usage metrics

15.2.1.1 Kraken

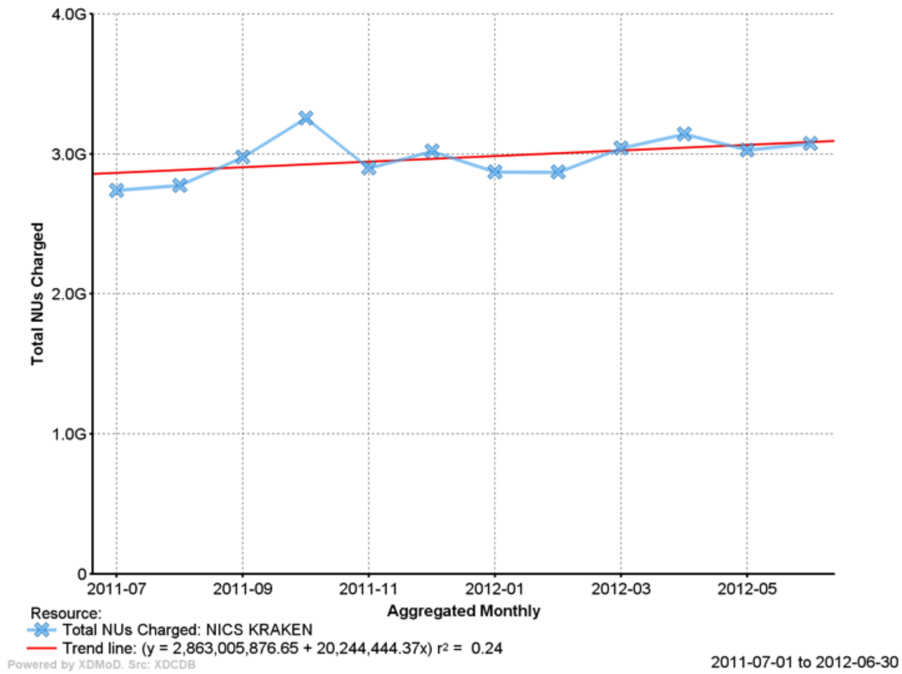


Figure 6: Total NUs charged on Kraken in XSEDE PY1.

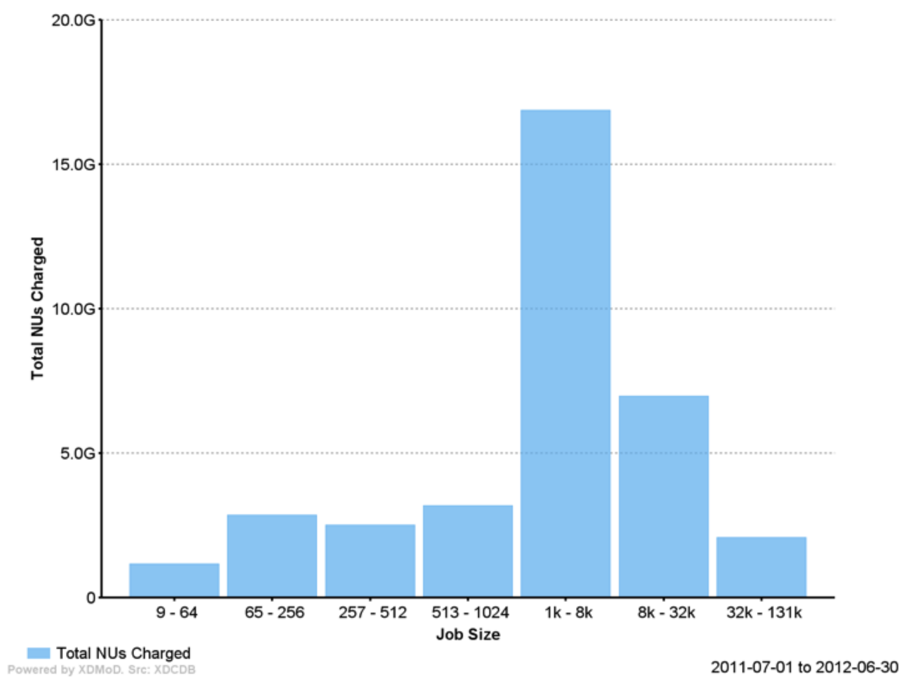


Figure 7: Total NUs charged by job size to Kraken in XSEDE PY1.

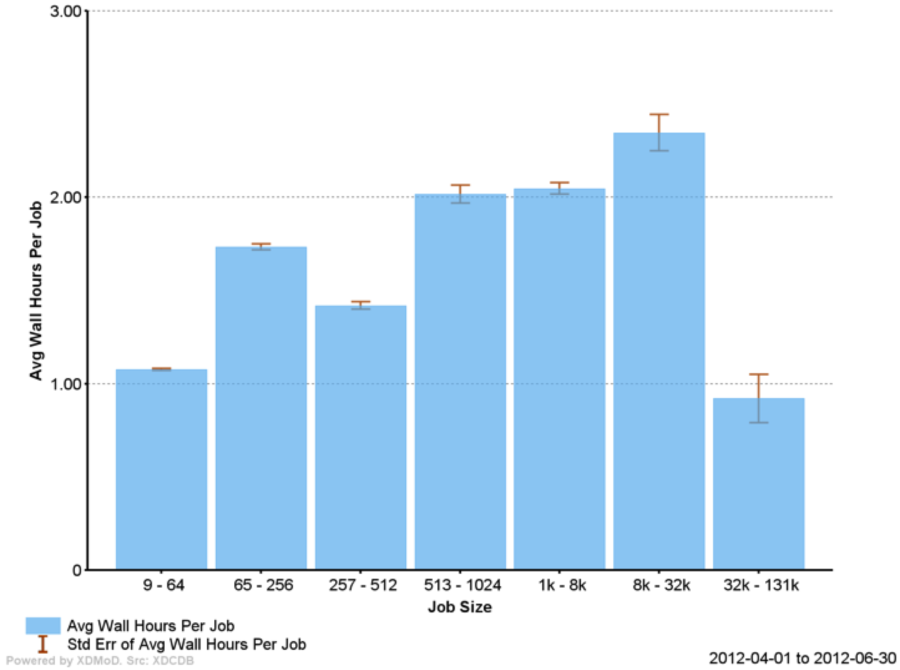


Figure 8: Average wall hours per job by job size on Kraken in PY1.

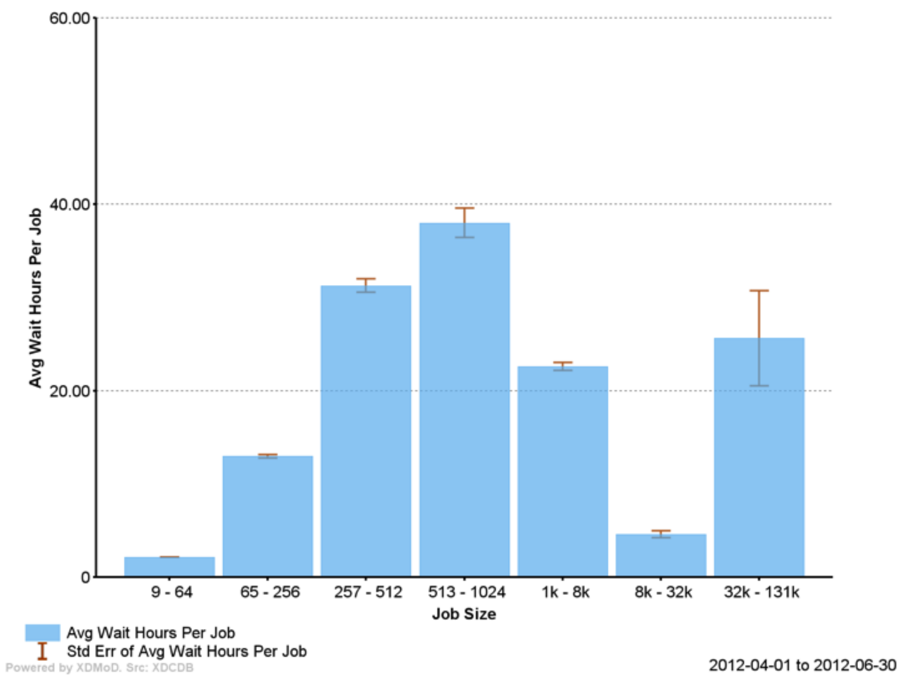


Figure 9: Average Wait Hours per job by job size on Kraken in PY1.

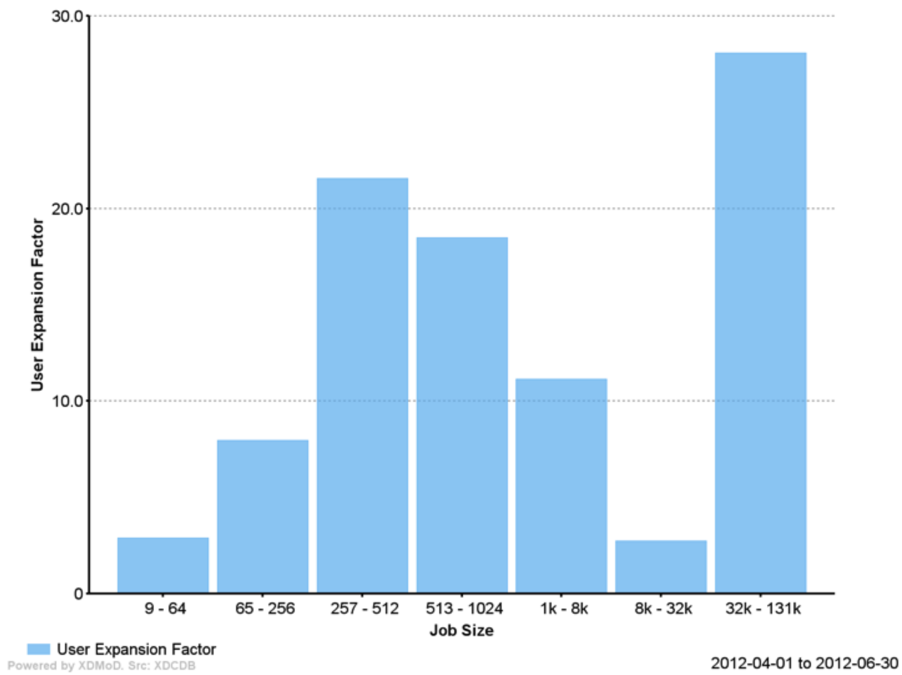


Figure 10: User expansion factor by job size on Kraken in PY1.

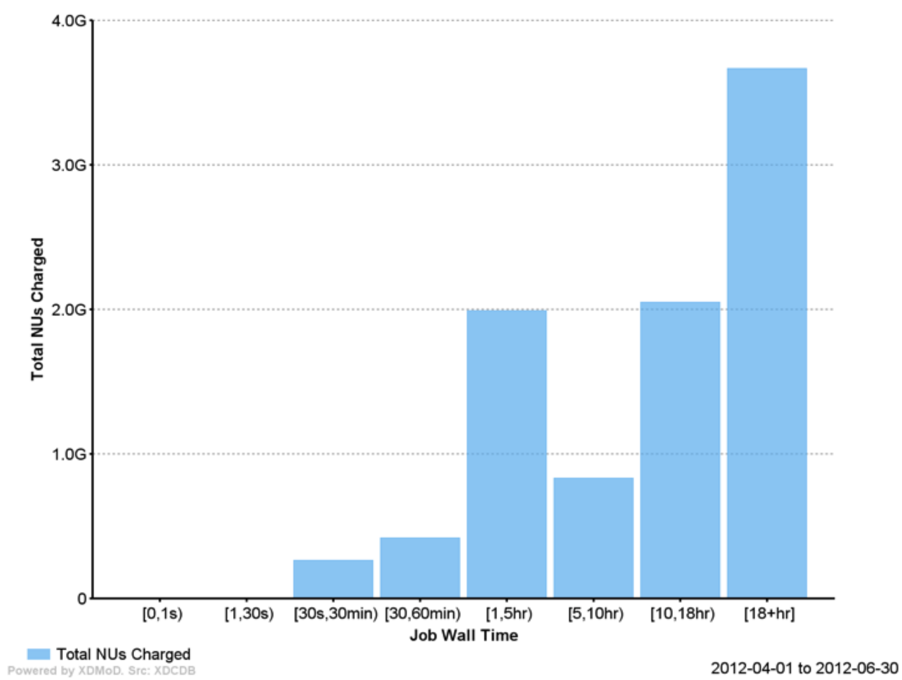


Figure 11: Total NUs charged by job wall time on Kraken in PY1.

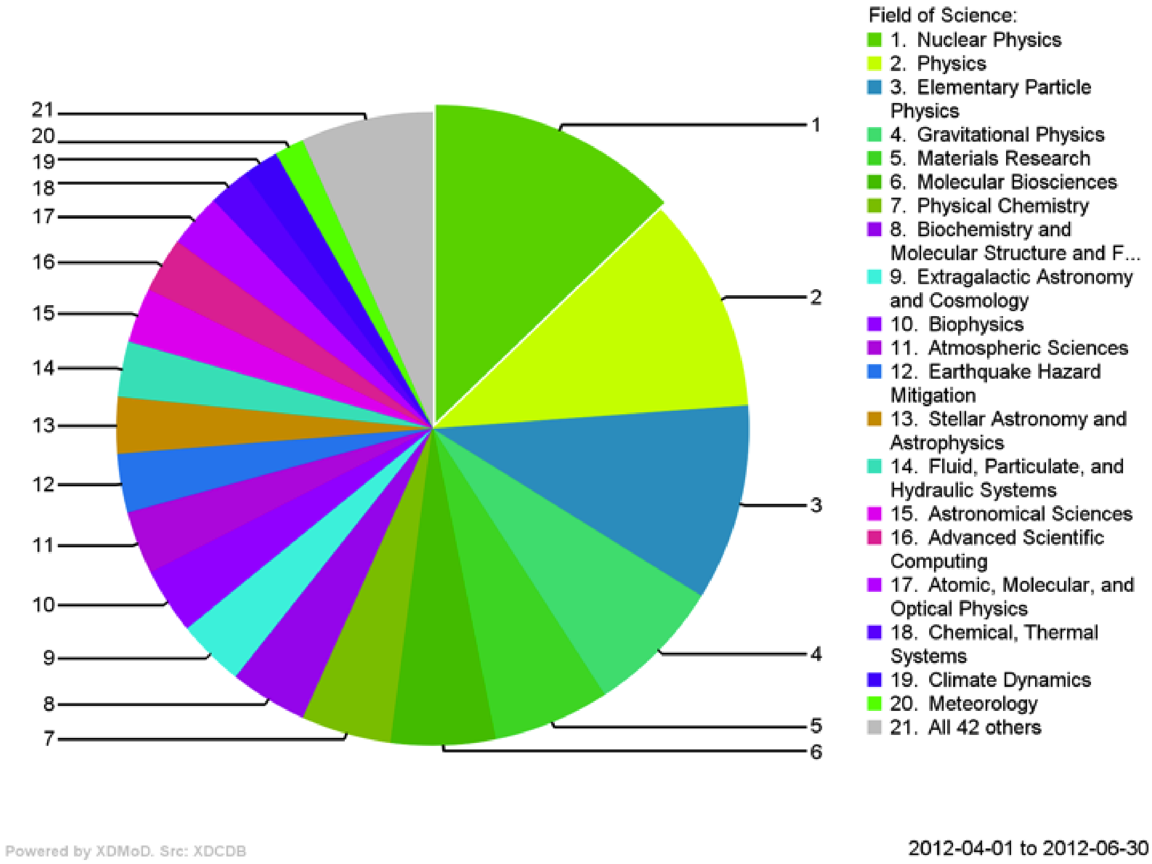


Figure 12: Total NUs charged by field of science on Kraken in PY1.

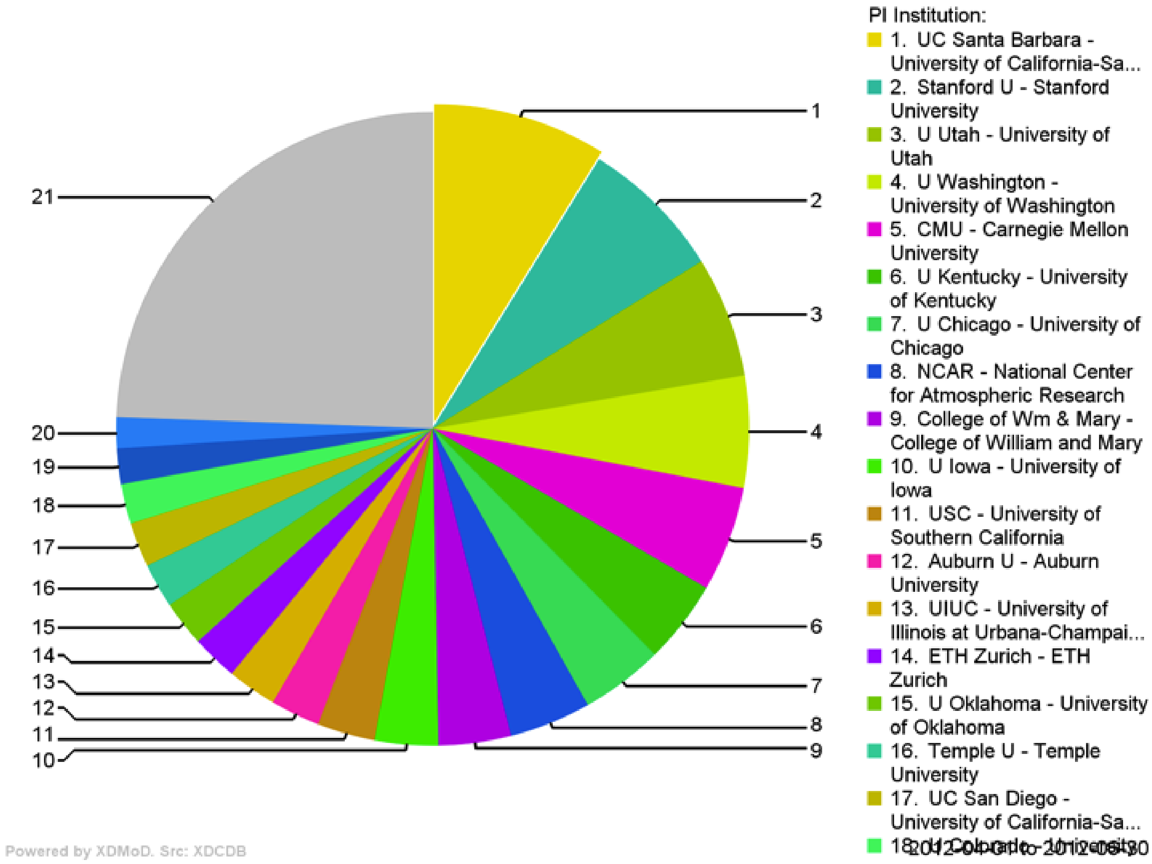


Figure 13: Total NUs charged by PI Institution on Kraken in PY1.

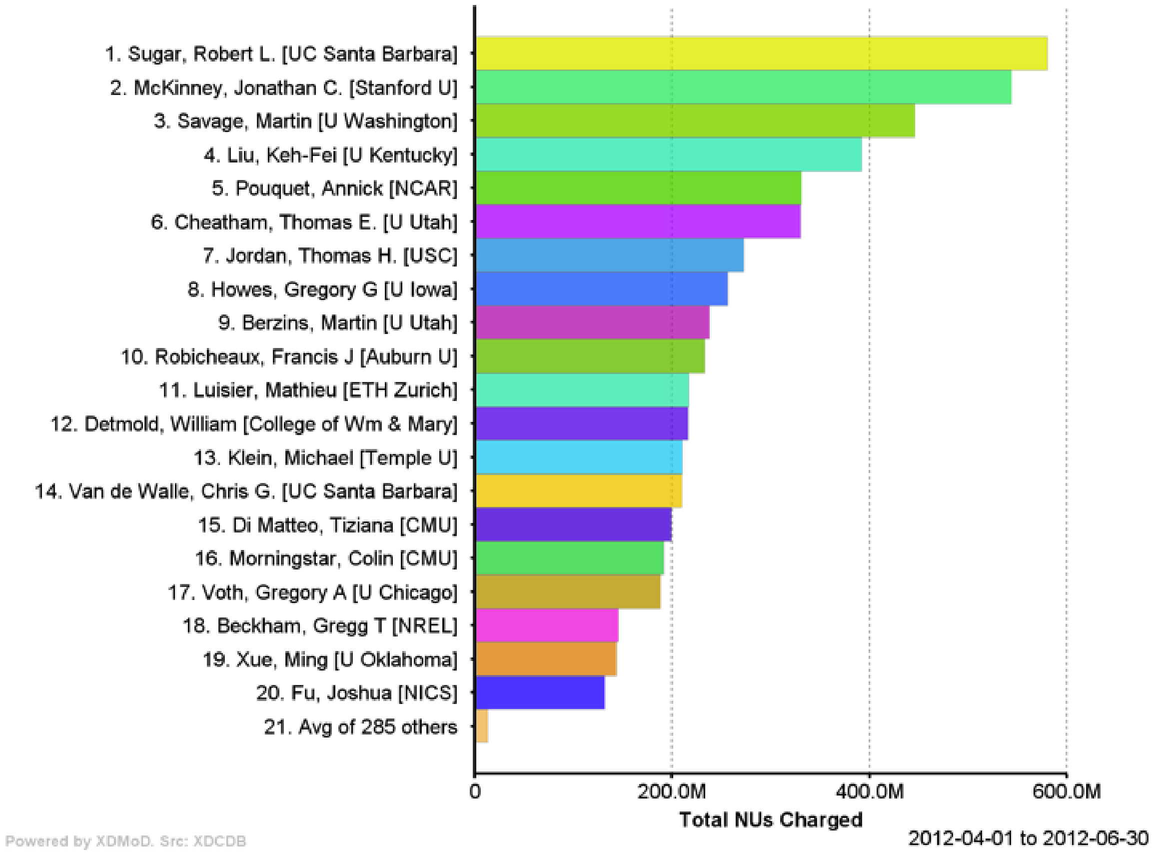


Figure 14: Total NUs charged by PI on Kraken in PY1.

15.2.1.2 Nautilus

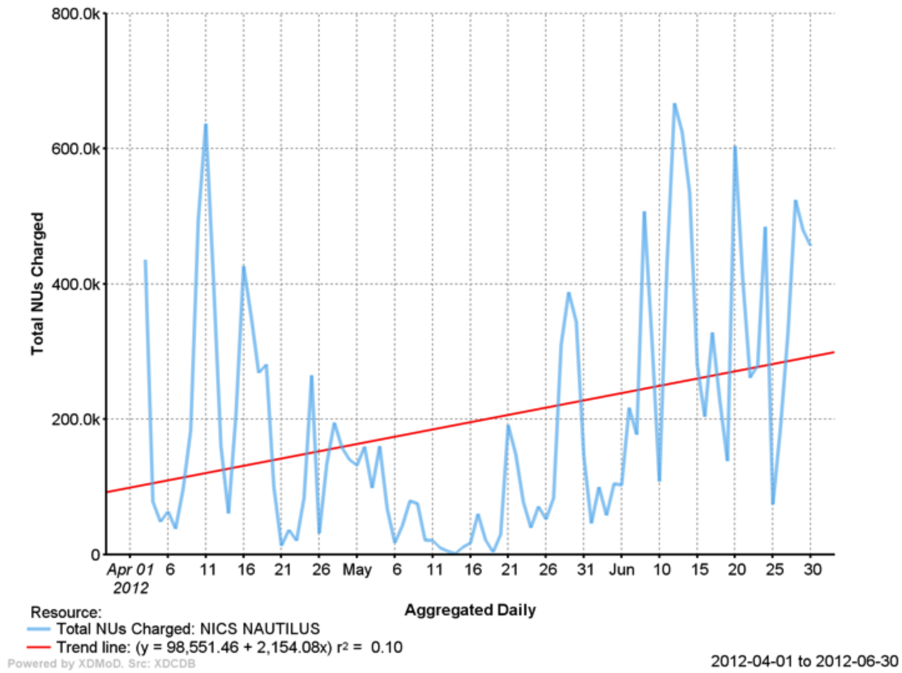


Figure 15: Total NUs charged to Nautilus in PY1.

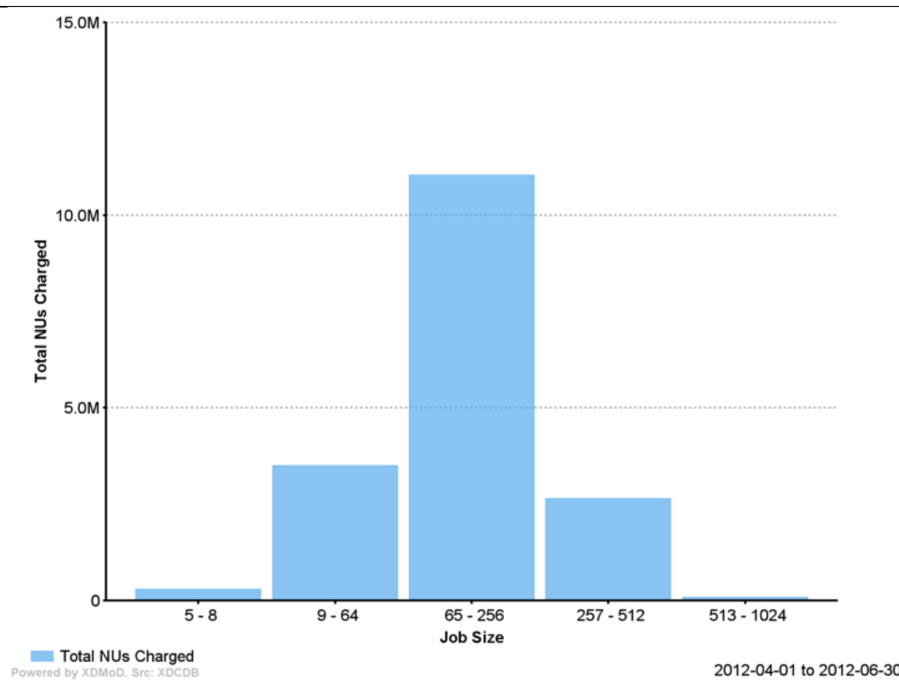


Figure 16: Total NUs charged by job size on Nautilus in PY1.

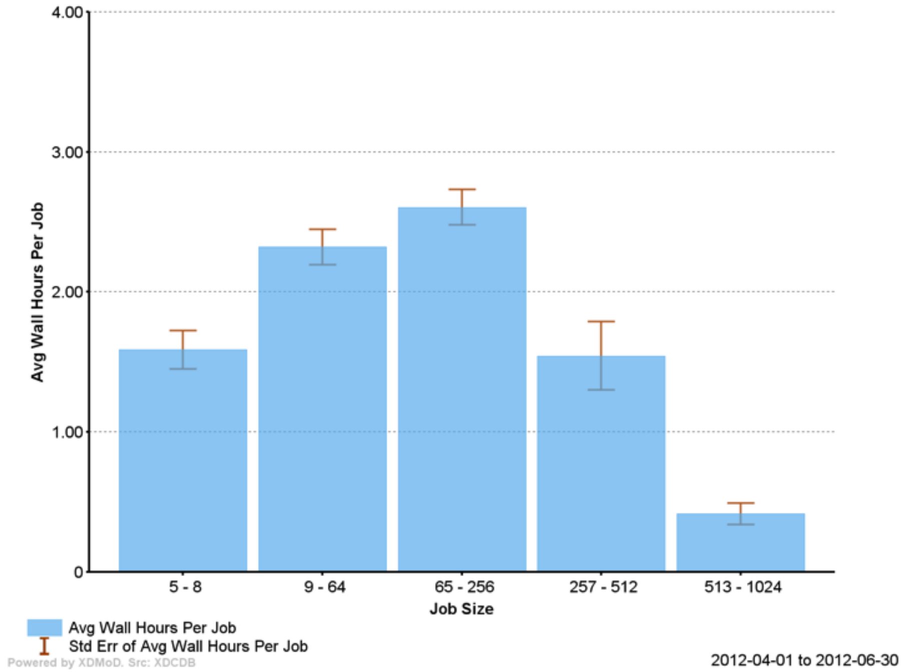


Figure 17: Average wall hours per job by job size on Nautilus in PY1.

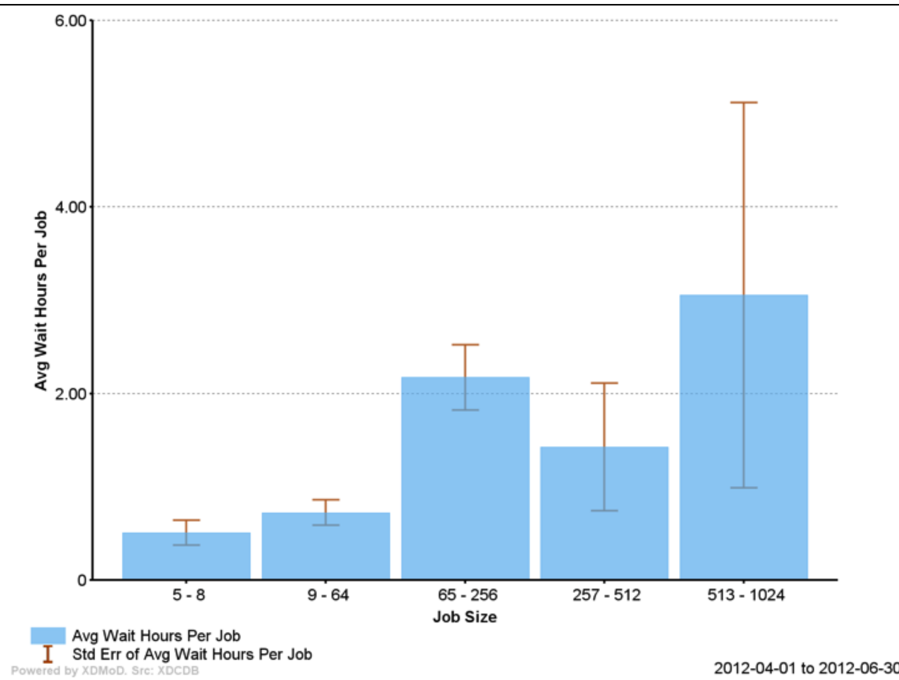


Figure 18: Average wait hours per job by job size on Nautilus in PY1.

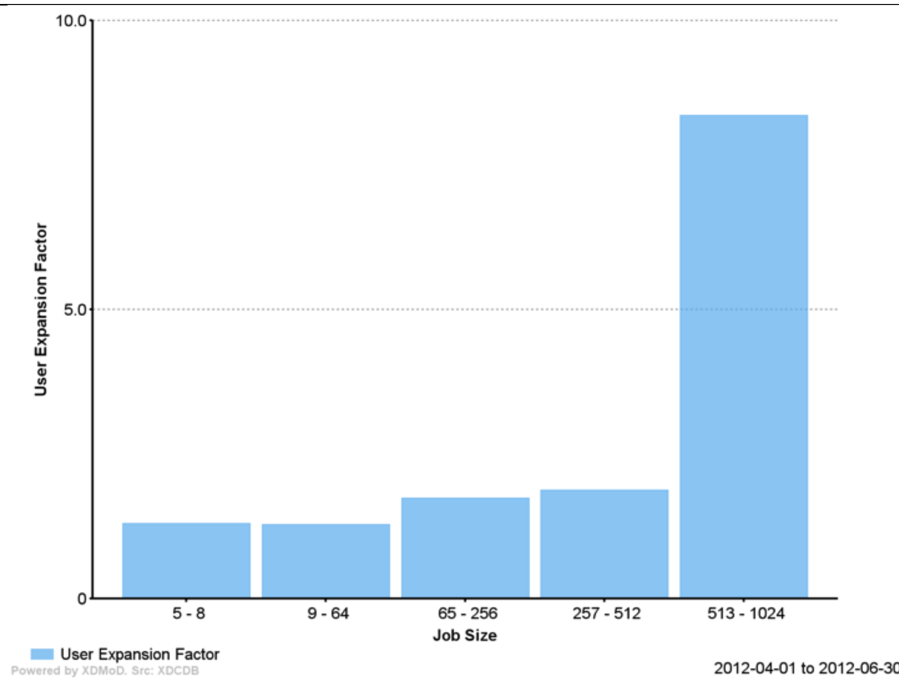


Figure 19: User expansion factor by job size on Nautilus in PY1.

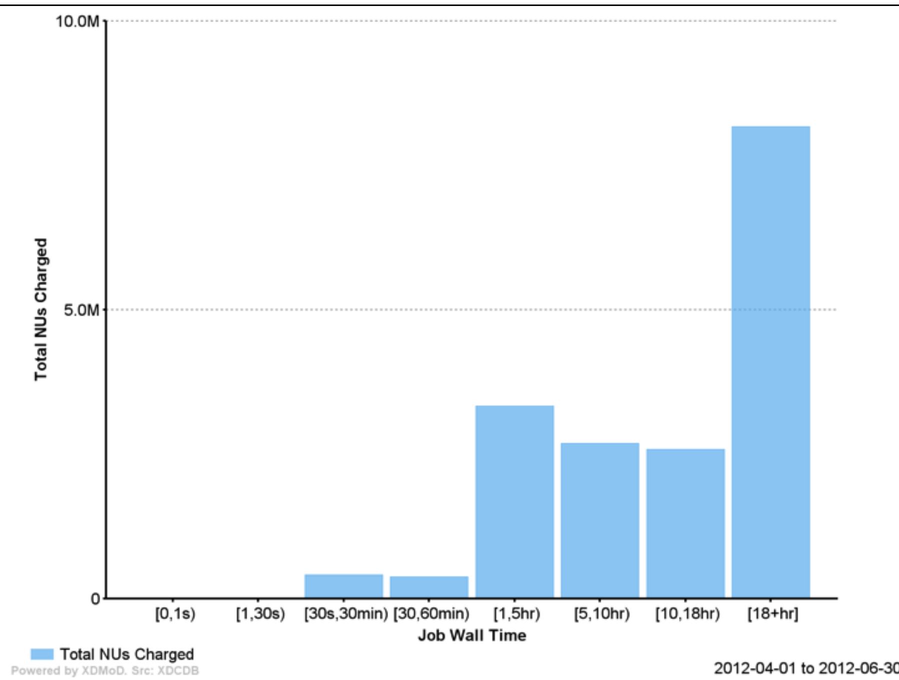


Figure 20: Total NUs charged by job wall time on Nautilus in PY1.

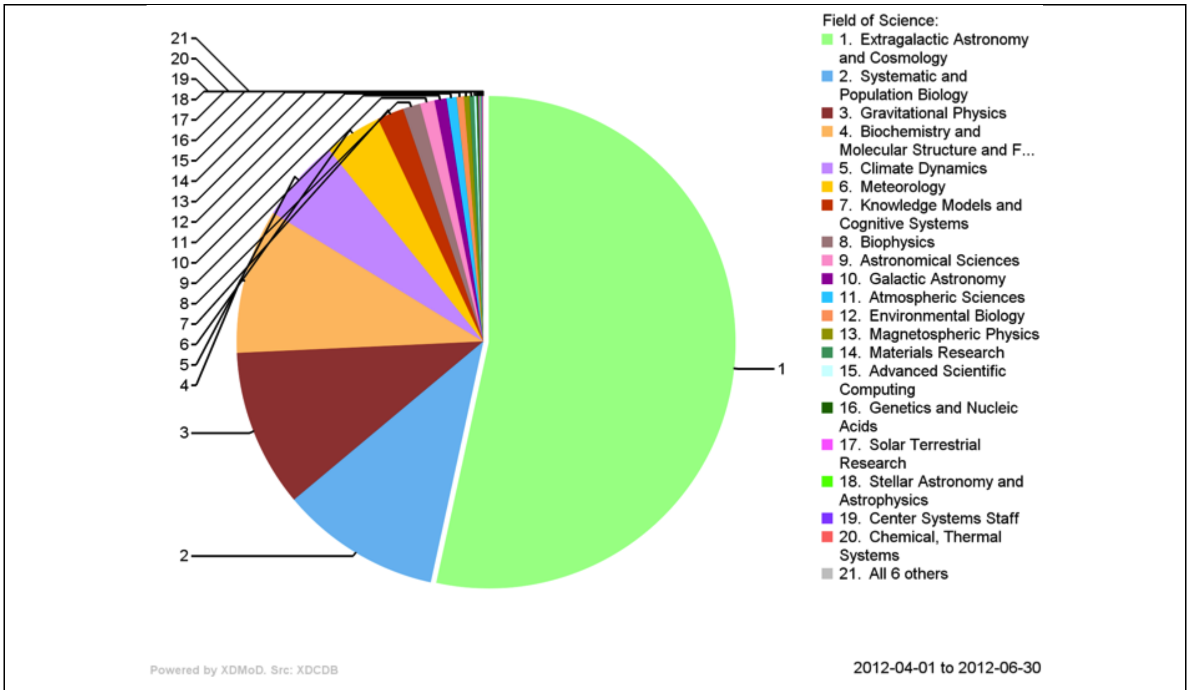


Figure 21: Total NUs charged by field of science on Nautilus in PY1.

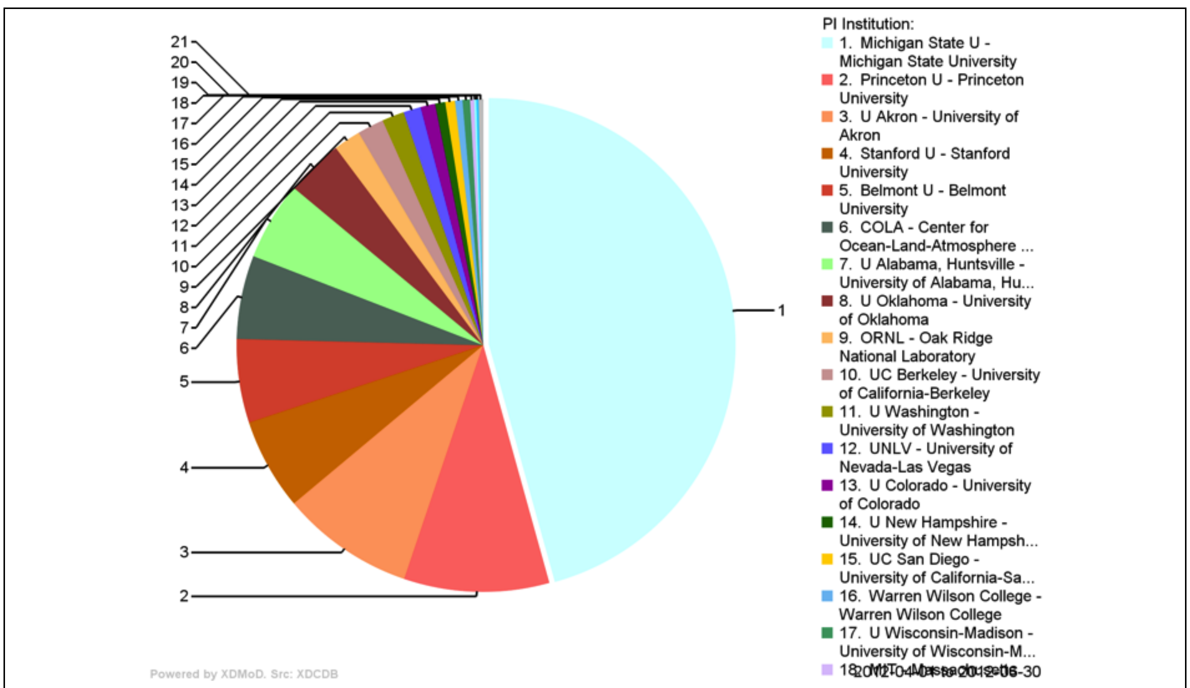


Figure 22: Total NUs charged by PI institution on Nautilus in PY1.

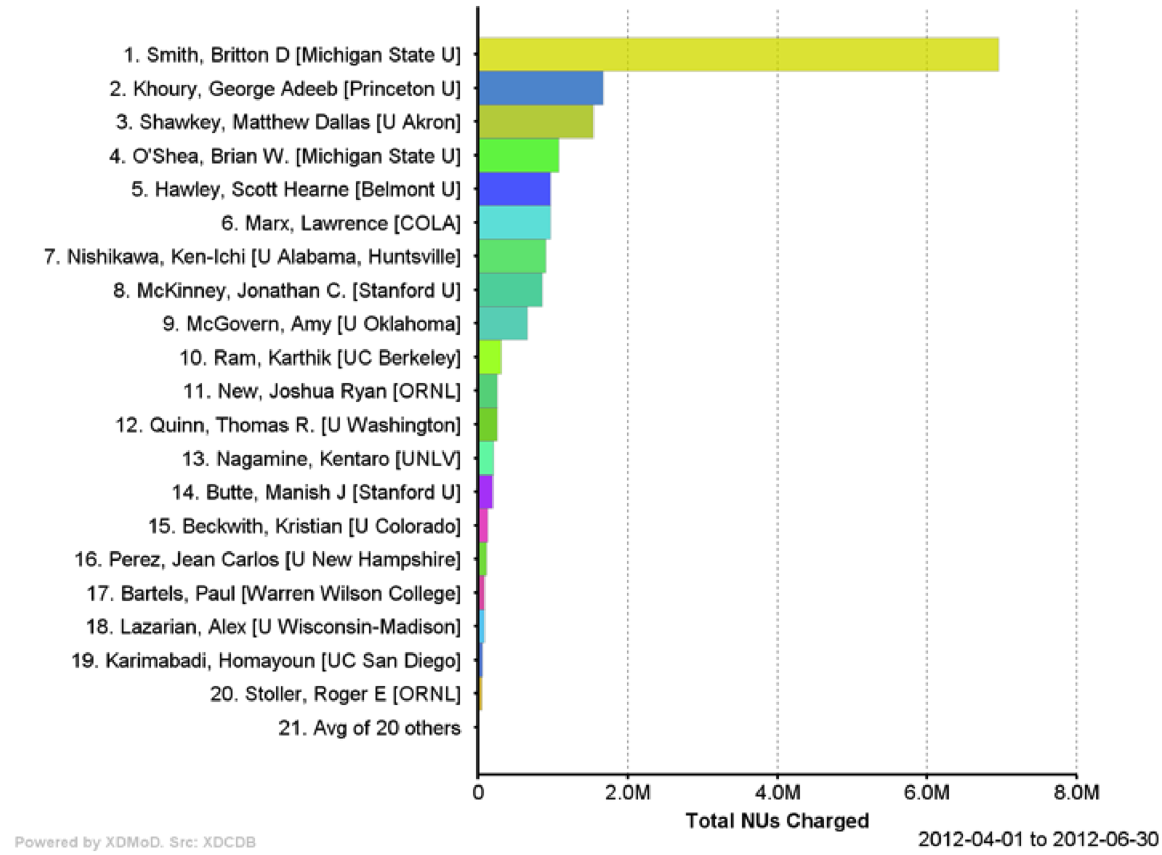


Figure 23: Total NUs charged by PI on Nautilus in PY1.

15.2.2 SP metrics

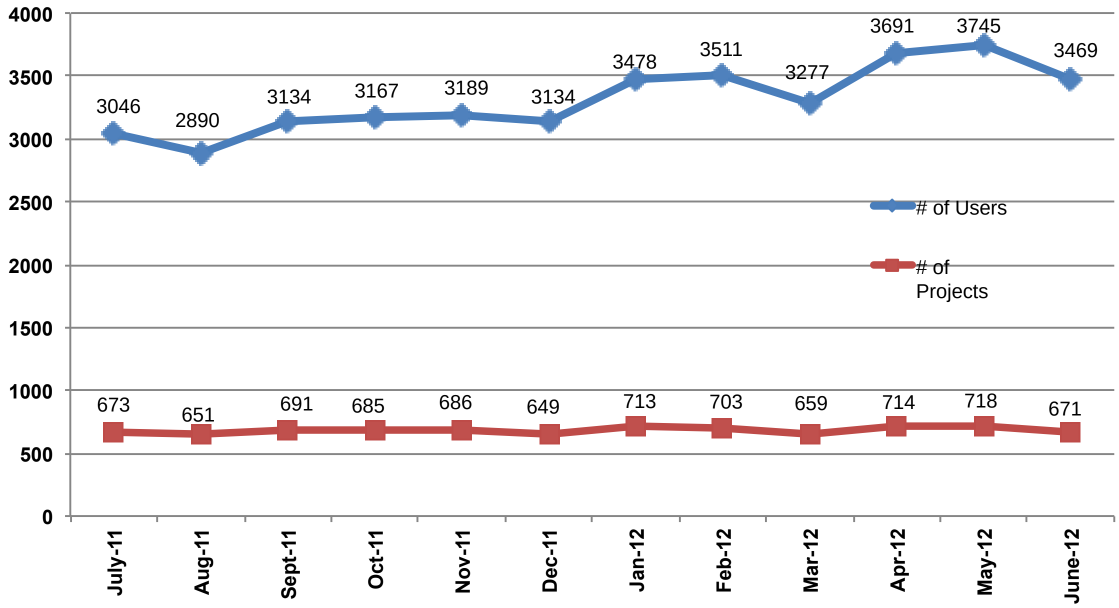


Figure 24: The number of users and projects with active accounts on NICS resources in PY1.

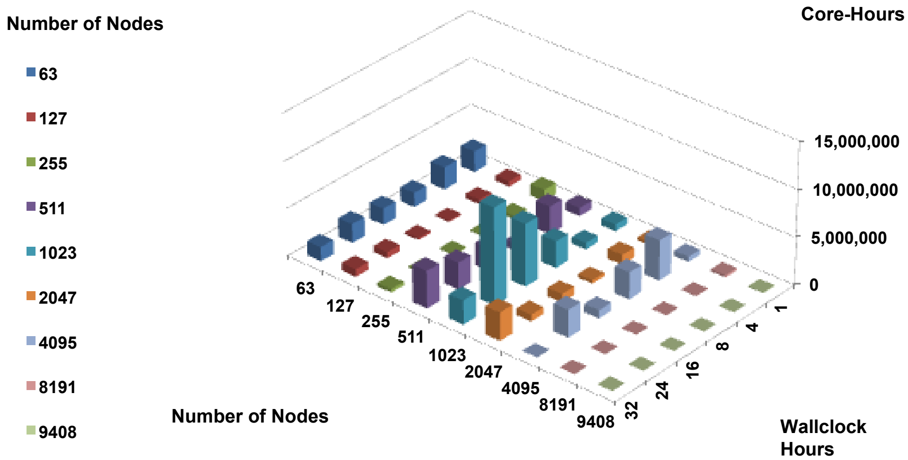


Figure 25: Usage trends on Kraken in PY1.

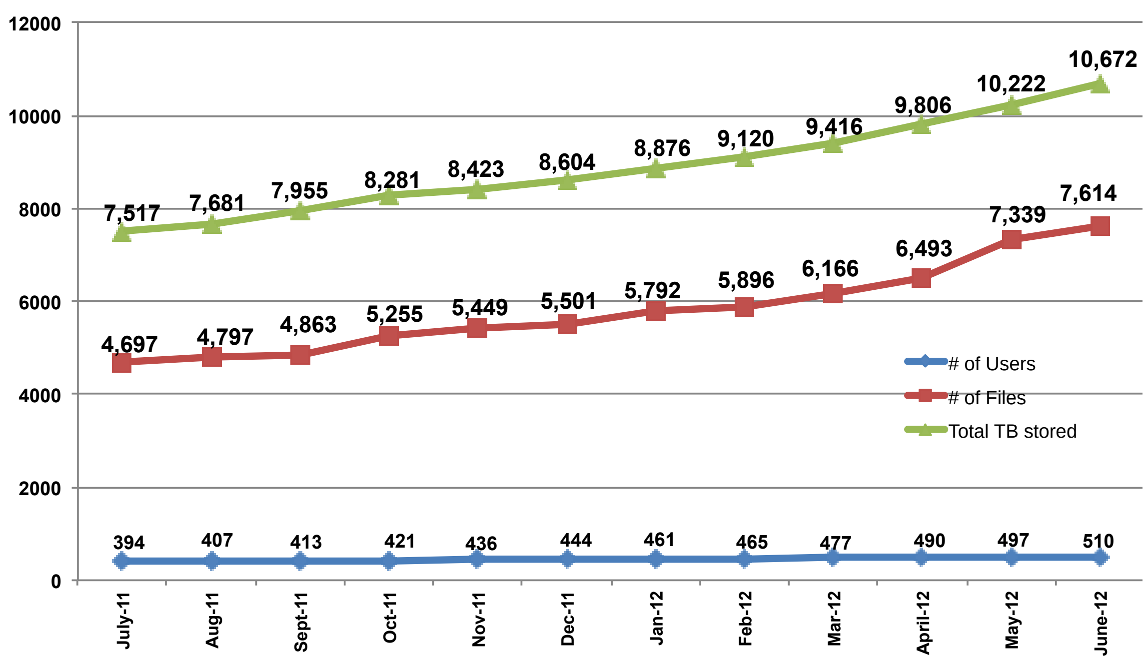


Figure 26: The number of files and their total footprint in the HPSS archive at the close of PY1.

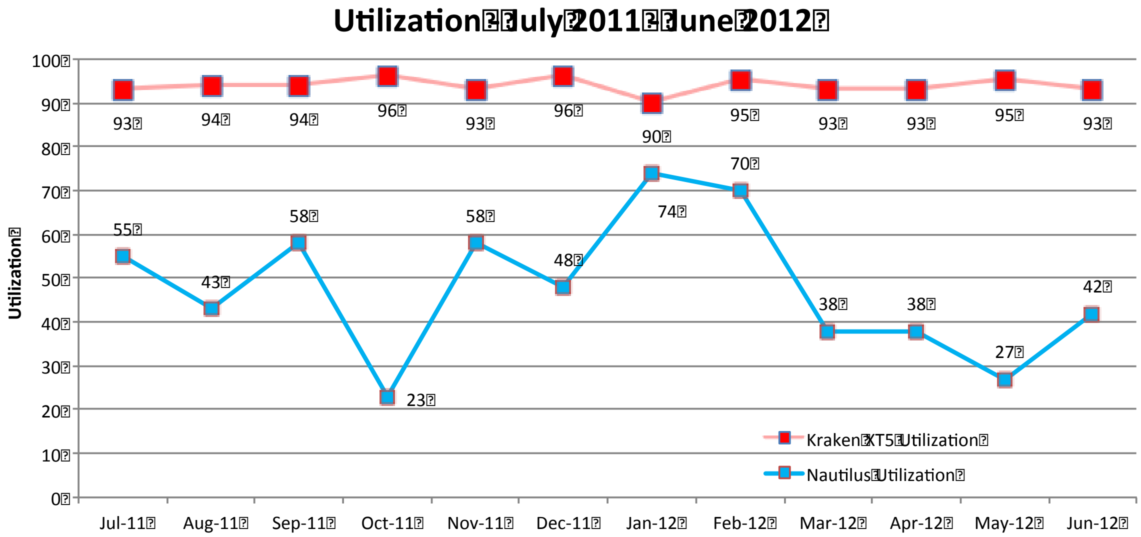
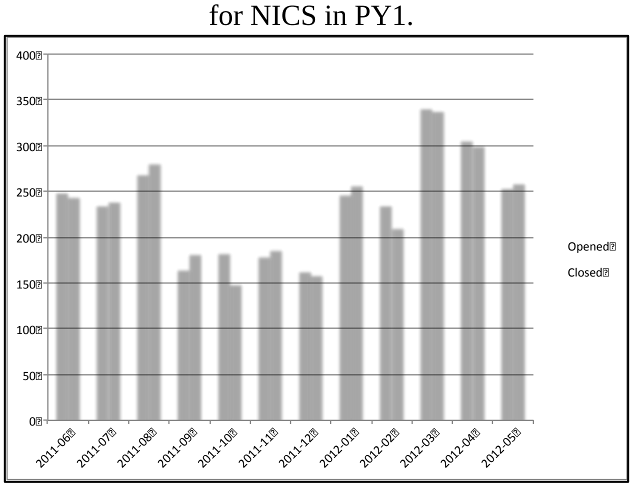


Figure 27: Monthly utilization of Kraken and Nautilus for PY1.

15.2.3 Standard User Assistance Metrics

Figure 28: Ticket volume and resolution by month for NICS in PY1.



The frontline user support at NICS responded to 2,838 newly opened tickets in PY1 (Figure 28) and closed 2,816 tickets--of these 2,737 tickets were marked as resolved. The majority of those tickets closed, but not resolved were abandoned by the user. The resolved tickets corresponded to a variety of issues (Table 4) with the majority falling into two groups: login/access issues and jobs/batch queues. Open tickets experienced a MTTR of 12.2 hours for PY1 with most tickets being resolved within 24 hours.

Table 4: Ticket resolution times by category for PY1.

	account issues	filesystems	gateways	grid software	inca messages	jobs / batch queues	login / access issues	mss / data issues	network issues	other	reservation request	software / apps	system issues	Grand Total
0-1 Hour	12	19	-	5	-	113	343	31	1	12	-	51	12	**599**
1-24 Hours	43	47	1	11	3	270	417	45	4	31	1	109	36	**1018**
1-7 Days	37	38	-	11	1	233	175	39	3	16	-	146	22	**721**
1-2 Weeks	2	6	-	7	-	73	37	15	1	4	-	62	8	**215**
> 2 Weeks	5	5	-	2	-	61	12	20	2	7	-	65	5	**184**
Grand Total	**99**	**115**	**1**	**36**	**4**	**750**	**984**	**150**	**11**	**70**	**1**	**433**	**83**	**2737**

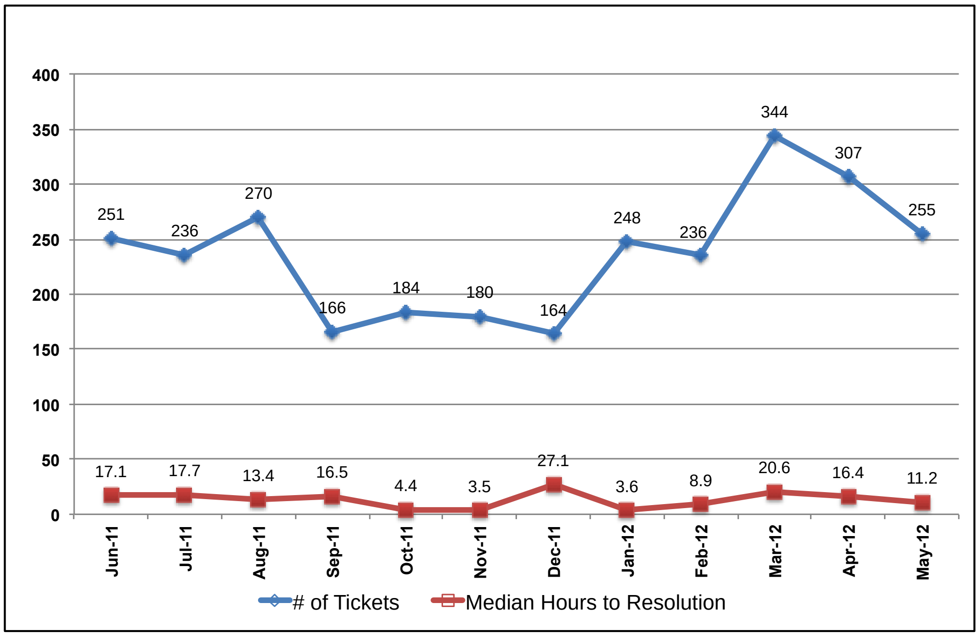


Figure 29: The number of tickets opened and the mean time to resolution of those resolved each month for PY1.

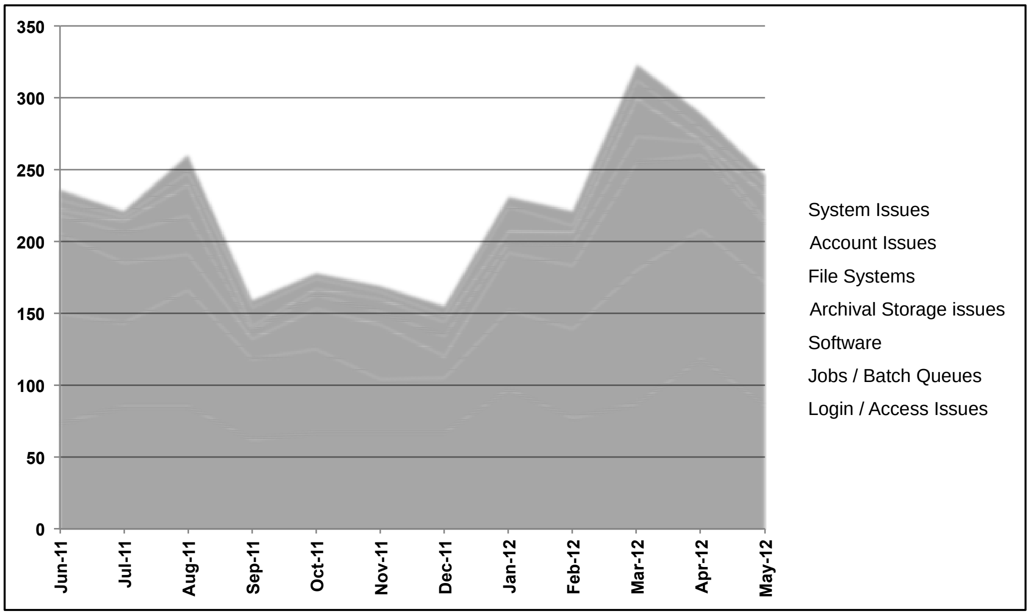


Figure 30: The distribution of monthly tickets by issue type for PY1.

15.2.4 SP-specific Metrics

NICS' resources provided 59% of computational cycles that were delivered to the NSF community in PY1 (Figure 31).

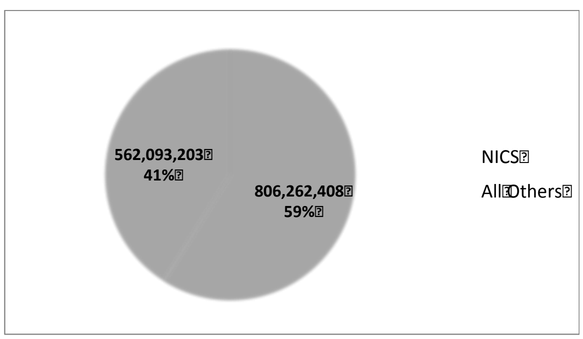


Figure 31: NICS as a percentage of total CPU hours provided by XSEDE in PY1.

16 Pittsburgh Supercomputing Center - Service Provider Quarterly Report

16.1 Executive Summary

The Pittsburgh Supercomputing Center (PSC) operates and supports, *Blacklight*, a powerful and unique resource for the national research community. *Blacklight*, an SGI Altix UV 1000 acquired with the assistance of an NSF grant and operated as an XSEDE resource, is the world's largest shared-memory system, providing two partitions of 16TB each. *Blacklight*, as well as other PSC systems, is supported by a central file system, archival storage and extensive LAN, MAN and WAN infrastructure. For persistent storage such as archiving files, hosting data collections, etc., PSC operates Data SuperCell, a scalable, disk-only file repository that provides fast access to files. Its initial deployment has four petabytes. With operational funding from NIH, PSC also operates *Anton*, a special-purpose computer for molecular dynamics which is used by many NSF-supported researchers.

PSC resources enabled significant progress in many areas; e.g., genomics and molecular biology. Users are finding *Blacklight's* operating characteristics to be very valuable. Jobs that use more than 2 terabytes are now commonplace. Also, Panagiotis Foteinos of Old Dominion University says, "Our application is memory intensive but not bandwidth-bound. There are many remote accesses per second of just a few bytes. Therefore, latency is what bothers us. *Blacklight's* NUMAlink inter-blade connection offers very few latency cycles from hop to hop." Matthew MacManes of the University of California, Berkeley says, "Without [*Blacklight*], I would simply be unable to complete the requisite analyses… Currently there is simply no better resource for this type of work."

Via MATLAB Distributed Computing Server (MDCS), users are running MATLAB in parallel on *Blacklight*. Since MATLAB is heavily used by researchers who are not already users of HPC, this, the only instance of MDCS in XSEDE, is expected to bring new communities of users to HPC. PSC user support people worked with many users on their applications and continue to earn high praise for their efforts. Similarly PSC's Three Rivers Optical Exchange is helping NRAO and WVnet with their connections to the Internet. PSC networking staff members with collaborators at the National Center for Supercomputing Applications are developing Web10G, a tool that will enable ordinary users to effectively use advanced networks. PSC systems administrators helped the University of Victoria with operation of their UV 1000.

PSC engaged in a broad range of Training, Education and Outreach activities, which included enlisting new communities into HPC, several major STEM education programs, and HPC training workshops. PSC has a program of supporting undergraduate interns which has contributed to their career development. For instance, PSC's mentorship of Duquesne University physics student Alex Arrico, a summer intern at PSC in 2010 and 2011, was instrumental in Arrico's acceptance as a graduate student in the Materials Science Department at the University of Tennessee.

In collaborative activities, the Emmons group at Albert Einstein College of Medicine has completed alignment of the nerve ring of the C. elegans male using the AlignTK software developed by Greg Hood of PSC's National Resource for Biomedical Supercomputing. A team from the Swanson School of Engineering, the Pittsburgh Supercomputing Center and Westinghouse will utilize an $876,000 grant to develop a comprehensive experimentally validated computational framework for the turbulent mixing in the lower plenum of next generation high temperature gas reactors (HTGRs).

16.1.1 Resource Description

PSC provides a range of computing and storage platforms for the national science community.

For applications requiring very large shared memory, high-productivity programming models, and/or moderate parallelism with a high-performance system-wide interconnect, PSC operates *Blacklight*, an SGI UV 1000 cc-NUMA shared-memory system comprising 256 blades. Each blade shares 128GB of local memory, and holds two Intel Xeon X7560 (Nehalem) eight-core processors, for a total of 4096 cores and 32 TB across the whole system. Each core has a clock rate of 2.27 GHz, supports two hardware threads and can perform nine Gflop/s for a total system floating point capability of 37 Tflop/s. *Up to 16 TB of this memory is accessible as a single memory space to a shared-memory program.* Message-passing and PGAS programs can access all 32 TB on the system. *Blacklight* is part of the National Science Foundation XSEDE integrated national system of cyberinfrastructure.

Additionally, PSC has an SGI Altix 4700 system called *Salk*, smaller than *Blacklight*, which is also targeted at applications requiring large shared memory, high-productivity programming models, or moderate parallelism with a high-performance, system-wide interconnect. *Salk* is administered for the NIH-funded National Resource for Biomedical Supercomputing (NRBSC) and offers 144 Montvale processors providing a peak aggregate speed of 0.96 Tflop/s with 288 GB shared memory. This system supports advanced programming languages and models including UPC and Star-P.

PSC operates an *Anton* special-purpose supercomputer for molecular dynamics (MD) simulation that performs up to 100 times faster than conventional supercomputers. Designed by D. E. Shaw Research (DESRES) and provided to PSC without cost by DESRES, it is available for non-commercial research use by universities and other non-profit institutions. This machine, the only *Anton* computer operated outside DESRES, is hosted by PSC and is available to the national biomedical community with funding from NIH's National Institute of General Medical Sciences. Computing time on *Anton* is allocated by a peer-review committee convened by the National Research Council. A large number of *Anton* users are NSF supported investigators. The *Anton* computer is supplemented by a high performance file storage system for simulation trajectories and an analysis cluster (*Kollman*). Each of the four nodes in the analysis cluster consists of two Intel Westmere six-core processors and 96 GB of memory. The high-performance file storage system consists of a 500-TB Lustre file system. The file system and the analysis cluster nodes are interconnected over Quad Data Rate (QDR) InfiniBand. Availability of the *Anton* system has been extended at least until September 2013.

PSC operates several Linux clusters for scientific research as well as several high-end servers and powerful workstations for development, analysis, and visualization tasks.

The production workload on all of the PSC computing platforms is managed by PBS/Torque. Several scheduler policy modules used include a locally-developed module, *Simon*, and the Maui scheduler.

All of the PSC computing platforms have access to *Brashear*, PSC's shared, central file system using the Lustre file system architecture. It comprises eight storage nodes and 350 TB of direct-attached disks, forming a large I/O cluster globally accessible within the PSC site. Access to the file system is provided by InfiniBand, 10-Gigabit Ethernet and 1-Gigabit Ethernet. Each node in the I/O cluster is a Lustre Object Storage Server (OSS) hosting multiple Object Storage Targets (OSTs).

PSC is a partner in the Lustre *Albedo* Wide Area File System project along with five other XSEDE sites. PSC is taking the lead by managing the metadata service for the *Albedo* file system in addition to providing a portion of its bulk object storage. In addition to *Albedo*, PSC has installed, debugged and made available a test file system running pre-release Lustre software incorporating advanced security with Kerberos authentication. *Albedo* is mounted on *Blacklight*.

PSC's archive file repository is a disk-only system that is less costly than a disk-tape system and provides much faster file access. Each building block in the repository has one petabyte of useable disk storage, which is managed by the ZFS file system and the PSC-developed SLASH2 replicating distributed file system. ZFS and SLASH2 provide multiple layers of robust data integrity checking to protect user data against data corruption. This building-block architecture will enable the repository to scale well beyond its initial deployment of four petabytes.

Users can access the repository from within PSC using the familiar PSC file archiving utility, *far*. From outside PSC, users can employ a variety of well known file transfer methods such as scp and gridftp. These transfers are handled by a series of dedicated data transfer servers.

PSC network facilities consist of production and research Local Area Network (LAN), Metropolitan Area Network (MAN), and Wide Area Network (WAN) infrastructures.

Local Area Network Infrastructure: The LAN infrastructure consists of switched Ethernet with speeds up to 10 Gb/s. The LAN architecture was constructed to overcome issues of buffer contention in data center Ethernet switches on the Science DMZ[1]. This allows for higher bandwidth data transfers to the data transfer nodes.

3 Rivers Optical Exchange: PSC operates and manages the 3 Rivers Optical Exchange (3ROX) a regional network aggregation point that provides high-speed commodity and research network access, primarily to sites in Western and Central Pennsylvania and West Virginia. While the primary focus of 3ROX is to provide cost-effective, high-capacity, state-of-the-art network connectivity to the university community, this infrastructure also provides well-defined network services to both community (K-12, government) and commercial entities in Western Pennsylvania. University member sites currently include Carnegie Mellon University, the Pennsylvania State University, the Pittsburgh Supercomputing Center, the University of Pittsburgh, WVnet, and West Virginia University.

3ROX Metropolitan Area Network Infrastructure: 3ROX MAN infrastructure is DWDM-based and supports multiple 10-Gigabit Ethernet waves. It is capable of supporting 40 and 100-Gigabit waves as the need arises. This DWDM network connects four different locations around the city that include long haul service providers, a co-location hotel, a campus based co-location facility, and the Northern Pike machine room.

3ROX Wide Area Network Infrastructure: 3ROX WAN infrastructure has both Commodity Internet and Research and Education components. Explicit routing is used to maintain the acceptable use policies associated with the various production and research network infrastructures.

The 3ROX Commodity Internet component consists of multiple high-performance WAN connections to major Internet service providers, including a Gigabit Ethernet connection to Cogent and a 10 Gigabit Ethernet connection to Global Crossing. In addition, the Internet2 Content Peering Service is also used to offload Commodity Internet traffic.

The 3ROX Research and Education component includes a 10-Gigabit Ethernet connection, with 5 Gb/s of bandwidth, to the Internet2 network. In addition to the Internet2 connection, 3ROX also has a 10-Gigabit Ethernet connection to National LambdaRail's PacketNet service, a nationwide routed IP network; a 10-Gigabit Ethernet connection to the XSEDE backbone network via National LambdaRail; a 10-Gigabit Ethernet connection to National LambdaRail's FrameNet

[1] From http://fasterdata.es.net/science-dmz/: The Science DMZ is a portion of the network, built at or near the campus or laboratory's local network perimeter that is designed such that the equipment, configuration, and security policies are optimized for high-performance scientific applications rather than for general-purpose business systems or "enterprise" computing.

service, a nationwide Layer2 network; a 10-Gigabit Ethernet connection between PSC's offices and its remote supercomputing machine room at 4350 Northern Pike; and a 10-Gigabit Ethernet connection to Penn State University (PSU) to provide XSEDE connectivity to PSU.

16.2 Science Highlights

In addition to major science accomplishments that are highlighted in the XSEDE report, we present selected others specific to PSC.

16.2.1 Behavioral Genomics: Using Next-Generation Sequencing to Better Understand Social Behavior in the Tuco-Tuco (Matthew D. MacManes & Eileen A. Lacey, University of California, Berkeley)

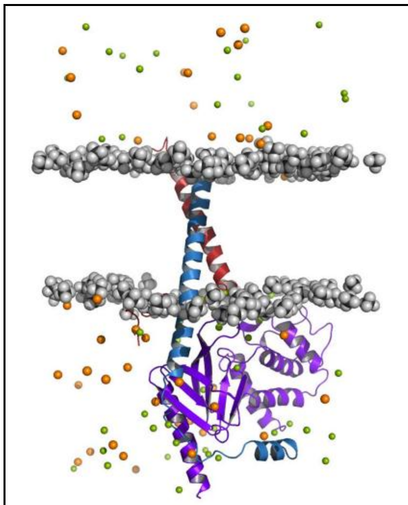
Matthew MacManes is interested in how differences in gene expression are related to the transition from social to solitary living in the colonial tuco-tuco, a species of burrowing rodent. MacManes and colleagues have exploited *Blacklight's* shared memory and its documented advantages for genomic analysis (see *GenomeWeb*, February 1, 2012) for this study. The tuco is a unique case for investigating the genetics of social behavior in that some of these animals live in groups while others leave their burrow system to live in solitary conditions. The study uses messenger RNA extracted from the hippocampus (a brain region implicated by prior research in social behavior) of captive tucos housed (at Berkeley's Museum of Vertebrate Zoology), with control groups in social and solitary conditions. The extracted nucleotide tissue was sequenced (in an Illumina sequencer), yielding 56 billion base-pairs of raw data — 56 million 100 base-pair reads. MacManes used XSEDE resources at several sites (including TACC's *Ranger* and *Longhorn*), but found *Blacklight* to be the most effective hardware for RNA assembly and analysis. XSEDE consultant Phil Blood of PSC installed pre-compiled modules of about 20 programs typically used in genomics assembly and analysis, which MacManes found to be extremely helpful. Using 640 gigabytes of RAM (80 cores), the assembly required 14 days of computing, with subsequent analysis extending for months. MacManes' analysis identified a number of genes that are differentially expressed according to tuco social behavior, and the researchers have a manuscript reporting their findings in preparation. "*Blacklight* is a key resource for my analyses of next-generation sequence data," says MacManes. "Without it, I would simply be unable to complete the requisite analyses. I feel so strongly about *Blacklight* that I have referred colleagues and collaborators. Currently there is simply no better resource for this type of work."

The Patagonian colonial tuco-tuco (Ctenomys sociabilis) in its native environment (volcanic ash). The tuco-tuco is a species of subterranean rodent, related to the common guinea pig, of interest in behavioral genomics because they exhibit a dichotomy between social and solitary living conditions.

16.2.2 Molecular Biology: Mechanistic Insight into the "Inside-Out" Signaling of Integrins (Marta Filizola, Mount Sinai School of Medicine)

Integrin is a receptor, composed of two subunits, that mediates the attachment between a cell and surrounding tissue, including other cells and the extracellular matrix. It often functions as a signaling protein, responsive to changes in the extracellular environment, and is involved in immune patrolling and cell migration and is prominently involved in the chain of interactions that instigates blood clotting in response to a wound. Using PSC's *Anton* system, Filizola and

colleagues carried out simulations that broaden current understanding of the mechanism by which integrin is activated by an intracellular protein called talin. The simulations modeled a form of integrin called *αIIbβ3,* composed of *αIIb* and *β3* subunits. Like other integrins, αIIbβ3 can signal bi-directionally through interaction of its extracellular domain with extracellular ligands ("outside-in" signaling) or its cytoplasmic tail with intracellular proteins ("inside-out" signaling). To obtain rigorous insights into αIIbβ3 "inside-out" signaling at an unprecedented level of molecular detail, Filizola and colleagues used *Anton* for microsecond-scale, all-atom molecular dynamics simulations of experimentally-based models of αIIbβ3 in an explicit phospholipid bilayer, and in the presence of the F2-F3 subdomains of the intracellular activator talin-1. With *Anton,* the researchers were able to simulate about six microseconds of trajectory for integrin interacting with talin, which revealed that talin stabilizes the intracellular integrin domains, preventing an otherwise tilted orientation. From their results, the researchers also observed a significant bending in the *β3* transmembrane helix due to talin binding, and found a rotation of the F3 with respect to the F2 talin domain. These observations confirm recent experimental work to the effect that talin anchors to the membrane phospholipids. These and other detailed findings point to a more complex molecular mechanism involved in the integrin transmembrane helical change from inactive to active conformation than proposed to date by experimental studies.



Transmembrane regions of integrin αIIb (red) and β3 (blue) subunits interacting with the intracellular protein talin (purple). Lipid head groups and ions are represented as spheres. For clarity, the lipid chains and the water have been removed.

16.3 User-facing Activities

16.3.1 System Activities

PSC networking staff members worked with the Greenbank NRAO facility to understand the end-to-end networking issues they have been seeing for the past year or so. While their traffic does not currently go through PSC's Three Rivers Optical Exchange (3ROX), we offered to help them debug their connection in anticipation of their eventually connecting through 3ROX once their 10-GE link, funded through their BTOP grant, is connected to WVU in Morgantown. The collaboration determined that the bottleneck was local to their facility and they are now working on making local changes. A 10-GE connection to WVNet was deployed and tested during the reporting period. We expect a significant increase in our traffic load to both the commodity and R&E networks from this connection.

16.3.2 Services Activities

Blacklight users continue to make good use of its characteristics in their work. *Blacklight* weekly usage often exceeds 80% of the maximum possible, which is remarkably good given that our scheduling approach treats all jobs including those from *Ember* in the same way (i.e., no preference to jobs with desirable utilization characteristics). Jobs that use more than 2 GB have become commonplace. One job used 7.1 TB and ran for 10.02 hours on 1,536 cores. The NUMAlink interconnect is also useful. For example, Panagiotis Foteinos at the Computer Science Department of Old Dominion University, whose jobs often use 2 terabytes of memory, writes:

> Our application is memory intensive but not bandwidth-bound. There are many remote accesses per second of just a few bytes. Therefore, latency is what bothers us. *Blacklight's* NUMAlink inter-blade connection offers very few latency cycles from hop

> to hop. Our application is tightly coupled (lots of synchronization). We chose the shared memory model (against distributed) for our code in order to decrease complexity. Getting good performance on distributed systems is messier and sometimes impossible. *Blacklight* is the largest NUMA machine we are aware of. It supports our shared memory model, it has many cores, and it also gives us the opportunity to implement "distributed" improvements that other people have done in the traditional distributed machines.

Via MATLAB Distributed Computing Server (MDCS), users are running MATLAB in parallel on up to 32 cores on *Blacklight*. This is the only instance of MDCS in XSEDE. Given the widespread use of MATLAB in communities that have not been heavily involved in high performance computing, the availability of MDCS in XSEDE will bring new users to HPC. Since this is a new capability, we are monitoring users' experiences and will address any issues that may arise.

Headlined "Three Rivers Optical Exchange: One of the Leading Resources in the World for Network Know-How," an article in the Quilt Circle magazine discusses PSC's Three Rivers Optical Exchange (3ROX) and its networking support for PSC's National Resource for Biomedical Supercomputing (NRBSC). See http://www.thequilt.net/index.php/resources.

PSC networking staff members with collaborators at the National Center for Supercomputing Applications are developing Web10G, a tool that will enable enable ordinary users to effectively use advanced networks. Web10G was accepted as a tutorial for the Summer Joint Techs meeting to be held at Stanford July 15-19. The tutorial is designed to acquaint the participants with the utility of Web10G, explain the transition from Web100 to Web10G, and provide an environment for the participants promoting hands-on development and creation of simple Web10G clients using the newly retooled *userland* API. In addition to the tutorial, a talk, "Web10G: Meeting the Promises of Web100" was accepted as a presentation for the Joint Techs meeting.

Users continue to appreciate the support they receive from PSC staff members. For instance, Niel Henriksen of Tom Cheathams's group at the University of Utah wrote the following to Marcela Madrid:

> Thank you so much for all the help you've given us with our calculations on *Blacklight*. We really appreciate the individualized support for our unique needs and the timely responses to our questions. Our large QM calculations seemed like an ideal fit with the *Blacklight* machine and we hope to continue using this resource in the future to provide reference data for improving MD simulations of RNA. It has been great working with you and the staff at PSC.

Similarly, Tomekia M. Simeon (Department of Chemistry, Northwestern University), Sonja Braun-Sand (Department of Chemistry and Biochemistry, University of Colorado), and Archana Dubey (Physics Department, University of Central Florida) sent notes of thanks.

16.4 Security

Although PSC systems are under frequent attack, PSC had no security incidents.

PSC uses a virtual private network (VPN) service that allows staff to protect themselves from eavesdropping and other threats by tunneling them back to PSC while they are traveling or at a remote location. While machines that are physically behind PSC's security firewall are scanned for security vulnerabilities by our Nessus security scanners on a periodic basis, machines that use the VPN are scanned every time they come in through it.

16.5 Education, Outreach, and Training Activities

Type	Title	Location	Date(s)	Hours	Number of Participants	Number of Under-represented people	Method
Presen-tation	Advanced Bioinformatics	SciTech Academy of Pittsburgh Public Schools Pittsburgh, PA	17 APR 2012	2.5	26	12	S
Workshop	Accelerating Applications with OpenACC	PSC Pittsburgh, PA	18-19 APR 2012	14	25	5	S
Workshop	Prof. Dev. for Teachers on PSC's BEST and CMIST Programs	Oakland Catholic Girls High School Pittsburgh, PA	25 APR 2012	3	30	30 (all girls)	S
Workshop	Towards High-Volume Data Computing at Petascale – 2nd International Workshop on High Volume Data Modeling and Visualization	King Abdullah University of Science and Technology Saudi Arabia	04-06 MAY 2012	16	30	0	S
Workshop	Computational Methods for Spatially Realistic Microphysiological Simulations	PSC Pittsburgh, PA	07-09 MAY 2012	24	18	4	S
Tutorial	Accessing Free, Large-scale Computation and Data Resources for Bioinformatics Through the eXtreme Science and Engineering Discovery Environment (XSEDE)	Great Lakes Bioinformatics Conference, University of Michigan Ann Arbor MI	15-17 MAY 2012	2	60	N/A	S
Presen-tation	Careers in the Computational Biosciences	SciTech Academy of Pittsburgh Public Schools Pittsburgh, PA	21 MAY 2012	2.5	26	12	S
Presen-tation	Careers in the Computational Biosciences	Oakland Catholic Girls High School Pittsburgh, PA	01 JUN 2012	3	30	30 (all girls)	S

Type	Title	Location	Date(s)	Hours	Number of Participants	Number of Under-represented people	Method
Presen-tation	ISchool Inclusion Institute	PSC Pittsburgh, PA	05-JUN-2012	2	20	20	S
Workshop	Computer Simulation of Biomolecular Dynamics and Reactions	PSC Pittsburgh, PA	05-08 JUN 2012	32	24	10	S
Workshop	MARC – Bioinformatics Summer Institute	PSC Pittsburgh, PA	19-27 JUN 2012	68	27	27 (participants from minority serving institutiuons)	S
Course	Extreme Scaling	International Summer School on HPC Challenges Dublin, Ireland	24-28 JUN 2012	1	90	25	S
Course	Supercomputing in Public Health Policy	International Summer School on HPC Challenges Dublin, Ireland	24-28 JUN 2012	1	35	10	S
Tutorial	Performance Engineering of Parallel Codes	International Summer School on HPC Challenges Dublin, Ireland	24-28 JUN 2012	3.5	60	22	S

In recognition of his mentorship of Duquesne physics student Alex Arrico, a summer intern at PSC in 2010 and 2011, which was instrumental in Arrico's acceptance as a graduate student in the Materials Science Department at the University of Tennessee, Professor Simonetta Frittelli, Chair of the Department of Physics at Duquesne University and the advisor of the Duquesne Chapter of the National Physics Honor Society (Sigma Pi Sigma), notified PSC's Yang Wang that he has been elected to membership in Sigma Pi Sigma.

James Komianos, a Carnegie Mellon University physics undergraduate who was mentored by PSC's Yang Wang and Roberto Gomez last summer, had a poster on his work on the use of CUDA GPU in physics applications at CMU's "Meeting of the Minds - Undergraduate Research Symposium" on May 9, 2012. See http://www.cmu.edu/uro/MoM/index.html for more information about the symposium. This summer Yang Wang supervised intern Dominic DeBiasio, an undergraduate at Duquesne University, on a project called, "Finding the Diagonal

Blocks of the Inverse of Large Matrices," which has application to the software package LSMS. Troy Wymore supervised intern Jacob Riddle, an undergraduate at the University of Pittsburgh, on a project called, "Quick-SNA^{5}PS: An online Server for Detecting Signatures of Natural Selection."

On June 19-27, PSC's National Resource for Biomedical Supercomputing (NRBSC) held its annual MARC-sponsored Bioinformatics Summer Institute with participants from minority serving institutions across the country. The goal of the workshop is to assist minority serving institutions with developing bioinformatics programs. The workshop focuses on the analysis of a gene and/or protein family through the construction and subsequent mining of a high accuracy multiple sequence alignment along with assembling and analyzing data from next-generation sequencers. XSEDE campus champion Eduardo Socolovsky from Norfolk State attended the workshop. We will do a two-day bioinformatics workshop at Norfolk State with Socolovsky in the fall. Ten of the workshop participants are also part of NRBSC's MARC summer internship program, during which the interns work with NRBSC staff on individual science projects.

Pallavi Ishwad continues to lead BEST and CMIST, outreach programs developed by NRBSC. BEST (Better Educators of Science for Tomorrow) is a program that prepares teachers to refocus their teaching strategies towards exposing and encouraging students to become aware of emerging and exciting biomedical careers. It provides a high school level bioinformatics curriculum and ongoing support for teachers. CMIST (Computational Modules in Science Teaching) brings innovative science tutorials into secondary school classrooms, focusing on integrative computational biology, physiology, and biophysics. In contrast to other teaching tools, CMIST modules include high quality, biologically realistic 3-D animations produced with cutting-edge simulation and visualization software developed at NRBSC.

For the past several years, PSC has hosted a team that placed among the winners of the Science Bowl, an event that is organized by the National Energy Technology Laboratory each year. On Monday June 18 we hosted the Science Bowl team from Mt. Lebanon High School. As usual, the team started with a tour of the machine room and then heard talks from a few PSC scientists/engineers to get them excited about careers in science and technology. Pictured at right in the PSC machine room are, l to r, students Nick Brown, Kevin Skolnick, Jerry Zhang and teacher Joshua Bilak.

PSC's Dave Moses, Pallavi Ishwad, Yang Wang and Markus Dittrich were Sponsor Judges in the Pittsburgh Regional Science and Engineering Fair on March 30. PSC sponsored two awards, which went to the following two students:

- Jarrod Cingel - Bethel Park High School - Software Parallelization and Machine Translation
- Rishi Mirchandani - Fox Chapel Area High School - When Irrelevance Becomes Relevant: An Election Investigation

The (Intel-sponsored) International Science and Engineering Fair (ISEF) took place May 13-18 at the Pittsburgh Convention Center. ISEF is the largest and longest-running competition for high-school students in science and engineering disciplines. Award winners receive scholarships funds for college - the top award for the best project across all disciplines is $100,000. See http://www.societyforscience.org/intelisef2012. PSC's Nathan Stone was Co-Chair of Physics "Grand Awards" Judges for ISEF. PSC's Derek Simmel served as a Grand Awards Judge in the Environmental Sciences category.

Regarding the OpenACC workshop in the table, this was a hands-on workshop presented jointly by PSC, NVIDIA and PGI. The workshop allowed time for participants to have one-on-one consultations with PSC staff members. 12 attendees responded to a detailed survey questionnaire. Of these, 4 have already implemented OpenACC into their research. 1 achieved greater than 15-fold speedup, 1 achieved between 2- and 4-fold speedup, and 2 achieved less than 2-fold speedup. 7 have been delayed in implementing OpenACC but are either working on it or intend to do so. 1 realized that OpenACC is not suitable for accelerating his application.

Workshop attendee Galen Arnold of NCSA wrote to Ralph Roskies, "John [Urbanic], Tom [Maiden], Marcela [Madrid] and the rest of the PSC staff did a great job with the OpenACC workshop this week. For me it was time well spent." Maiden, who organized the workshop, reports that Arnold made similar comments to him during the workshop.

Attendee Rich Loft of NCAR said, "After two days of hacking on a kernel 'toy' prototype of the gradient routine from our Discontinuous Galerkin shallow water model, I measured a 10x speedup (30 Gflops vs 3 Gflops) ... using OpenACC."

Regarding the iSchool Inclusion Institute (i3) presentation listed in the table, i3 is a program of Pitt's School of Information Science that addresses a critical problem within the information sciences: a lack of diversity among its students and faculty. To foster a culture of creativity, innovation, and collaboration across disciplines, the iSchools actively recruit and develop students and faculty from underrepresented segments of the population. This is the second year PSC presented to the Pitt iSchool. According to the i3 Director "the students continued discussing some of the ideas/concepts during our evening activities. You all struck a nice balance for the range of majors in the room."

Regarding the 2nd International Workshop on High-Volume Data Modeling and Visualization listed in the table, Yang Wang's invited talk, "Towards High-Volume Data Computing at Petascale", introduced the international audience to PSC and XSEDE services and efforts. There were 15 speakers altogether, including people from LBNL and Beijing University and the local hosts David Keyes and Tom DeFanti. The workshop provided a framework for dialog among people with various scientific backgrounds and from different parts of the world.

PSC's Phil Blood and Shawn Brown contributed to the International Summer School on HPC Challenges that is listed in the table. Brown taught a class on "Extreme Scaling" and one on using supercomputing to help set public health policy. Blood gave a tutorial on "Performance Engineering of Parallel Codes." This is the third time Blood has given this training at this summer school, which is consistently rated one of the most useful presentations at the school by the students. Last year this was the highest rated class at the summer school. Both Brown and Blood were on the planning committee for the summer school this year. Blood thinks "the summer school went very well, probably the best we have had so far."

16.6 SP Collaborations

The Emmons group at Albert Einstein College of Medicine has completed alignment of the nerve ring of the C. elegans male using the AlignTK software developed by Greg Hood of PSC's National Resource for Biomedical Supercomputing.

A team from the Swanson School of Engineering, the Pittsburgh Supercomputing Center and Westinghouse will utilize an $876,000 grant to develop a comprehensive experimentally validated computational framework for the turbulent mixing in the lower plenum of next generation high temperature gas reactors (HTGRs). These high-efficiency reactors are utilized for electricity production and a broad range of process heat applications. The team includes Principal Investigator Mark Kimber, PhD, Assistant Professor, Department of Mechanical Engineering and Materials Science; John Brigham, PhD, Assistant Professor of Structural Engineering and

Mechanics in the Department of Civil and Environmental Engineering; Anirban Jana, PhD, Sr. Scientific Specialist, Scientific Applications and User Support at the Pittsburgh Supercomputing Center; and Milorad Dzodzo, PhD, Westinghouse Electric Company. Through computational fluid dynamics (CFD) modeling and experimental validation, the results from this project will lay the groundwork for future stress analysis, failure and fatigue studies, and uncertainty quantification for HTGR systems.

16.7 SP-Specific Activities

PSC's Chad Vizino, J. Ray Scott, and Brian Johanson made suggestions on how to more effectively administer the UV1000 at the University of Victoria, BC, Canada. The system is part of the Compute Canada/WestGrid/University Systems. Their UV1000 is, like one of our two nodes, a 16TB system with IB-based Lustre I/O. They have been having many problems getting jobs of various types to run. They are working on implementing our RAM-disk management solution. The last message from their systems person said, "Thank you very much, J. Ray. The job I submitted with the changes worked! Your and Chad's inputs were very effective."

16.8 Publications

P. Blood and A. Ropelewski, "Accessing Free, Large-scale Computation and Data Resources for Bioinformatics Through the eXtreme Science and Engineering Discovery Environment (XSEDE)", Great Lakes Bioinformatics Conference, University of Michigan, May 15, 2012.

P. Blood, "Performance Engineering of Parallel Applications", 2012 European-US Summer School on HPC Challenges in Computational Sciences, Dublin, Ireland, June 27, 2012.

Croft, Rupert; Di Matteo, Tiziana; Khandai, Nishikanta; Springel, Volker; Jana, Anirban; Gardner, Jeffrey; Dark matter halo occupation: environment and clustering; eprint arXiv:1109.4169.

Feng, Yu; Croft, Rupert A. C.; Di Matteo, Tiziana; Khandai, Nishikanta; Sargent, Randy; Nourbakhsh, Illah; Dille, Paul; Bartley, Chris; Springel, Volker; Jana, Anirban; Gardner, Jeffrey; Terapixel Imaging of Cosmological Simulations; The Astrophysical Journal Supplement, Volume 197, Issue 2, article id. 18 (2011).

Feng, Yu; Croft, Rupert A. C.; Di Matteo, Tiziana; Khandai, Nishikanta; Sargent, Randy; Nourbakhsh, Illah; Dille, Paul; Bartley, Chris; Springel, Volker; Jana, Anirban; Gardner, Jeffrey; Gaepsi: Gadget Visualization Tookit; Astrophysics Source Code Library, record ascl:1108.005.

Khandai, Nishikanta; Sethi, Shiv K.; Di Matteo, Tiziana; Croft, Rupert A. C.; Springel, Volker; Jana, Anirban; Gardner, Jeffrey P.; Detecting neutral hydrogen in emission at redshift $z \simeq 1$; Monthly Notices of the Royal Astronomical Society, Volume 415, Issue 3, pp. 2580-2593.

C. Vizino, "Update on PSC's UV 1000: Blacklight", SGI User Group Conference, Orlando, Apr. 17-19, 2012.

16.9 Metrics

16.9.1 Standard User Assistance Metrics

Numbers in Table 1 refer to all tickets handled by the PSC help desk in PSC's local ticket system.

Table 1: Distribution of times to resolution for the 334 tickets that were created as well as resolved between 4/1/2012 and 6/30/2012.

Ticket Type	0-1 hr	1-24 hrs	1-7 days	1-2 wks	>2 wks
System issues	2	14	29	6	0
Software/apps	9	19	49	17	16

Ticket Type	0-1 hr	1-24 hrs	1-7 days	1-2 wks	>2 wks
Other	0	4	9	0	3
Login/access	1	2	3	0	0
Jobs/batch queues	4	55	29	4	2
File_systems	1	7	8	1	0
Account issues	0	11	26	3	0

Numbers in Table 2 refer to tickets relating to PSC that were handled in the central XSEDE Ticket System.

Table 2: Distribution of times to resolution

Time to Resolution	account issues	file systems	grid software	jobs/batch queues	login/access issues	mss/data issues	network issues	software/apps	system issues	other
0-1 hr					1					
1-24 hr				3	2			3		
1-7 d	4	1		4	6			8		1
1-2 wk		2		3	6	1		9		2
> 2 wk	3	1		9	4	1		13		
Still Open			1					1		

16.9.2 SP-specific Metrics

Key system statistics for *Blacklight* for 4/1 to 6/30/2012are shown in Table 3.

Table 3: Operational Statistics - *Blacklight*		
Number of unplanned outages	15	
Number of planned outages	3	
Total outages	18	
Number of jobs failures due to system faults	66	
Total time* in period (hours)	4368	100%
Scheduled Downtime (hours)	47.00	1%
Unscheduled Downtime (hours)	209.40	5%
Total Downtime (hours)	256.40	6%
Total time available to users (total-downtime)	4111.60	94%
% System Utilization	79.9%	

* On *Blacklight* a node is half the machine. Time values listed are expressed in node hours.

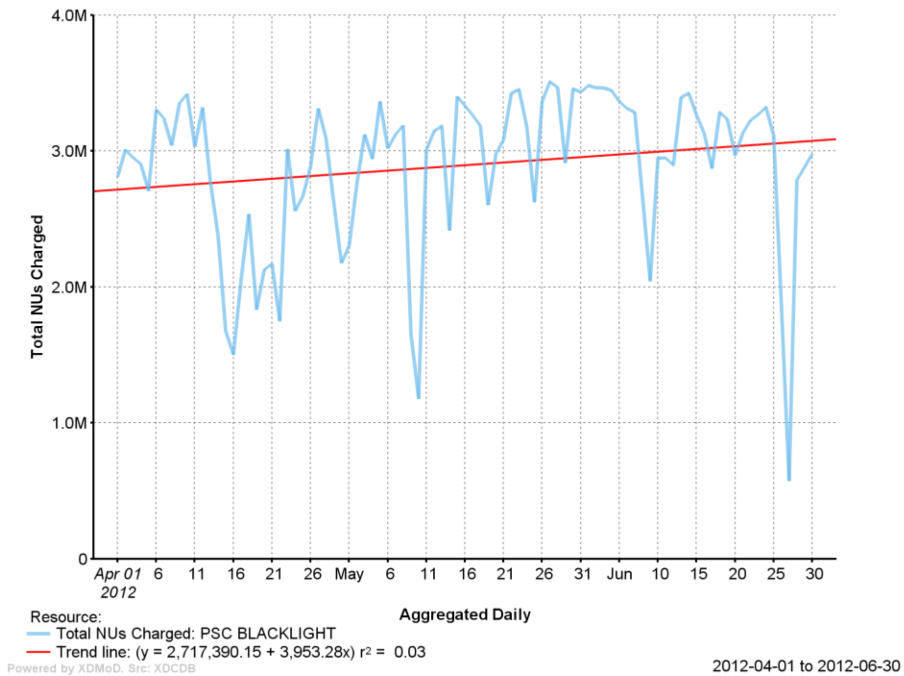
16.9.3 Standard systems metrics

The following charts of standard system metrics for *Blacklight* were provided by the Technology Audit Services team.

Total NUs Charged by Resource

Resource = PSC-BLACKLIGHT

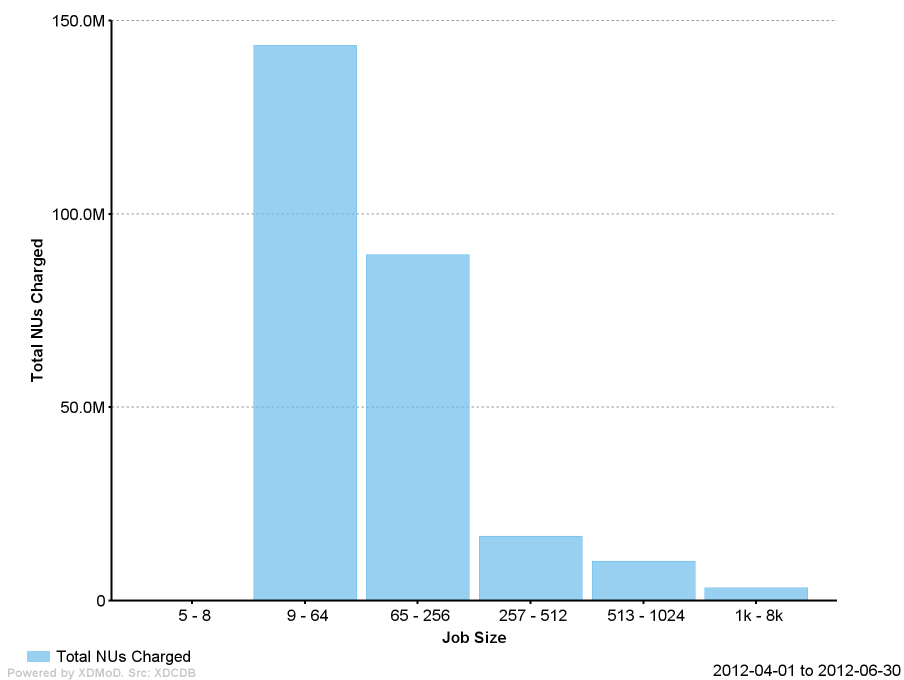
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = PSC-BLACKLIGHT

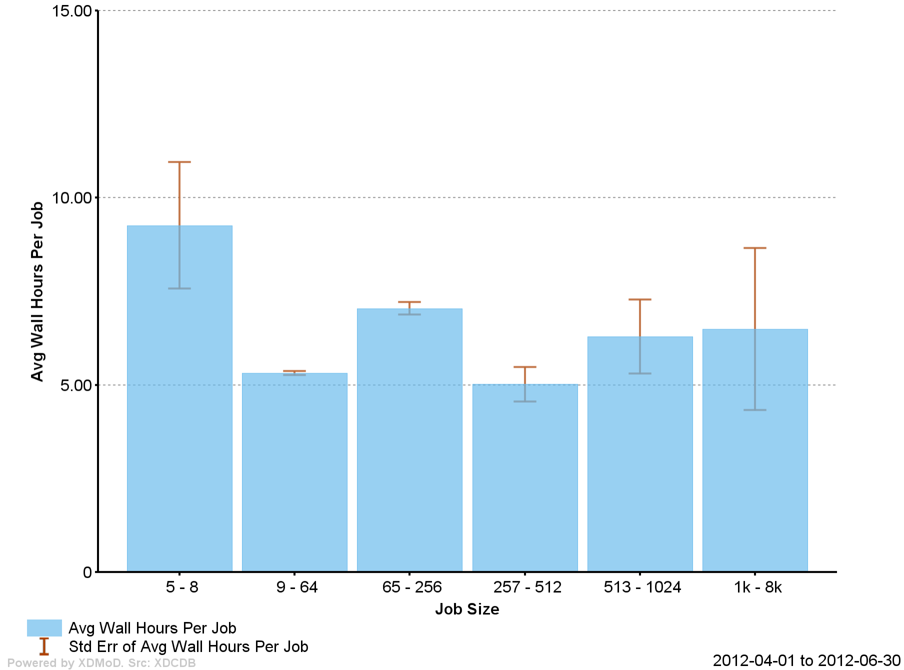
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = PSC-BLACKLIGHT

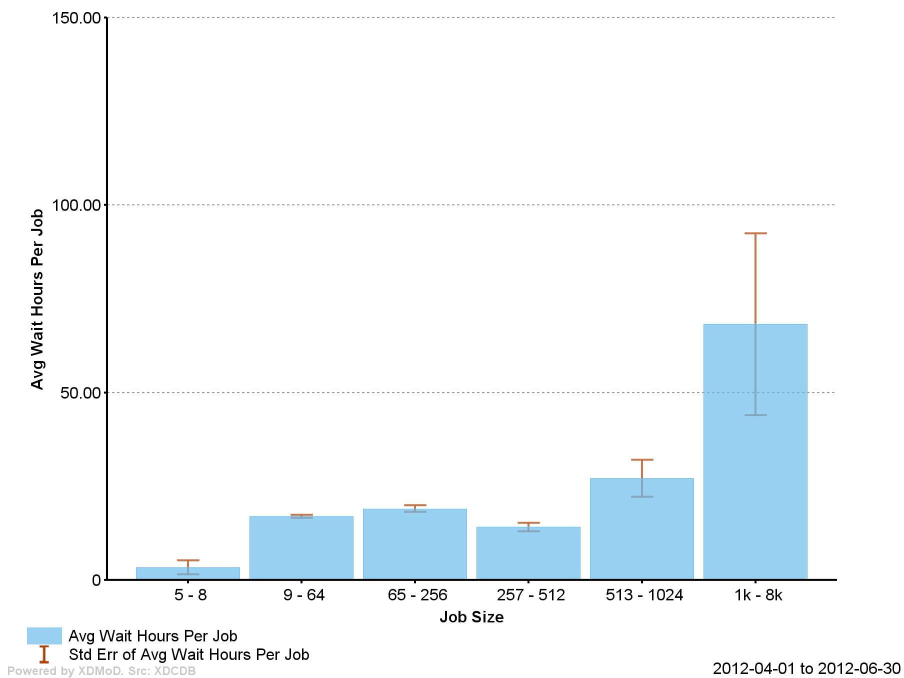
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = PSC-BLACKLIGHT

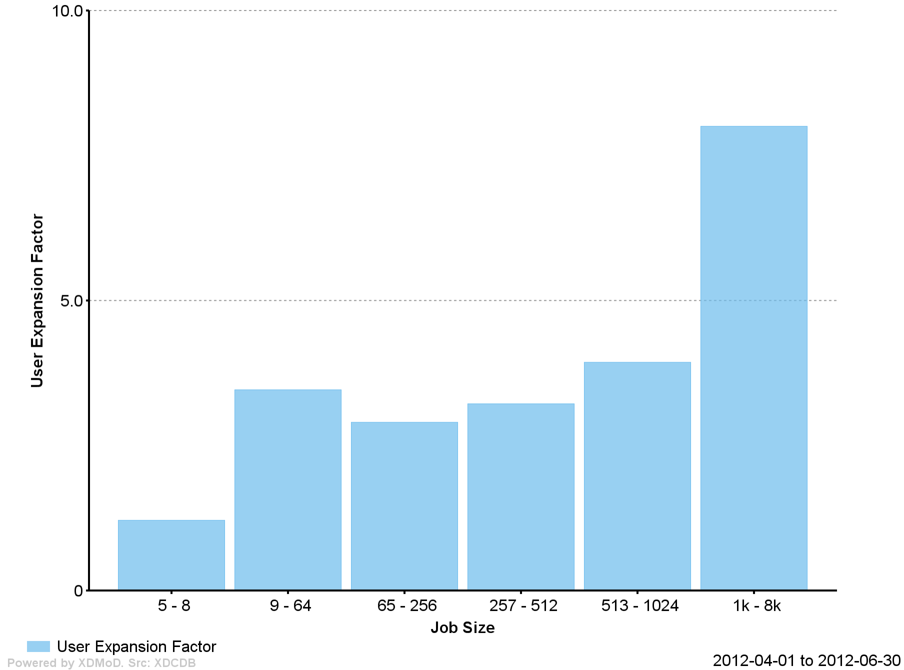
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = PSC-BLACKLIGHT

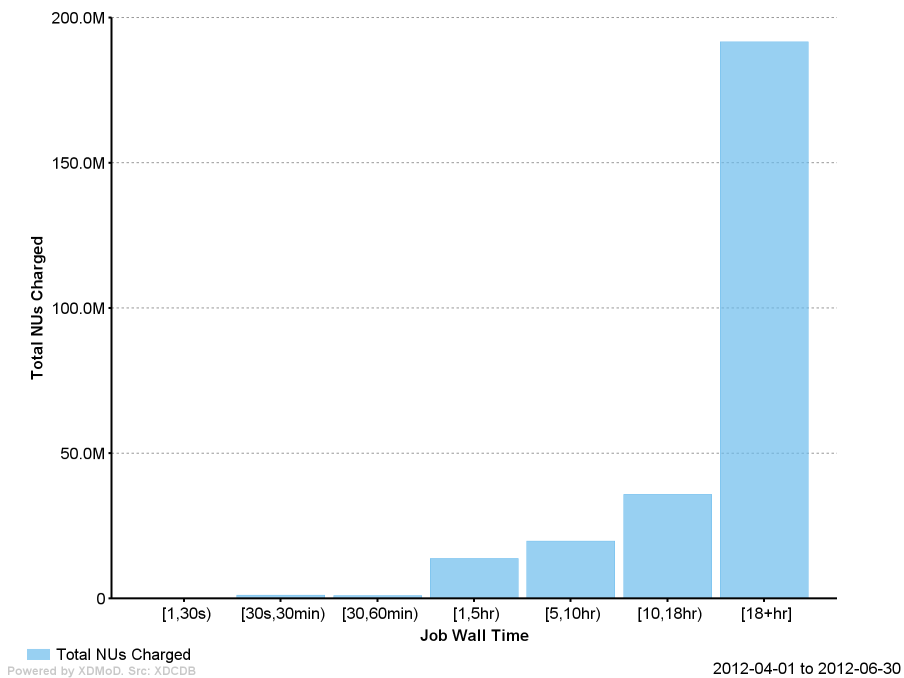
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = PSC-BLACKLIGHT

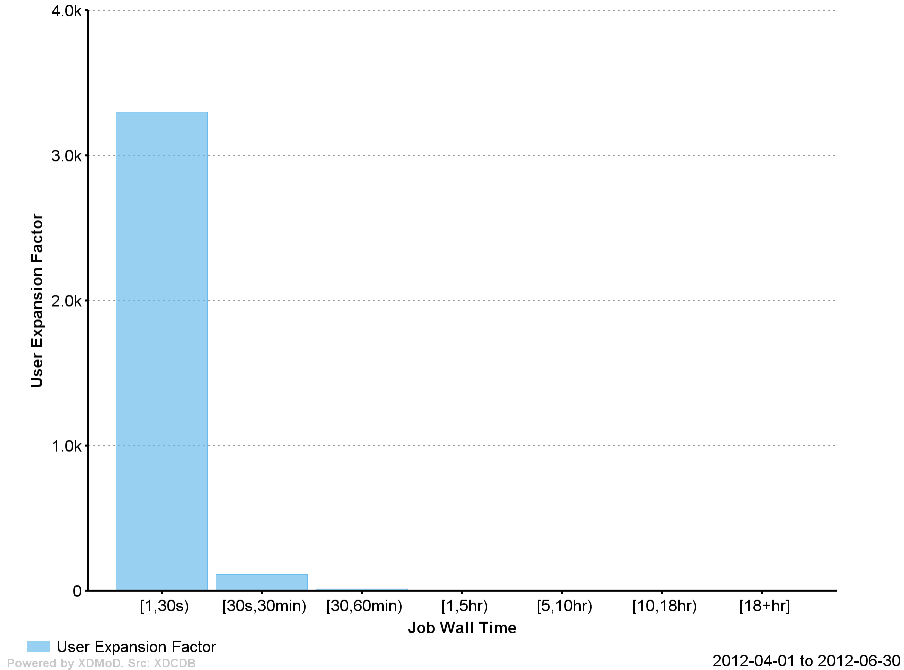
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = PSC-BLACKLIGHT

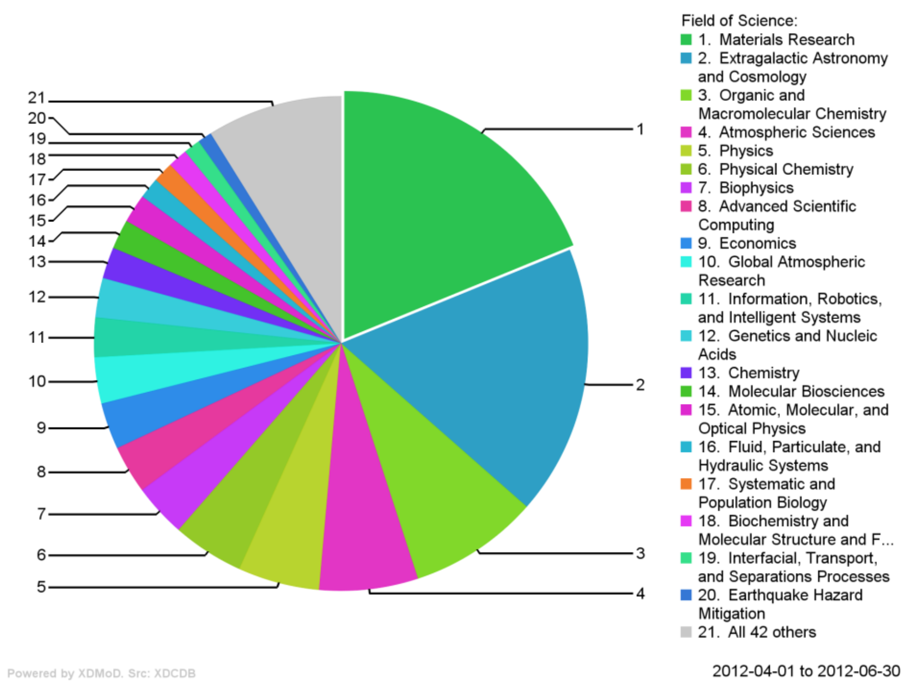
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = PSC-BLACKLIGHT

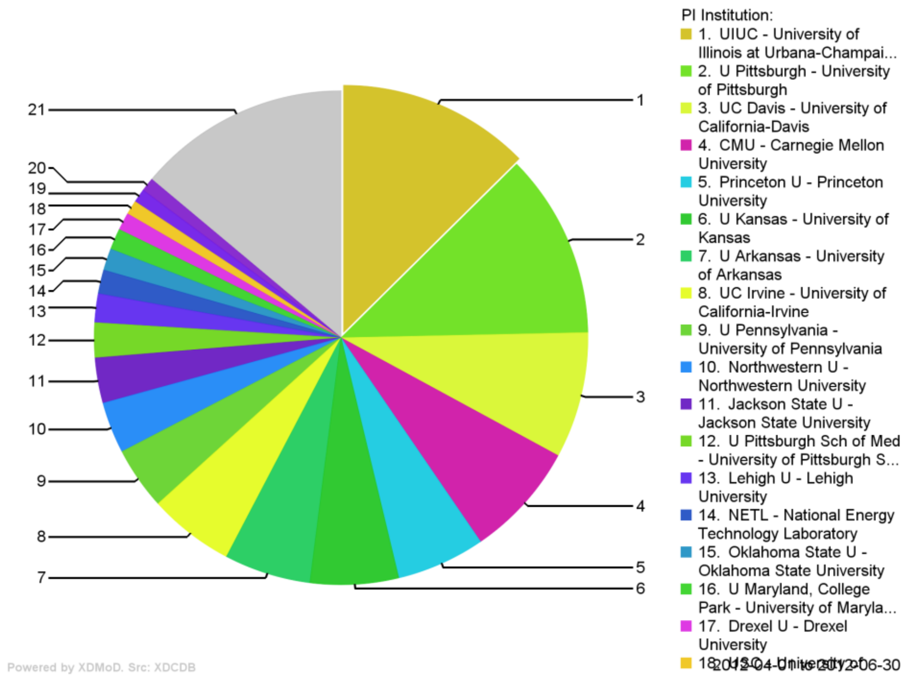
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = PSC-BLACKLIGHT

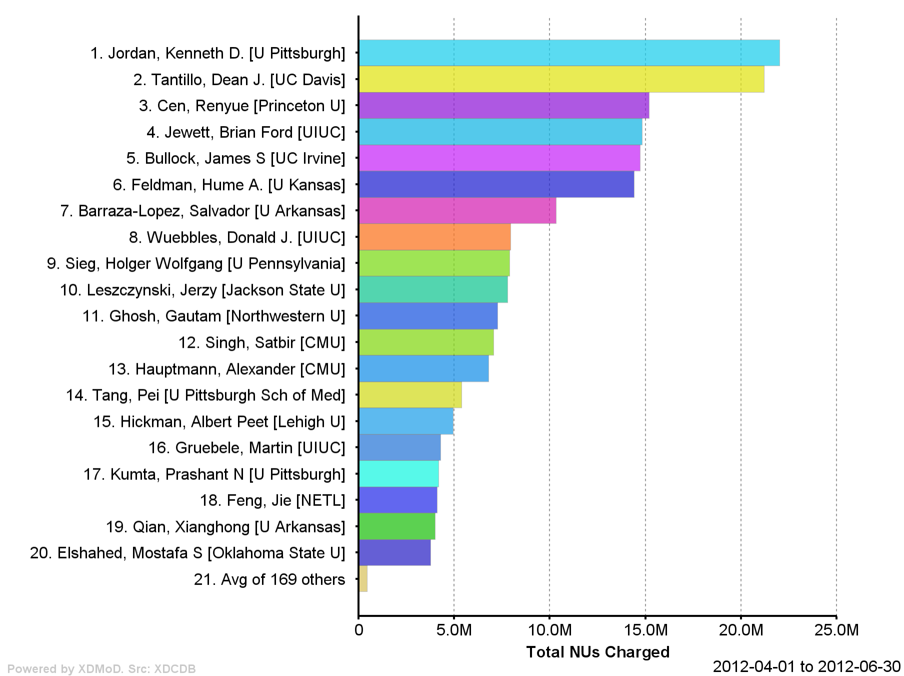
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = PSC-BLACKLIGHT

2012-04-01 to 2012-06-30



17 Purdue University - Service Provider Report

17.1 Executive Summary

Purdue continues to be a service provider (SP) to the XSEDE project as TeraGrid transitioned to XSEDE in July 2011, providing an HPC cluster (Steele), a high-throughput computing resource (the Purdue Condor pool), and a cloud resource (Wispy) to XSEDE users. The SP operates the systems and provides helpdesk and user support, as well as participate in XSEDE wide operations, security, software, training and outreach activities. Purdue contributes its expertise in HPC, high-throughput computing, virtualization and science gateway development to assist XSEDE users through training events, tutorials and demonstrations, as well as to the XSEDE ECSS staff on its conference calls and at conferences. These activities are funded by the NSF awards #0503992, #0932251. From Aug. 2011 to July 2012, Purdue XSEDE resources have supported 194 PIs with 203 allocations from 112 institutions.

As a service provider, Purdue continues to support a number of science gateways that utilize XSEDE data and computational resources, bridge OSG computation high-throughput HPC (HTHPC) jobs to XSEDE resource, and develop, deploy virtualization tools to support scientific users of cloud resources as part of the joint TeraGrid-OSG project ExTENCI. Purdue SP staff also supports and contributes to various XSEDE EOT activities and campus information technology community. For four years in a row (2009-2012), Purdue's Rosen Center for Advanced Computing has won the Campus Technology's Innovators awards to its DiaGrid (Condor pool), Community Cluster program, HUBzero and Purdue Studio of mobile applications for teaching and learning, including the resources that contribute to the national cyberinfrastructure TeraGrid/XSEDE.

Purdue's roles in XSEDE also include the following (funded separately by the XSEDE award):

- ECSS staff to provide expert support to user requests for advanced technical assistance.
- Lead for the XSEDE Campus Champion Program to further expand the XSEDE user base by reaching out to campuses, providing training and outreach activities, and getting feedback from campus users to improve XSEDE services.
- An XSEDE – OSG liaison to bridge the efforts of the two cyberinfrastructure projects by providing assistance and guidance to users in utilizing the OSG resources and identifying and organizing training and outreach activities to broaden user base.

These activities and impacts are reported in the overall XSEDE report.

17.1.1 *Resource Description*

Steele

The Steele cluster consists of 893 dual quad-core Dell 1950 compute nodes, running Red Hat Enterprise Linux, version 4. Each node thus has 8 64-bit Intel 2.33 GHz E5410 CPUs and either 16 GB or 32 GB of RAM. They are interconnected with either Gigabit Ethernet or InfiniBand. The machine offers access to the 120 TB scratch space. Steele's peak performance is rated at 66.59 TFLOPS. Steele cluster is well suited for a wide range of both serial and parallel jobs. Steele replaced the Purdue Lear cluster and was made available to TG users in May 2008. Its projected useful lifetime is through July 2013. In October 2009, Purdue RP has increased the TG *dedicated* portion of Steele from 22 nodes to nearly 200 nodes (1600 cores), which significantly reduced job wait time. Steele has no effective runtime limit on XSEDE jobs. Additionally, XSEDE users may leverage the larger Steele cluster by utilizing the standby queues with no node limit but subject to 4 or 8 hour runtime limits.

Condor Pool

The Purdue Condor pool is a shared resource among the resource owners (academic users at Purdue) and XSEDE/OSG users. Consisting of approximately 50,000 processor cores, the Condor pool is an opportunistic resource which allows Condor jobs access to machines that are not being used by their owners. The Purdue Condor pool is designed for high-throughput computing, and is excellent for parameter sweeps, Monte Carlo simulation, or most any serial application. In addition, some classes of parallel jobs (master-worker) may be run effectively in Condor. 30% of all Condor-usable cycles are available to XSEDE users at a minimum level of service. On average the Purdue Condor pool is able to provide up to 10 million CPU hours to XSEDE users per year.

The Purdue Condor resource, recently named DiaGrid, has expanded tremendously from a total of 7700 CPUs at the end of 2007 to its current size of about 50,000 cores (system information shown in Table 1). The Purdue Condor pool consists of nodes from 10 institutions, including Purdue's West Lafayette campus, University of Wisconsin-Madison, University of Nebraska-Lincoln, Indiana University, University of Notre Dame, and a number of Purdue's regional campuses. Memory on most of the compute nodes is 1 GB, 2 GB and 4 GB per core, while a small number of nodes have larger memory (e.g., 10GB per core). With a total of approximately 390 TFLOPS available, the Purdue Condor pool can provide large numbers of cycles in a short amount of time. All shared areas and software packages available on Steele are available on Condor. Available to TeraGrid/XSEDE users since 2006, the Condor pool is self-renewing as old machines in the pool are retired and new ones, e.g., from Purdue's community clusters, added over time.

Table 1: Purdue Condor pool information as of July 31, 2012

System Information	Cores	Total Memory	Local Interconnect	Processor Speed
X86_64 LINUX	49,100	129 TB	IB, 10 Gb or 1 Gb Ethernet	Various (2.1, 2.33, 2.5, 3.2 GHz)
INTEL & X86_64 WINDOWS	409	1.09 TB	1 Gb or 100Mb Ethernet	Various (2.13, 2,66, 3.6GHz)
INTEL LINUX	44	66 GB	1 Gb or 100Mb Ethernet	various
Total	**49,553**	**130 TB**		

Wispy

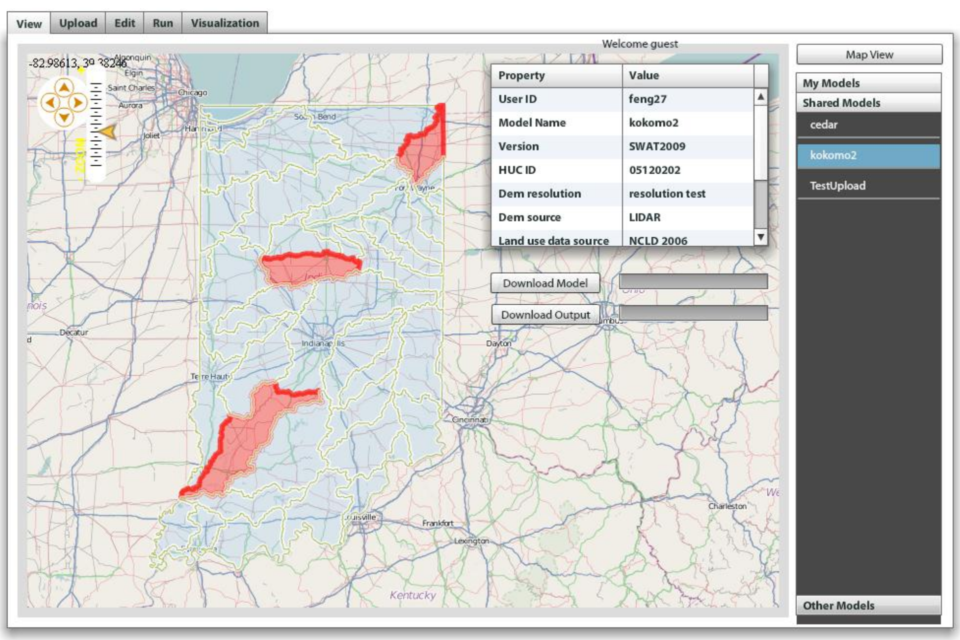
Purdue's *Wispy* is a special XSEDE resource, a cloud computing platform for research and education use. Wispy consists of 8 64bit, 16-core HP SL230 connected via 1 Gigabit Ethernet network with the capacity of supporting 128 VMs. Wispy runs KVM and the Nimbus cloud software. It provides users with the capability of packaging their applications and operating systems completely inside the Virtual Machine (VM) images, submitting these VMs to run in Wispy with up to 14 CPUs and 32GB of memory each, and have full control over the execution environment. Current usage includes small, instant, on-demand clusters for various tasks and running complicated or prepackaged applications on additional hardware resources.

17.2 Science Highlights

SWATShare – A Watershed Simulation and Collaboration tool for research and education
Venkatesh Merwade, Civil Engineering, Purdue University

Merwade leads a NSF funded CI-TEAM project (WaterHUB for Cyber Enabled Training, Education, and Research in Water Resources) that aims to support both research and education needs of the water community. A main goal: provide a much-needed cyberinfrastructure to train

and educate students, as well as the general public, on water-related issues. A product of the WaterHUB project, SWATShare is a new tool — powered by XSEDE computational resources and the HUBzero platform — that can be used by researchers studying water issues and by students who will become the next generation of scientists in the field. The Web-based model-sharing version of the Soil Water Assessment Tool (SWAT) not only enables students to run simulations and calibrations online, but also to publish, share and visualize results in studying how land use changes affect hydrology at watershed scale. SWATShare also allows students to make use of a plethora of publicly available water data rarely used in the classroom from the National Climatic Data Center, the U.S. Geological Survey and other sources. SWATShare utilizes the XSEDE HPC resources to calibrate and simulate using SWAT, a watershed-scale hydrologic model used by researchers around the world for simulating hydrologic cycle, water quality and agricultural management practices. Traditionally, the SWAT simulations are run on desktop computers. Calibration of such models is almost impossible on the desktops because it requires a lot of computation. Simulation of a large watershed is time-consuming on a desktop, too. In addition to these challenges, creating a SWAT input model involves complex steps and extensive knowledge of a set of tools, typically a process of days. Educators wish to use a powerful tool such as SWAT to teach the next generation hydrologists, but it has been impossible to use it in a classroom due to these difficulties. SWATShare not only supports the large scale computation that the watershed simulations require, but also support the sharing of SWAT input models – a user may publish a model that took days to create, and other users would be able to download the model, modify it and use it for their own simulations. SWATShare has made it possible to use real world simulation tools in a classroom. Students of a hydrology course (Purdue CE 549 Computational Watershed Hydrology), spring 2012, used SWATShare as a repository to share SWAT models.



The students have also participated in the tool development process through surveys to help define the user interface and workflow. Now researchers have created and made 25 SWAT Models available on SWATShare. Another hydrology class at Purdue will use this tool to learn in the fall of 2012. The SWATShare will be available to the public in the fall of 2012. The Purdue SP staff assisted the project by integrating SWATShare with XSEDE resource through the use of globus and an XSEDE community account. This work was presented at the XSEDE12 conference, July 2012 in Chicago, IL (See publication in 1.8). The students working on this project were funded by the XSEDE project to travel and participate in the conference.

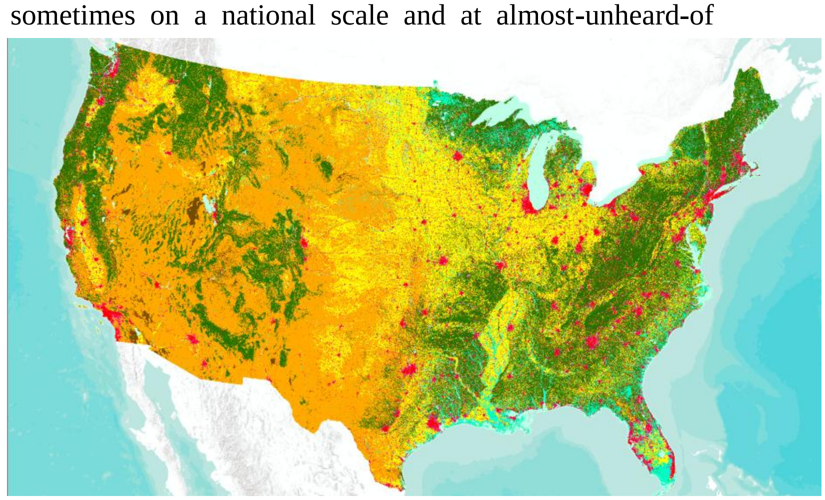
Figure: SWATShare tool interface. Available SWAT Models and their metadata are highlighted on the watershed shown on the map.

A large grid of model spectral energy distributions for young stellar objects
PI: Thomas Robitaille, Harvard University

Thomas Robitaille, formerly a Spitzer postdoctoral fellow at the Harvard-Smithsonian Center for Astrophysics now a researcher at the Max Planck Institute for Astronomy, studies galactic star formation, mostly through mid-infrared to millimeter wavelengths. In addition to instrumental data analysis, his work includes modeling multi-wavelength observations of young stars and studying the global properties of star formation in the Milky Way galaxy as they may apply to star formation in other galaxies that cannot be observed in the same level of detail as our own. In collaboration with Barbara Whitney of the University of Wisconsin and the Space Science Institute, he has computed a large grid of model spectral energy distributions of young stars, spanning a large range of evolutionary stages and stellar masses, and developed a tool to rapidly fit these models to a given set of observations. The models and the fitting tool are both publicly available and have been widely adopted by the star formation community. A new grid of models, addressing a number of issues with the current models, is to be released in 2012. He is using the Purdue Condor pool XSEDE resource to help develop this model grid, which will dramatically improve the parameter space coverage, the quality of the models and therefore the science that can be derived. This project has been the top Condor user on the Purdue Condor pool and has consumed more than 25M NUs in the past 12 months. (Robitaille, T. P., HYPERION: an open-source parallelized three-dimensional dust continuum radiative transfer code, Astronomy & Astrophysics, Volume 536, id.A79, 12/2011).

Windows HPC Cluster helping with data intensive computing
PI: Brian Pijanowski, Forestry and Natural Resources, Purdue University

Purdue SP staff has supported the Forestry and Natural Resources Professor Bryan Pijanowski's research to look at how people and land use play into the state of ecosystems, including issues like projecting loss of prime farmland to urbanization over the next 50 years and the potential effects of climate change on fish habitats. His lab examines such questions with computer modeling, among other techniques, sometimes on a national scale and at almost-unheard-of resolutions as fine as every 30 meters. The work can involve hundreds of gigabytes to terabytes of data to begin with, trillions of data points and software that runs — in some cases will only run — in Microsoft Windows. In the past, that could mean running a model for months on a desktop computer. But the new cluster supercomputer using the high-performance computing version of Windows has reduced the time it takes to get results to a few hours and is



enabling more detailed and accurate modeling. “We could actually run the entire world at 30-meter resolution if we wanted,” Pijanowski says.

While similar in functionality to Purdue’s other community cluster supercomputers, the Windows HPC cluster offers potential advantages for some users and developers. Users get an entry point to high-performance computing with the familiar Windows user interface. The Windows HPC cluster makes high-performance computing available to researchers who rely on Windows-only software like ArcGIS and Microsoft Excel. In addition, the new cluster can run Windows HPC versions of research software common on Purdue’s Linux community clusters, such as MATLAB and SAS.

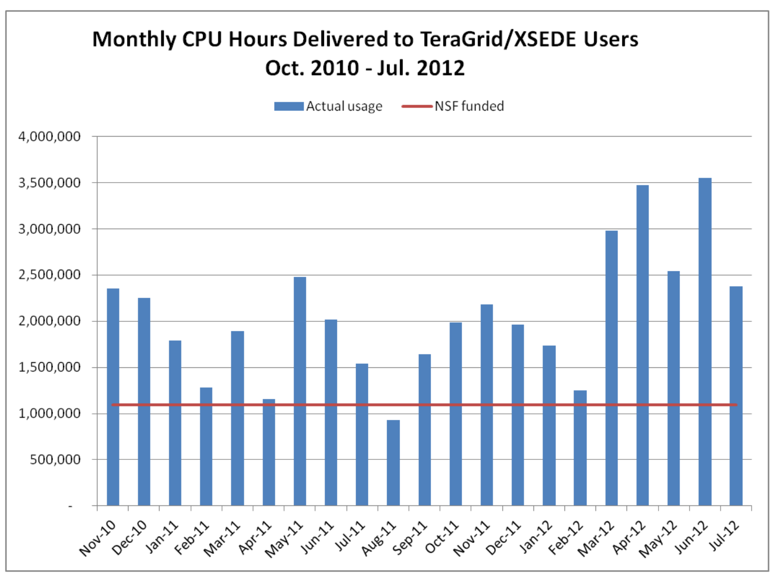
The SP staff also helped the group to visualize the data, showing the extent of urban (red) for the year 2040 across the lower 48 states. Also depicted is forest (green), agriculture (yellow), grassland and pasture (orange), barren (brown), wetlands/marshland (aqua) and surface water (light blue). Many cities double in their footprint over the next 30-40 years which would result in more impervious surface, less natural habitat for wildlife and more non-point source pollutants into fresh water systems. This visualization was showcased at the XSEDE12 conference’s Visualization Gallery in July 2012 in Chicago, IL.

17.3 User-facing Activities

17.3.1 System Activities

The Steele cluster continues to be busy and highly utilized by the XSEDE users in the past 12 months. It has maintained 98% uptime throughout the reporting period. In May 2012, Steele had a scheduled maintenance to update its scheduler to Moab and Torque. Steele continues to see higher allocation requests than its available cycles each quarter, and the trend is even higher demand in the September allocation cycle (as high as 62M SUs while only 8M SUs are available). XSEDE users access Steele through the XSEDE queues, and in addition to the NSF funded portion of the cluster, XSEDE users have access to the entire cluster through its standby queues with a wall clock limit of 4 hours for the jobs. In this manner, XSEDE users typically consumes as high as three times of the cycles allocated for XSEDE on a monthly basis. The figure to the right side illustrates the total usage on Steele by XSEDE users since November 2010.

The Purdue’s Condor pool continues to grow in size as Purdue’s community cluster program expands. All of the community clusters run Condor. Purdue’s Condor continues to serve the OSG community, as well as the XSEDE community, in addition to the campuses that are part of the DiaGrid initiative.



Wispy is an experiment cloud system in XSEDE. Its usage includes those from the joint TeraGrid/OSG ExTENCI project in the development, testing of virtualization technologies for science users. Wispy has recently been upgrade in July 2012 to new hardware based on Intel’s Sandy Bridge series of processors. The SP staff has revamped the authentication

on Wispy, simplifying the process for XSEDE users and providing clear documentation for Wispy access (https://www.xsede.org/web/guest/purdue-wispy). Currently, Wispy is available to XSEDE users. Non-XSEDE users can request access directly from Purdue SP.

Other systems: As part of its award-winning campus community cluster program, the Purdue Rosen Center deployed its fourth large cluster, Hansen, in September 2011, and its fifth and most powerful community cluster, Carter, which began full operation at the end of March 2012. Carter was built through a partnership among Purdue, HP, Intel and Mellanox using upcoming technologies. Carter was ranked 54th on the latest international Top500.org list of the world's most powerful supercomputers in November 2011. The new cluster features HP compute nodes with two eight-core Intel Xeon E-5 “Sandy Bridge” processors, 16 cores per node, 32 gigabytes of RAM, a 500 GB system disk and 56 gigabits per second InfiniBand interconnects. ***Though Carter is dedicated for campus use, its nodes, as well as nodes in all Purdue campus clusters, are in the Purdue Condor pool, available to XSEDE and OSG users***.

Purdue has recently deployed a new cluster that runs Windows HPC Server, the latest version of Microsoft’s operating system for high-performance computing. Windows HPC is an alternative to Linux-based clustering software packages, with similar features such as cluster management tools, a job scheduler and a message passing interface (MPI) library. This cluster is enabling many faculty, many in earth sciences, whose research requires data intensive computing on a Windows platform (see science highlight in Section 1.2).

17.3.2 Services Activities

Purdue SP provides both helpdesk support and consulting support to XSEDE users. The SP user support staff worked with many XSEDE users during the year. Categories of user issues and requests are summarized in the table in Section 1.9.2. In many cases, SP staff interacts with the user over an extended period of time from diagnosing user code problems to providing solutions and alternatives. Some of the recent highlights are included below:

The SP staff assisted the CMU group (PI Maria Kurnikova) to get started on Steele, consulted on storage systems and jobs submission. We made adjustment to the disk quota to enable them to run their molecular simulations of the ion channel and receptor proteins in their study of atomistic details of the low pH triggered membrane insertion pathways of diphtheria toxin with a goal of better understanding of the interactions underlying the membrane protein assembly and stability. They also assisted the Boston U group (PI John Straub) that studies the sequence-dependent properties that lead to aggregation of peptides and proteins to better understand diseases such as Alzheimer's disease, type II diabetes, and Parkinson's disease. The user had questions about NAMD on Steele post-scheduler conversion, observed runtime differences on Steele for same size jobs, connection errors when submitting NAMD jobs. Our staff assisted them with answers about the scheduler, and provided workaround for submitting NAMD jobs, and recommended them to use entire nodes for their computation. The staff also helped an Iowa State group (PI Abhijit Chandra) access Steele and provided information regarding job submission, using Condor, and running MATLAB. Using Steele, this group aims at developing an integrated modeling framework for product quality and product integrity in the Chemical Mechanical Planarization (CMP) used in IC manufacturing process. The SP staff provided extended support to the Georgia Tech group (PI Cyrus Aidun) that is developing a coupled lattice-Boltzmann and finite element method for modeling fluid-structure interaction. The staff provided higher disk quota so he was able to run more jobs at once and investigated alternatives that would produce higher turnaround.

Purdue’s SP staff, also a Purdue campus champion (Kim Dillman), has identified 24 potential contacts to discuss possible access to XSEDE resources. A few of these are all students in a class

that might benefit from access to XSEDE resources. Her work with local users often translates into training materials for the larger XSEDE community. She actively participates in the face-to-face working meetings and phone meetings for both the Outreach and User Assistance working groups (with her dual role as Purdue Campus Champion and XSEDE Staff) and recently worked on various documents for XSEDE12 for these two groups.

The SP staff also assisted with completing the CESM (Community Earth System Model) gateway as a TeraGrid project and transitioning it into XSEDE. The original objective, set out by the TG GIG project, of developing a prototype of an integrated climate modeling environment that supports end-to-end model execution on TG/XD, metadata generation, data/metadata publishing to ESG and output analysis. The prototype has been completed, including the installation of the latest release of CESM plus a new model baseline with significantly enhanced metadata capability, a CESM gateway integrated with the ESG data and metadata publishing and advanced data analysis (using LAS – Live Access Server) and integrated data transfer capability for users to transfer large input and output files related to climate model simulations. The integrated workflow was demonstrated at an XSEDE science gateway meeting in May. The SP has set up a production site and a development site in preparation for the formal release of the CESM gateway. Purdue SP staff will provide support to the CESM gateway users, especially to students and instructors of classes that will use the gateway in teaching and learning. The SP staff helped fixed several problems in the C4E4 gateway caused by changes on the backend cluster system. The gateway was successfully used in a graduate class ABE 522 (Ecohydrology). Students ran SWAT auto calibration simulations which entails long running jobs that would not be feasible on regular lab computers. Positive feedbacks were received at the end of the semester.

Purdue SP also assisted the IsoMAP science gateway integration with XSEDE. The IsoMAP portal provides an online workspace that helps researchers access and integrate a number of disparate datasets, develop Isoscapes models over selected spatio-temporal domains, and predict maps for the stable isotope ratios of water, plants, and soils. It leverages the XSEDE computation resources to perform geospatial data processing and geostatistical model calculations. The portal was released in March 2011 and supports more than 500 registered users. In the past 12 months 2724 unique users visited the IsoMAP portal (Google Analytics). These users came from 50 countries (Figure 1). Among them, 96 registered users submitted 1013 model fitting and prediction jobs through the IsoMAP portal.

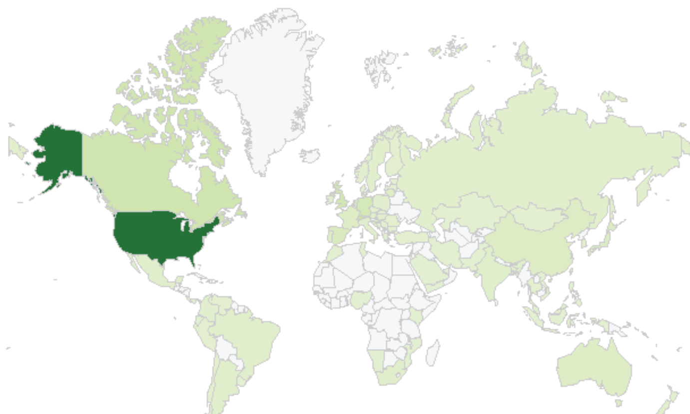
Location

Aug 1, 2011 - Jul 30, 2012

% of visits: 100.00%

Map Overlay

Site Usage



Visits	Pages / Visit	Avg. Visit Duration	% New Visits	Bounce Rate
4,920	**2.34**	**00:02:13**	**53.23%**	**64.37%**
% of Total: 100.00% (4,920)	Site Avg: 2.34 (0.00%)	Site Avg: 00:02:13 (0.00%)	Site Avg: 53.23% (0.00%)	Site Avg: 64.37% (0.00%)

	Country / Territory	Visits	Pages / Visit	Avg. Visit Duration	% New Visits	Bounce Rate
1.	United States	**2,568**	2.34	00:02:13	44.08%	63.40%
2.	Switzerland	**326**	1.31	00:00:24	64.72%	91.10%
3.	Canada	**259**	2.34	00:02:15	49.81%	62.93%
4.	United Kingdom	**222**	2.49	00:02:04	66.67%	59.46%
5.	Germany	**193**	3.11	00:02:39	65.28%	49.22%
6.	Netherlands	**129**	1.83	00:02:41	30.23%	74.42%

Figure: Demographic breakdown of IsoMAP users from Aug 1, 2011 to July 30, 2012.

17.4 Security

No security incidents were reported during this quarter.

17.5 Education, Outreach, and Training Activities

17.5.1 EOT Events

Type	Title	Location	Date(s)	Hours	Number of Participants	Number of Under-represented people	Method
Presentation	Development of an Integrated Earth System Climate Modeling Environment, by Lan Zhao, Chris Thompson, Carol Song	XSEDE Science Gateway conference call	May 11, 2012	0.5	Unknown	unknown	Presentation
Community newspaper	Cyberinfrastructure for Environmental Isotope Analysis and Modeling	AGU Eos Vol. 93, No. 19	May 8, 2012	n/a	Exposure to > 61,000 earth and space scientists	unknown	article

					worldwide		
Training	Campus Champions 101 training (aka New Champions Training) by Kim Dillman	New Champions training	June 12, 2012	1.5		unknown	Presentation
Student Training	Using XSEDE Resources Effectively, by Kim Dillman	webinar	June 26, 2012	1.5	10	unknown	Presentation
Tutorial	Selecting and Using XSEDE Resources for Maximum Productivity, by Kim Dillman	XSEDE12, Chicago, IL	July 16, 2012	4	Approx. 20	unknown	Presentation, hands-on
Tutorial	High-Throughput Computing With XSEDE, Ben Cotton, Kim Dillman, Mats Rynge, Miron Livny	XSEDE12, Chicago, IL	July 16, 2012	4	10	unknown	Presentation, hands-on
Visualization showcase	Simulating land use change at a national scale with fine resolutions Authors : G. Takahashi, J. Doucette, B. Pijanowski, D.Braun and C. Song	XSEDE12, Chicago, IL	July 18, 2012	2	200?	Unknown	demonstration
Panel	Campus Champions Panel, hosted by Kay Hunt	XSEDE12, Chicago, IL	July 19, 2012	1	175	~30 (estimated)	Panel discussion
Meeting	Campus Champions Breakfast Meeting (with people from Wales, UK, who look to start a similar program)	XSEDE12, Chicago, IL	July 19, 2012	1	10	0	Discussion
Meeting	Campus Champions meeting	XSEDE12, Chicago, IL	July 16, 2012	2	115	~25 (estimated)	Discussion
Meeting	Campus Champions Focus Group meeting	XSEDE12, Chicago, IL	July 17, 2012	2	120	~15 (estimated)	Presentation, discussion

17.5.2 Education

Purdue SP continues to recruit students to join the SP staff team and train them in HPC system administration, software maintenance, and scientific application support. The SP has funded one undergraduate student, Wesley Weber, to assist with system administration tasks. Wesley is a first year computer engineering student. He handles hardware fixes and front line support for the SP staff. We also have several graduate students to support the operations and users of the SP's XSEDE resources, as well as application support to improve services and capabilities, e.g., enabling online climate model runs using CESM, soil water assessment model (SWAT) and land data assimilation model (HRLDAS) on XSEDE resources. ***Several undergraduate students, including Alex Younts, Andrew Howard, Dan Dietz, Brian Raub, and John Michael, who assisted the project as SP staff in the past have gone on to have a career in the area of HPC and scientific computing applications.*** The SP staff has mentored several students through the

Purdue Discovery Park Undergraduate Research Internship (DURI) and the Purdue Summer Undergraduate Research Fellowship (SURF) programs throughout the school year. Five graduate students are on the Purdue SP staff. They helped users from running their tools on XSEDE resources to helping their research project integrate with XSEDE resources. One of the students assisted a CI-TEAM project and was awarded travel grant to attend XSEDE. A number of graduate students who were on SP staff have since graduated and went to work on computing projects at places such as Google and Twitter.

The SP staff has been supporting two classes in using XSEDE science gateways: a graduate class on ecohydrology (Purdue ABE 522) and a Civil Engineering graduate course (CE549 - Computational Watershed Hydrology), during the 2011-2012 school year. The staff worked with the instructors before the class began and throughout the semesters to ensure that their specific requirements are met by the gateways and that problems and issues resolved promptly. Among the lessons learned are that the need for data sharing and data transfer often varies from class to class, and timely solutions are critical during the semesters in order for the students to complete their assignments.

17.6 SP Collaborations

Purdue SP has contributed and continues to collaborate with OSG in multiple ways. Purdue's Condor pool of approximately 50,000 processors are an OSG resource to support high-throughput scientific applications. Purdue has also worked with OSG staff to support the OSG HTHPC (high-throughput HPC, aka OSG MPI, single core) jobs on Steele. OSG MPI jobs can flow to Purdue's Steele easily thanks to previous work Purdue staff did to standardize methods for advertising support for parallel (MPI) jobs on the OSG hence making site selection, compilation, and submission of MPI jobs easier for the end-user. During the period of 8/1/2011 through 8/1/2012, a total of 315,837 OSG MPI jobs ran on Steele and consumed 233.5K processor-core hours on Steele. During the same period, various OSG VO serial jobs, totally more than 4 millions, consumed a total of approximately 12 million processor hours in Purdue's Condor pool from 8/1/2011 to 7/31/2012.

Purdue SP is contributing to and coordinating the virtualization project within the ExTENCI project. The detailed report on Purdue activities related to ExTENCI is included in the overall XSEDE report.

17.7 SP-Specific Activities

The SP staff has investigated ways to make Wispy access easier for end users. Purdue has been collaborating with U of Wisconsin, Fermilab, and Clemson, etc on virtualization technologies to help XSEDE and OSG users. Wispy is being used in this project as a testbed for running scientific applications in virtual machines on it. The SP staff provided access to system staff at other sites and, through this process, we have identified issues and ways to improve access by automating the authentication steps. The SP is also developing tools for easier access to and control of virtual machines that run on Wispy. This part of the work is now integrated into the ExTENCI project.

WinHPC is a compute cluster operated by ITaP's Rosen Center for Advanced Computing (RCAC) and is a member of Purdue's Community Cluster Program. WinHPC went into production on December 1, 2011. WinHPC consists of HP compute nodes with two 12-core AMD Opteron 6172 processors (24 cores per node), 48 GB of memory, and 250 GB of local disk for system software and local scratch storage. All nodes have 10 Gigabit Ethernet interconnects. While similar in functionality to Purdue's other community cluster supercomputers, the Windows HPC cluster offers potential advantages for some users and developers. Users get an entry point to high-performance computing with the familiar Windows user interface. The Windows HPC

cluster makes high-performance computing available to a broader user community that includes researchers whose research computing jobs rely on Windows-only software like ArcGIS and Microsoft Excel. In addition, the new cluster supports Windows HPC versions of research software common on Purdue's Linux community clusters, such as MATLAB and SAS.

17.8 Publications

Publications by researchers who utilized SP resources (PI's name in **boldface** text**):**

1. **Robitaille**, T. P., HYPERION: an open-source parallelized three-dimensional dust continuum radiative transfer code, Astronomy & Astrophysics, Volume 536, id.A79, 12/2011
2. T. Kim, K.I. Lee, P. Morris, R.W. Pastor, O.S. Andersen, and **W. Im**, Influence of Hydrophobic Mismatch on Structures and Dynamics of Gramicidin A and Lipid Bilayers. *Biophys. J. 102:1551-1560* (2012).
3. X. Cheng and **W. Im**, NMR Observable-Based Structure Refinement of DAP12-NKG2C Activating Immunoreceptor Complex in Explicit Membranes. *Biophys. J. 102:L27-L29* (2012).
4. S. Park, T. Kim, and **W. Im**, Transmembrane Helix Assembly with Window Exchange Umbrella Sampling. *Phys. Rev. Lett. 108:108102* (2012).
5. K.I. Lee, S. Jo, H. Rui, B. Egwolf, B. Roux, R.W. Pastor, and **W. Im**, Web Interface for Brownian Dynamics Simulation of Ion Transport and Its Applications to Beta-Barrel Pores. *J. Comput. Chem.* 33:331-339 (2012). *[cover]*

Staff publications

1. L. Zhao, W. Feng, V. Merwade, C. Song. "WaterHUB – A Resource for Students and Educators for Learning Hydrology", XSEDE 2012 conference, July 16 – 20, 2012, Chicago, IL.
2. G. J. Bowen, J. B. West, L. Zhao, G. Takahashi, C. C. Miller, and T. Zhang, "Cyberinfrastructure for Environmental Isotope Analysis and Modeling", AGU Eos newspaper.

17.9 Metrics

17.9.1 Standard User Assistance Metrics

Q2-2012 Purdue Ticket resolution times by category from XSEDE Ticket System

Time to Resolution	account issues	file systems	grid software	jobs/batch queues	login/access issues	mss/data issues	network issues	software/apps	system issues	other
0-1 hr										
1-24 hr				2	5			1	3	4
1-7 d				12	6	1		1	14	8
1-2 wk				7		1		1	6	1
> 2 wk				16	1		1	5	4	5
Still Open				1		1		1		

17.9.2 SP-specific Metrics

None.

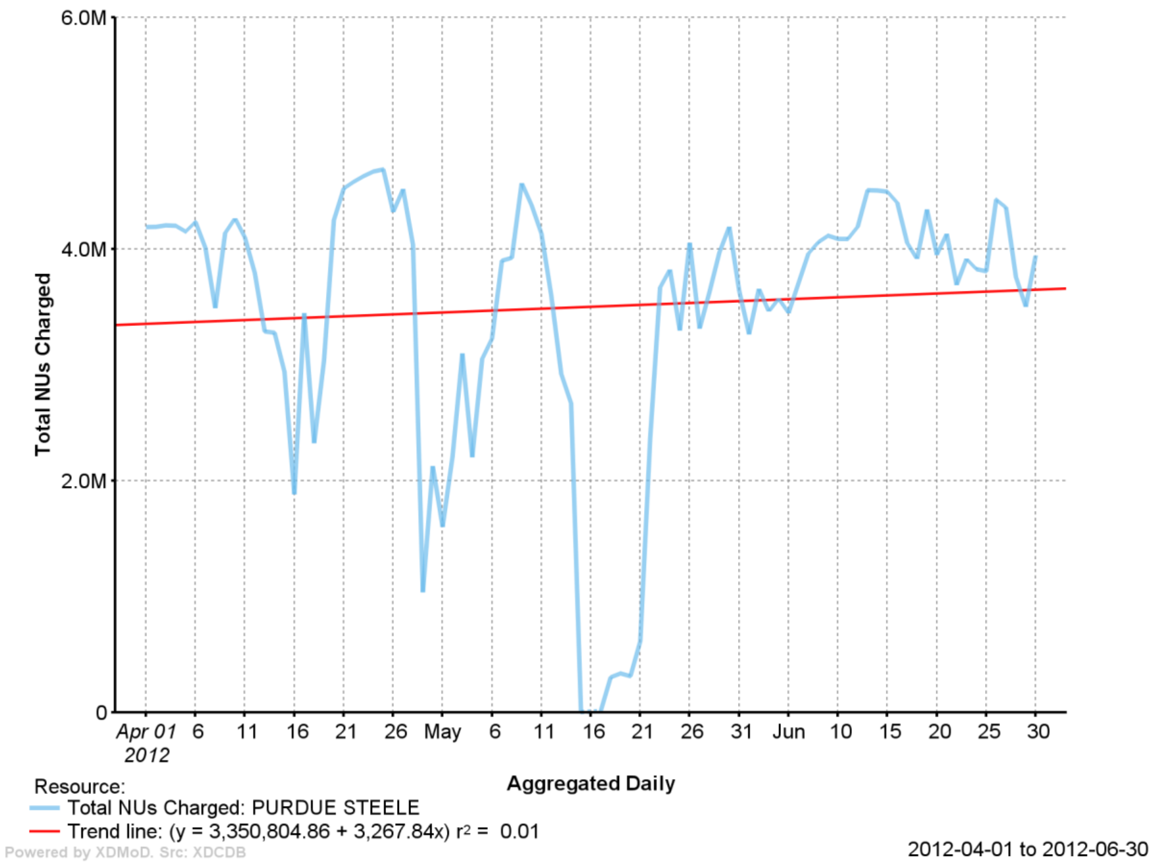
17.9.3 Standard systems metrics

See charts below.

Total NUs Charged by Resource

Resource = PURDUE-STEELE

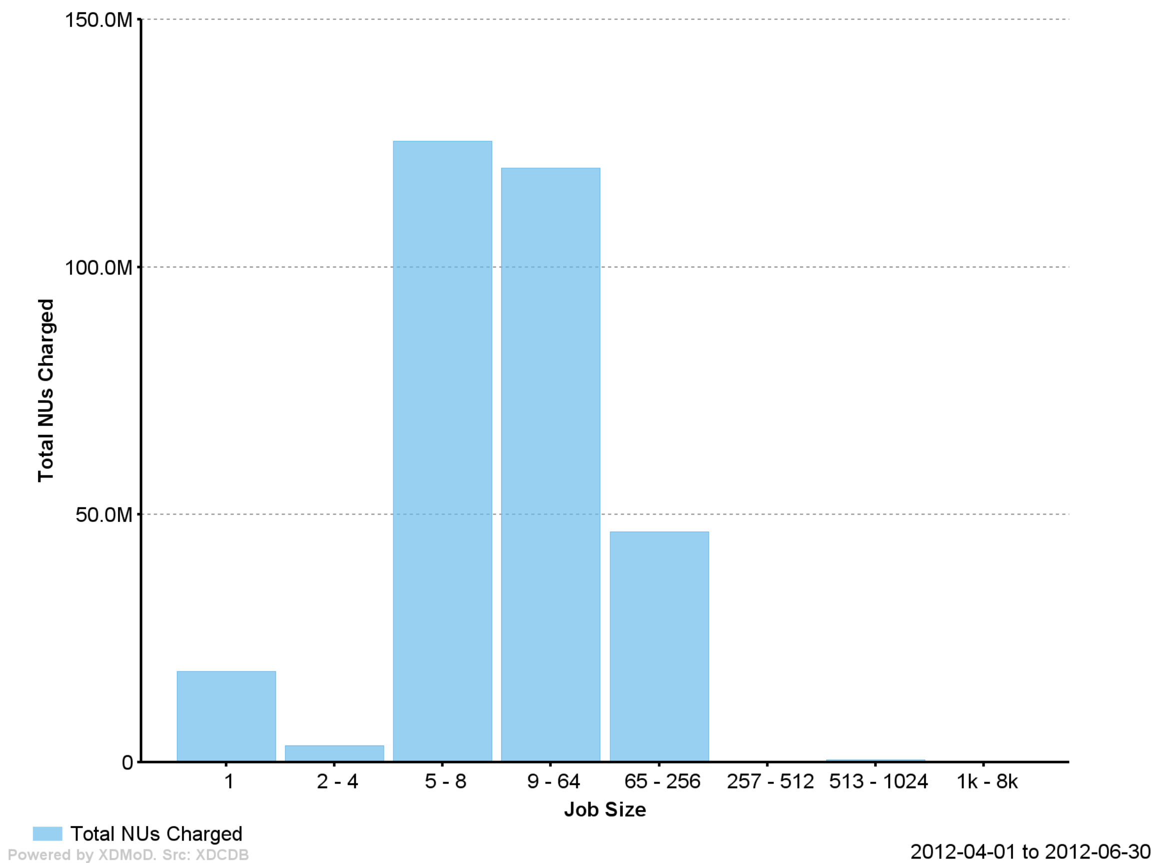
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = PURDUE-STEELE

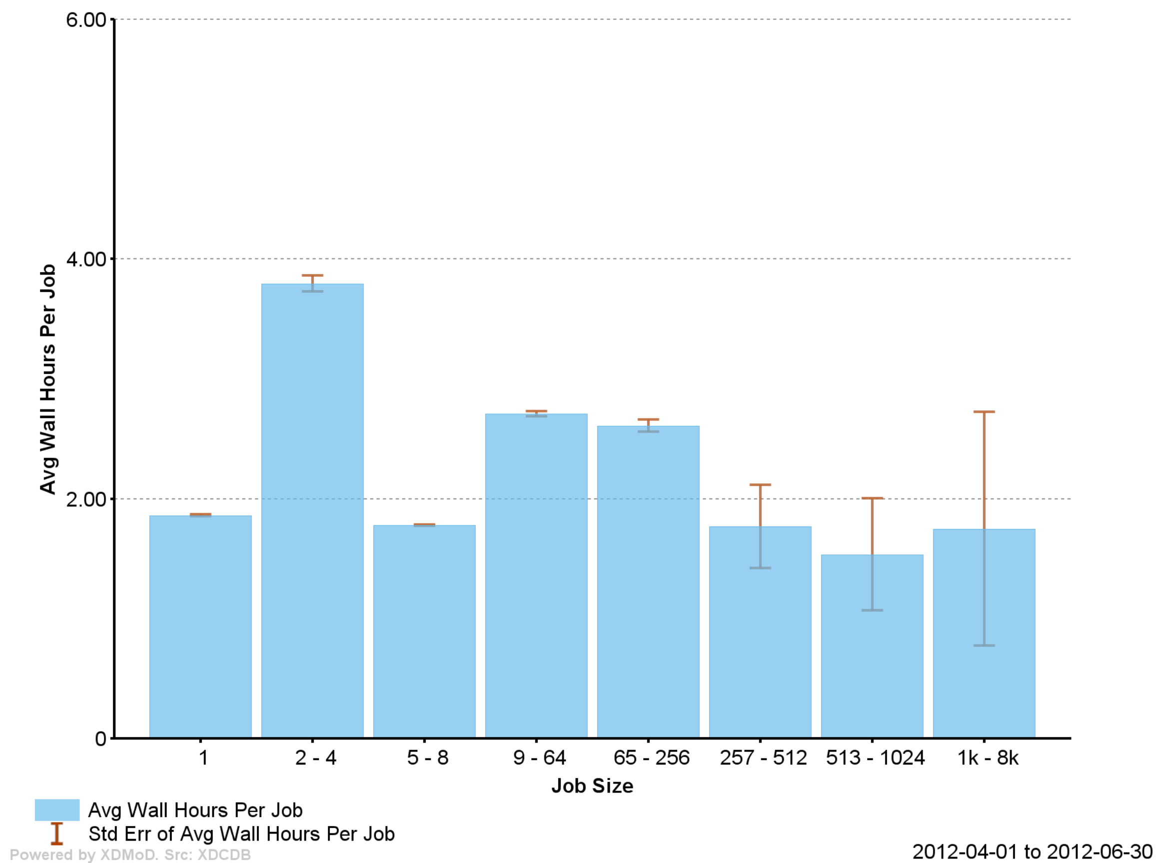
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = PURDUE-STEELE

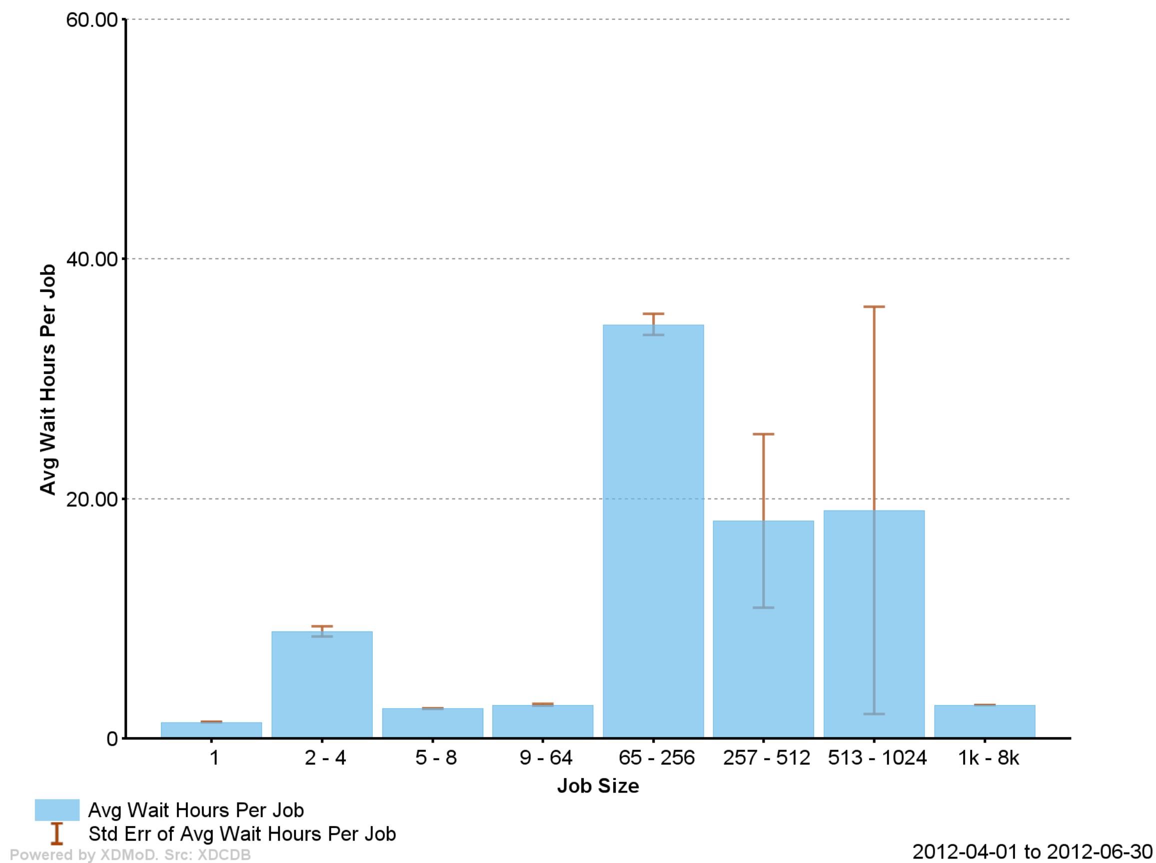
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = PURDUE-STEELE

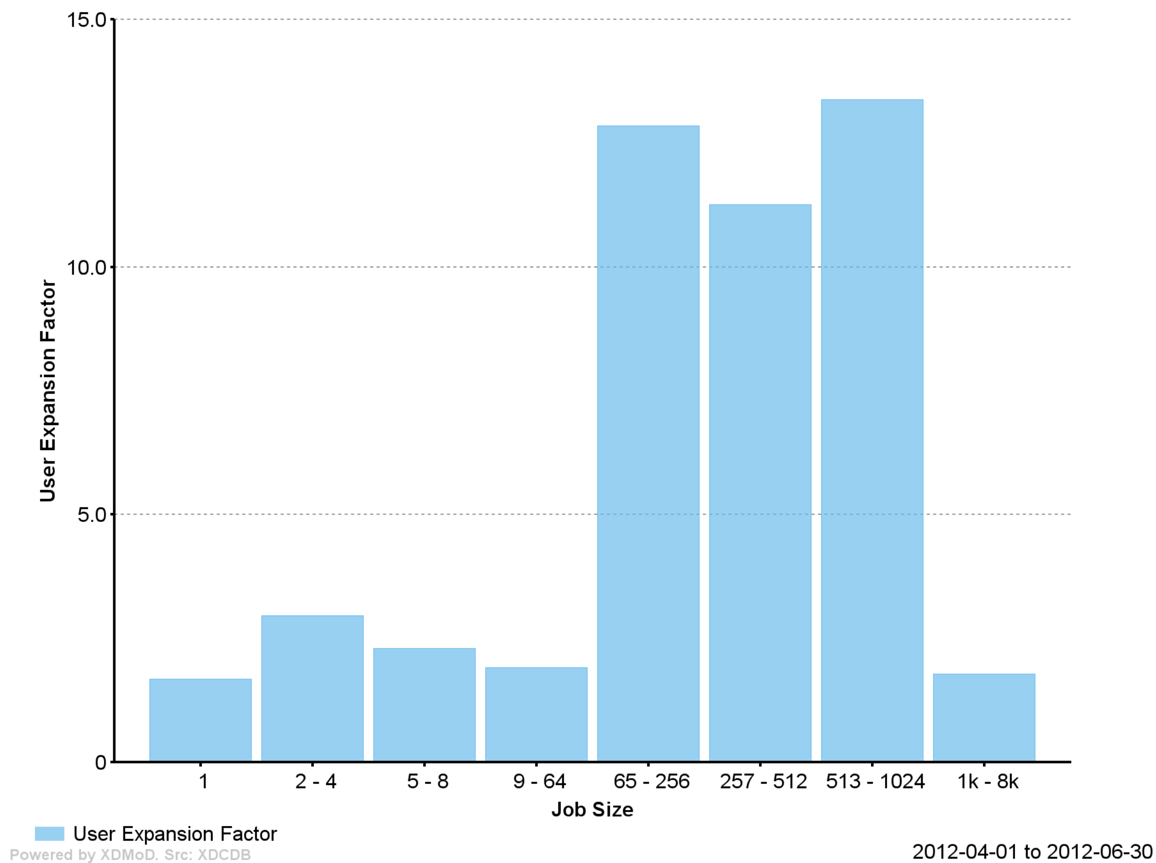
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = PURDUE-STEELE

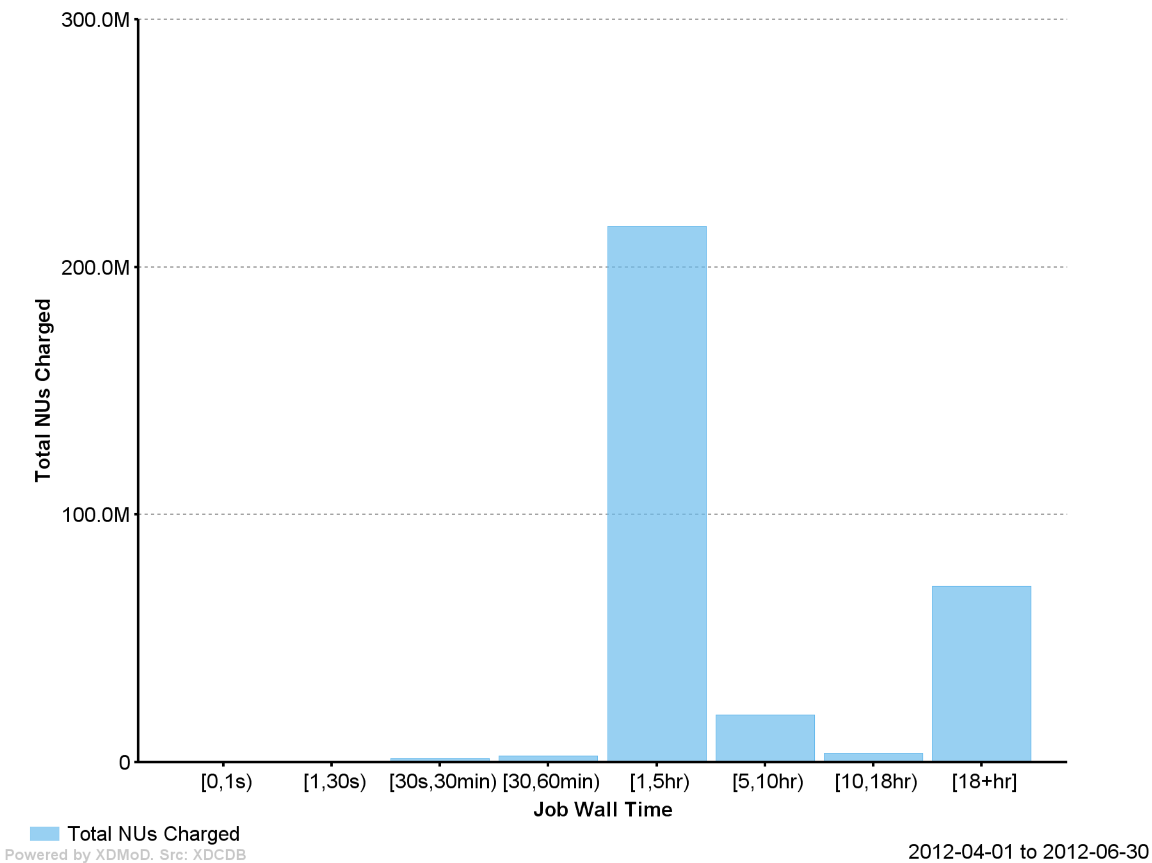
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = PURDUE-STEELE

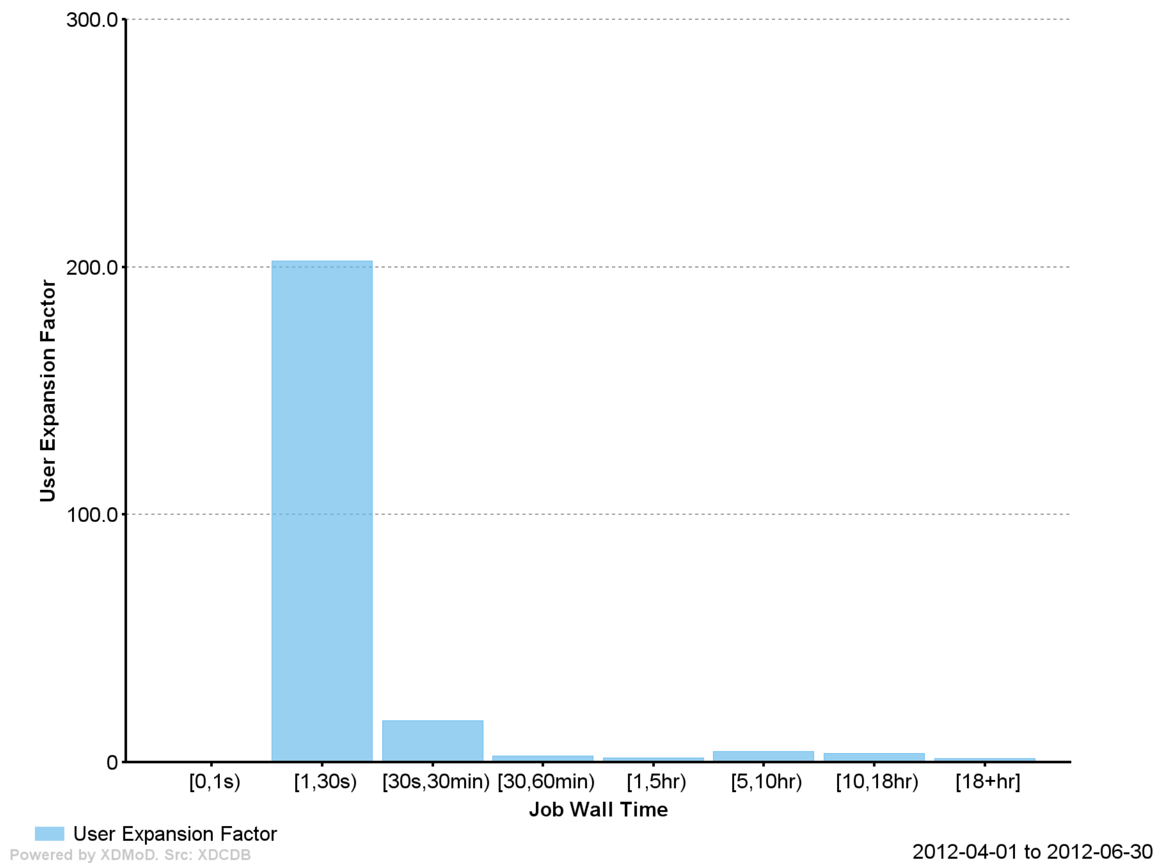
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = PURDUE-STEELE

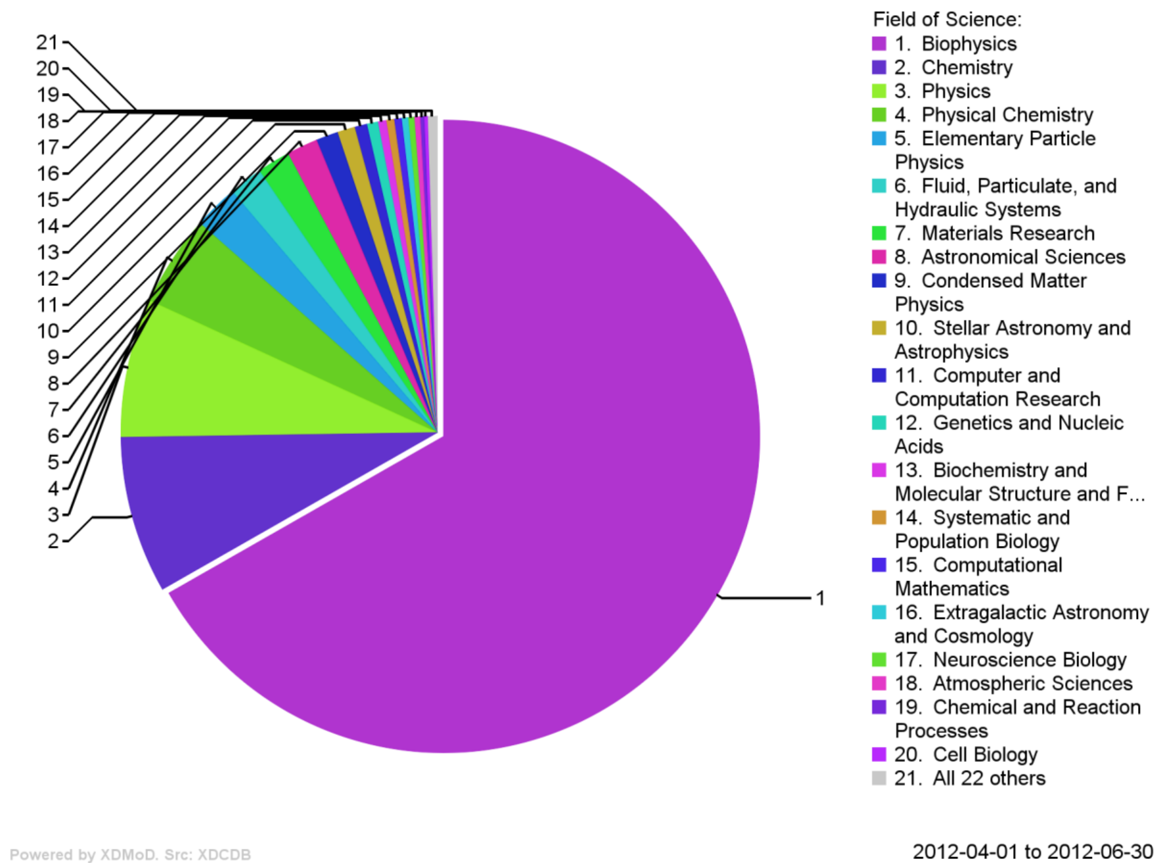
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = PURDUE-STEELE

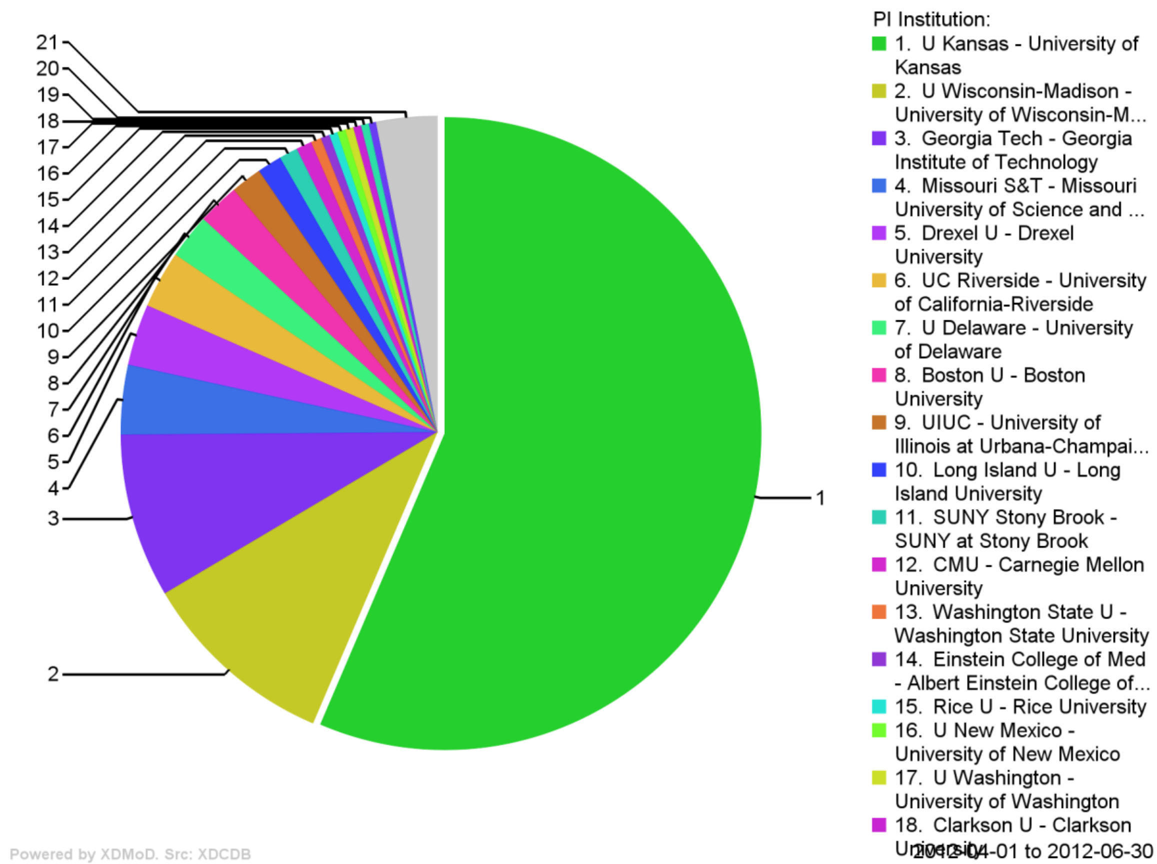
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = PURDUE-STEELE

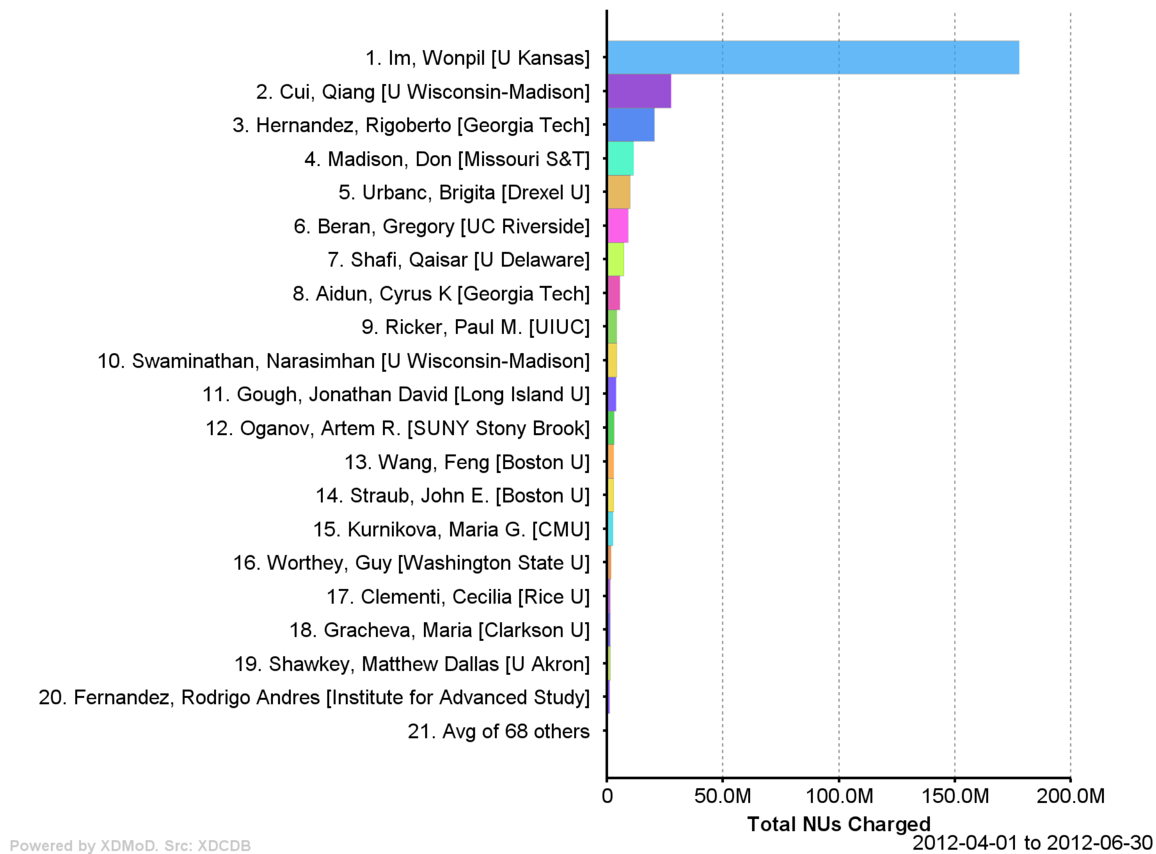
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = PURDUE-STEELE

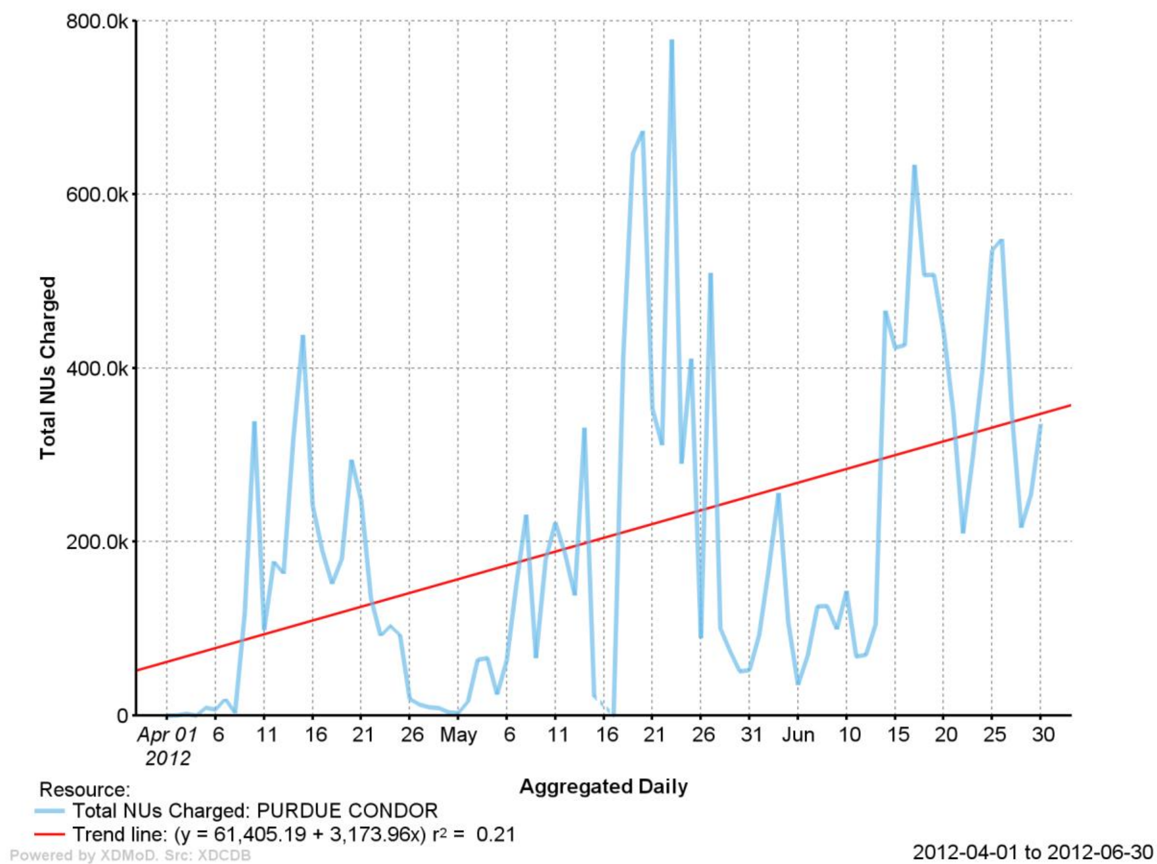
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = PURDUE-CONDOR

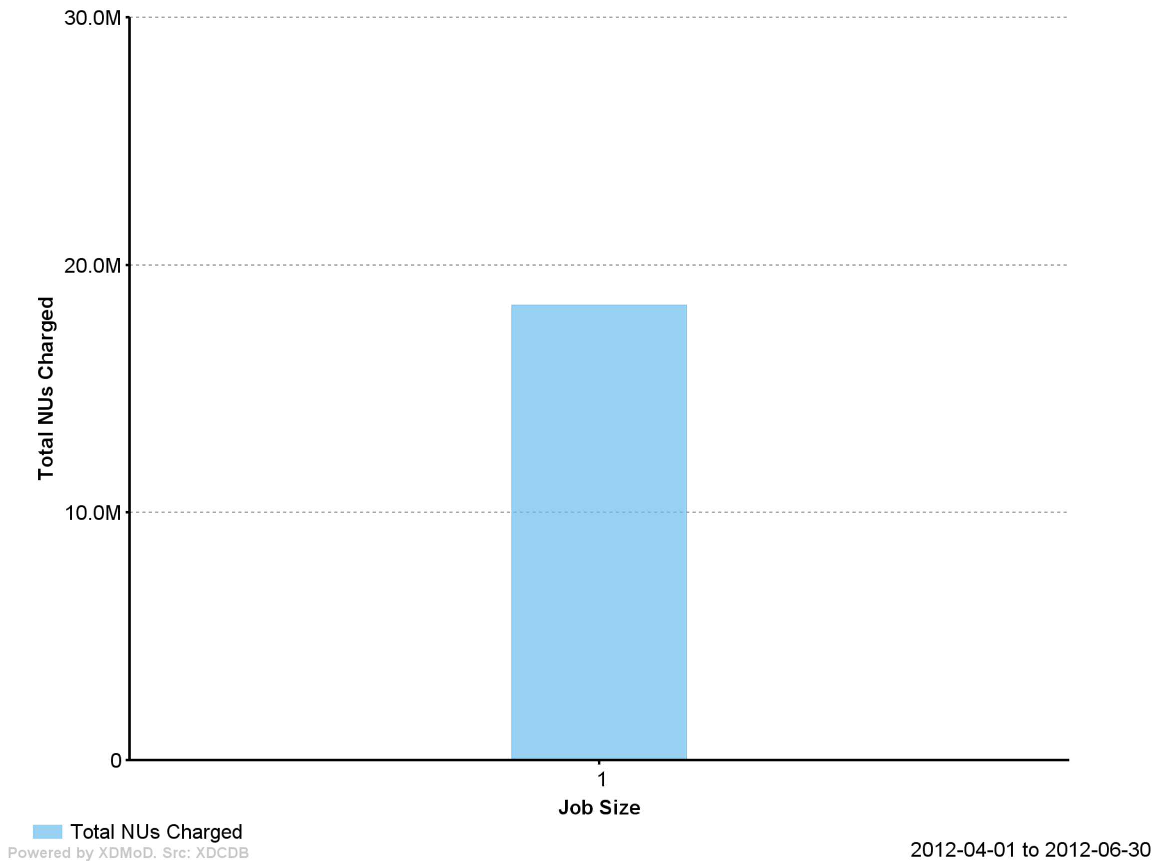
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = PURDUE-CONDOR

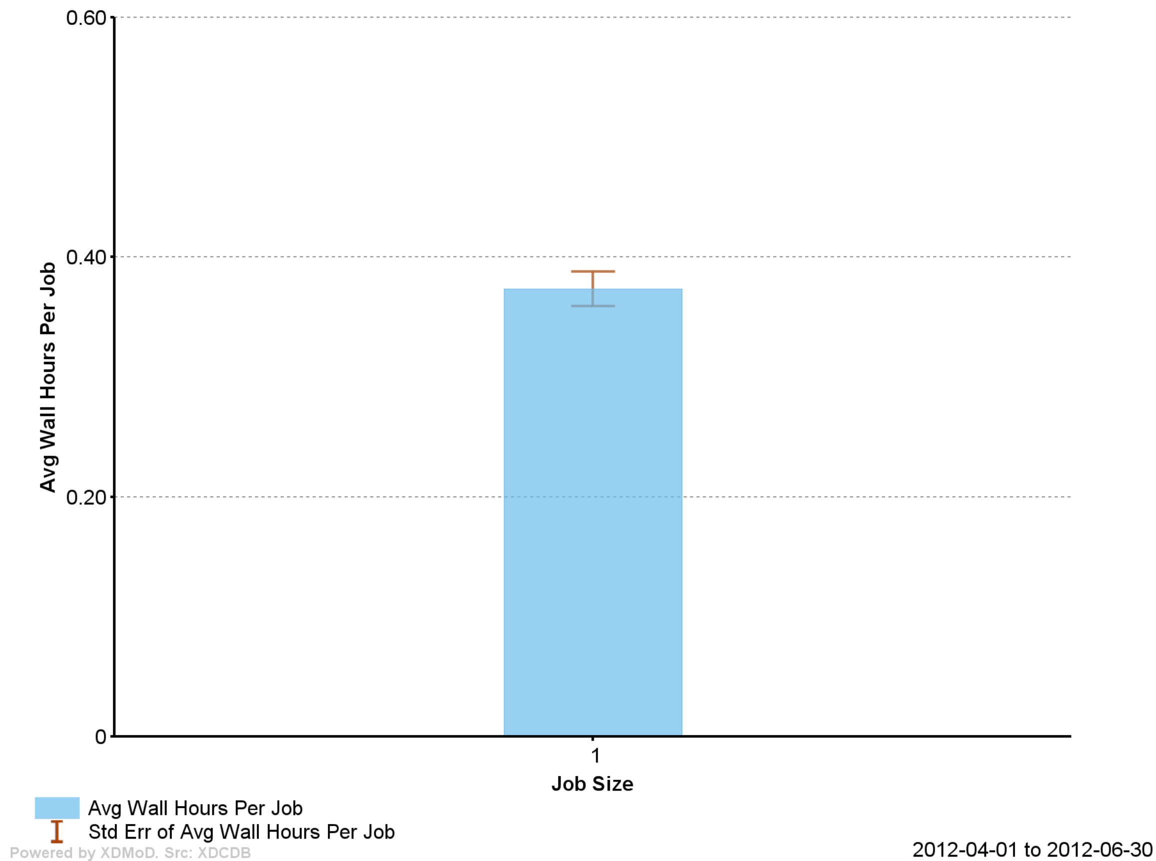
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = PURDUE-CONDOR

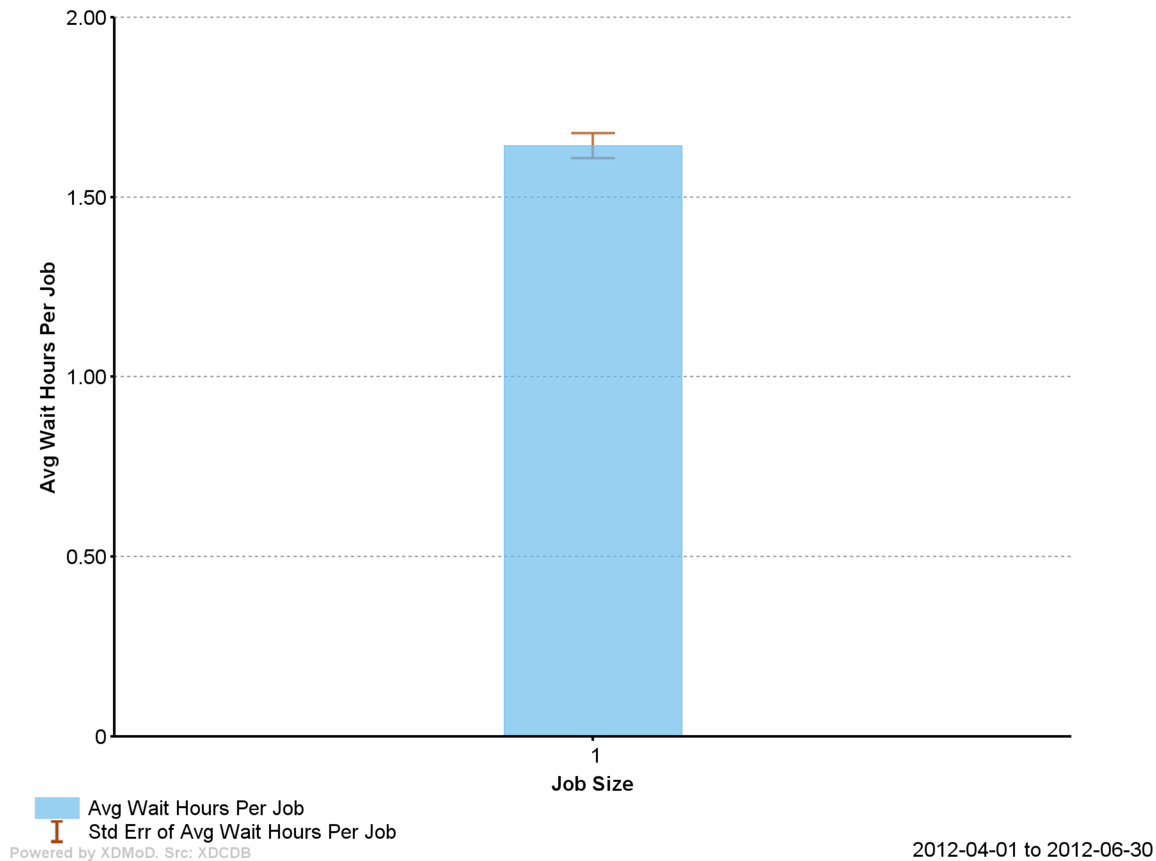
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = PURDUE-CONDOR

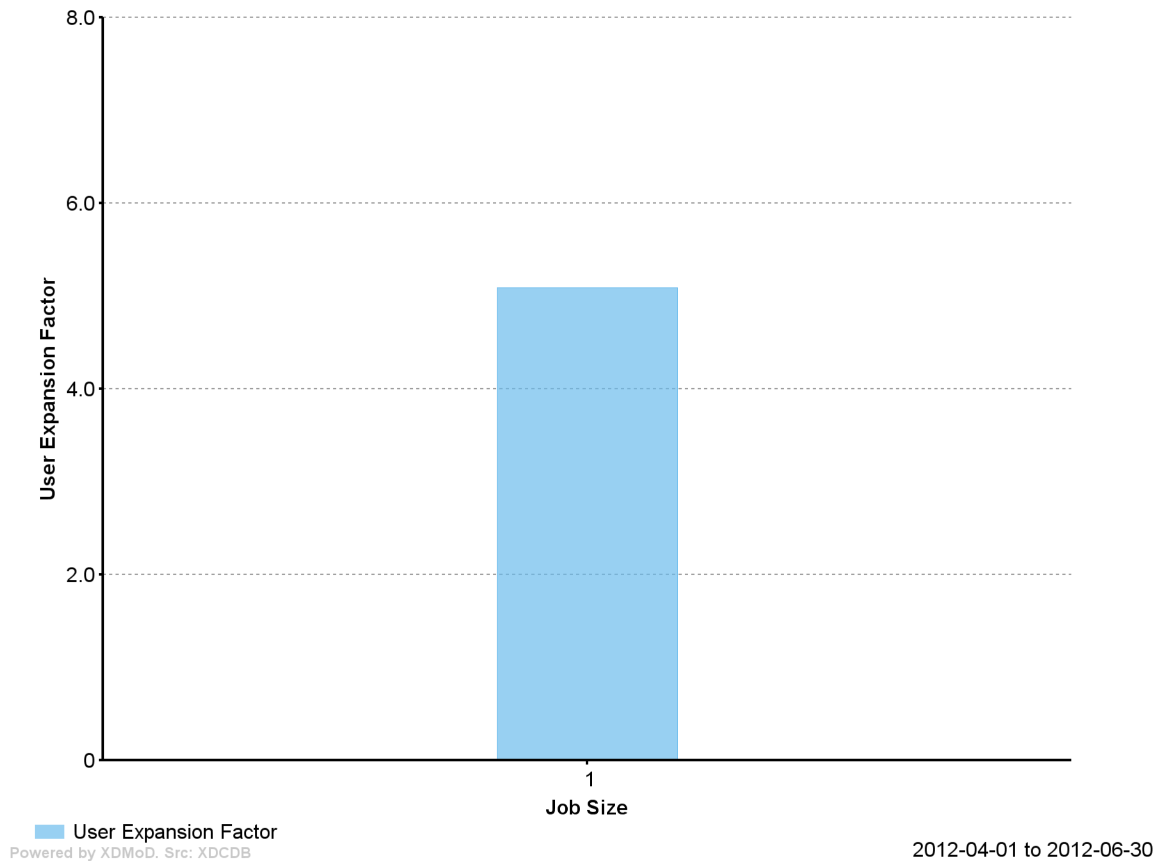
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = PURDUE-CONDOR

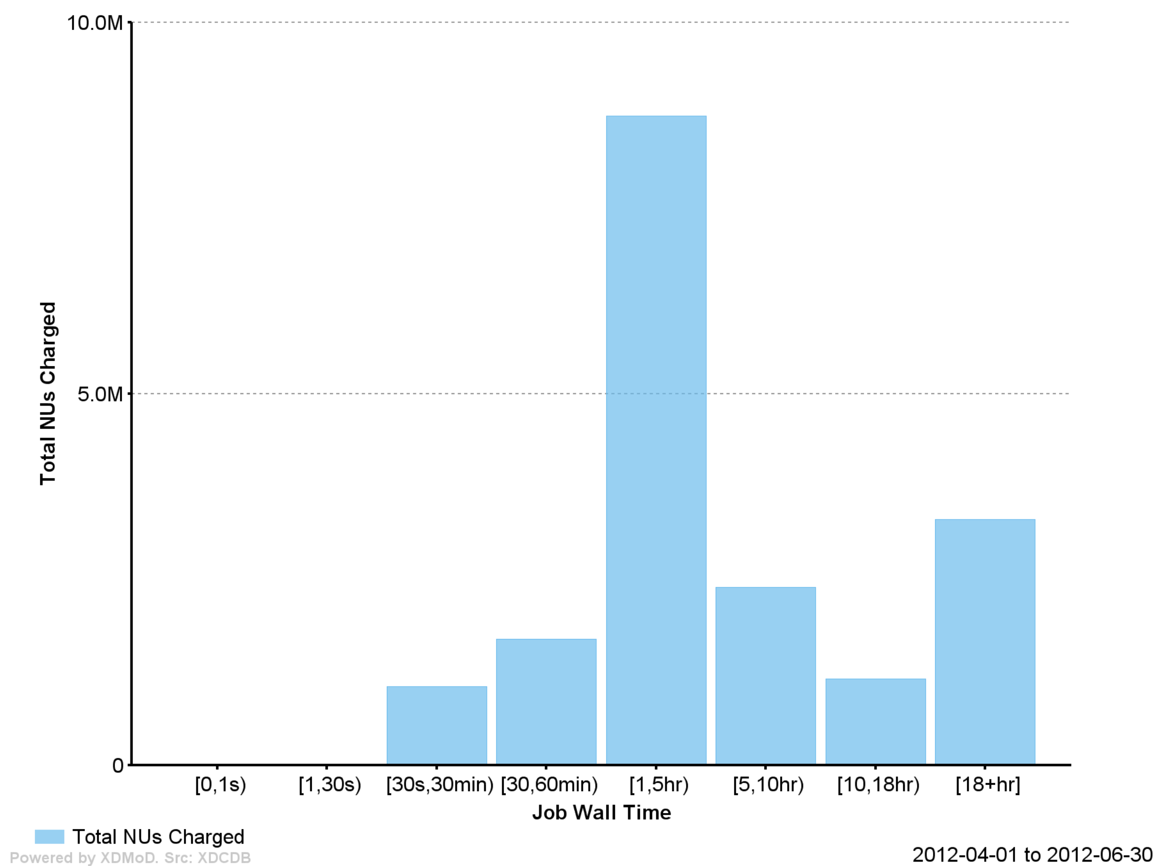
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = PURDUE-CONDOR

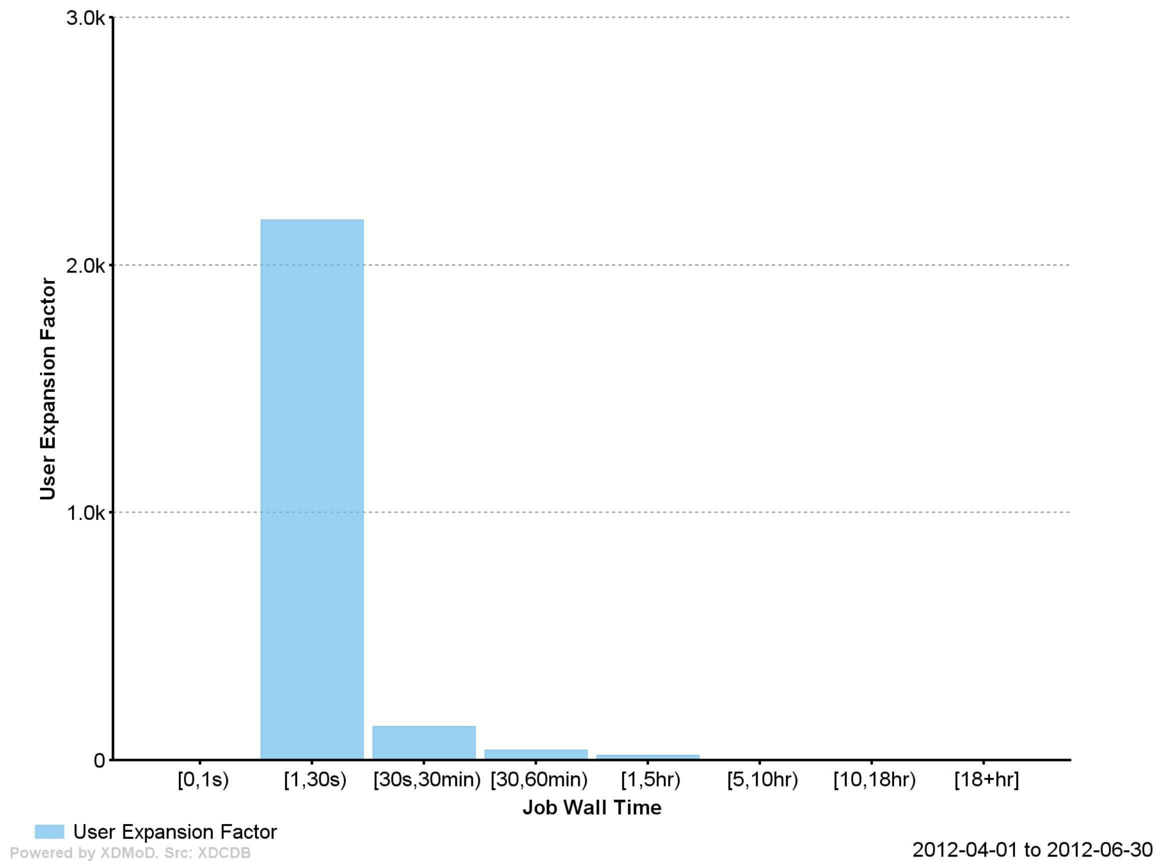
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = PURDUE-CONDOR

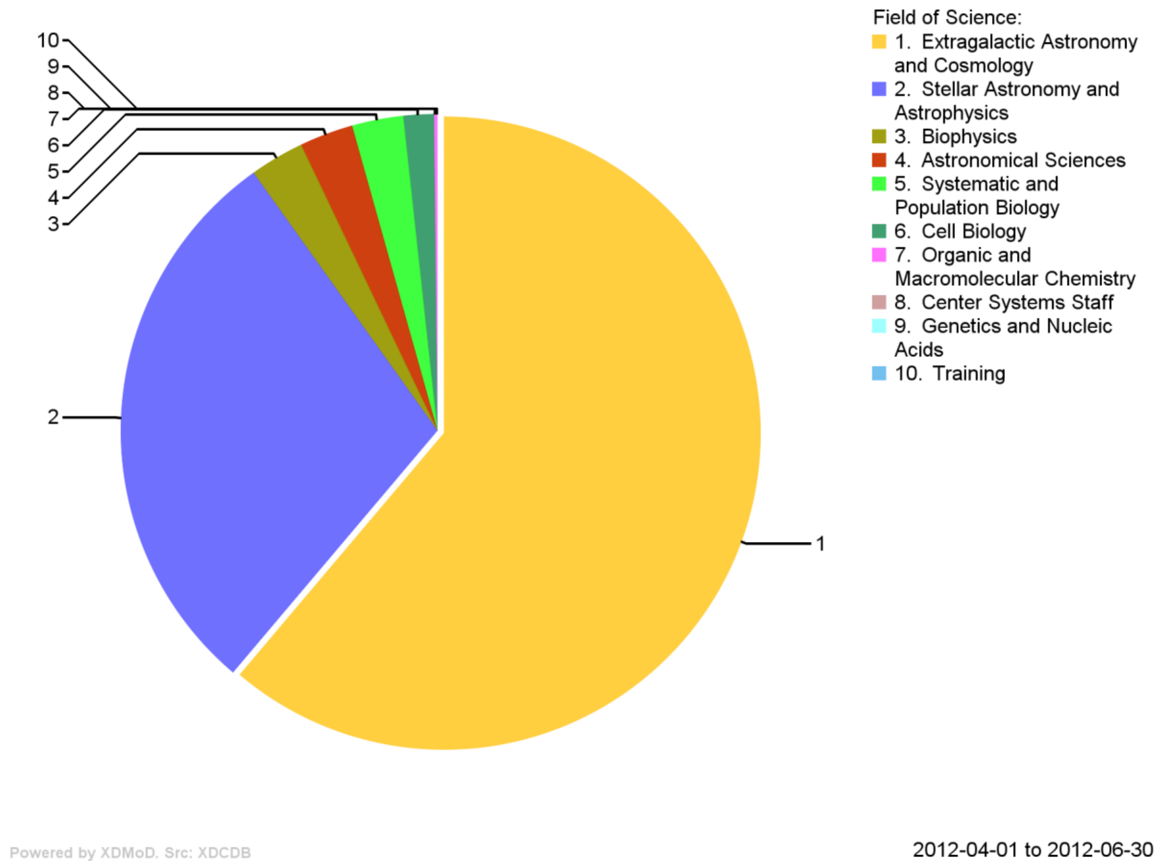
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = PURDUE-CONDOR

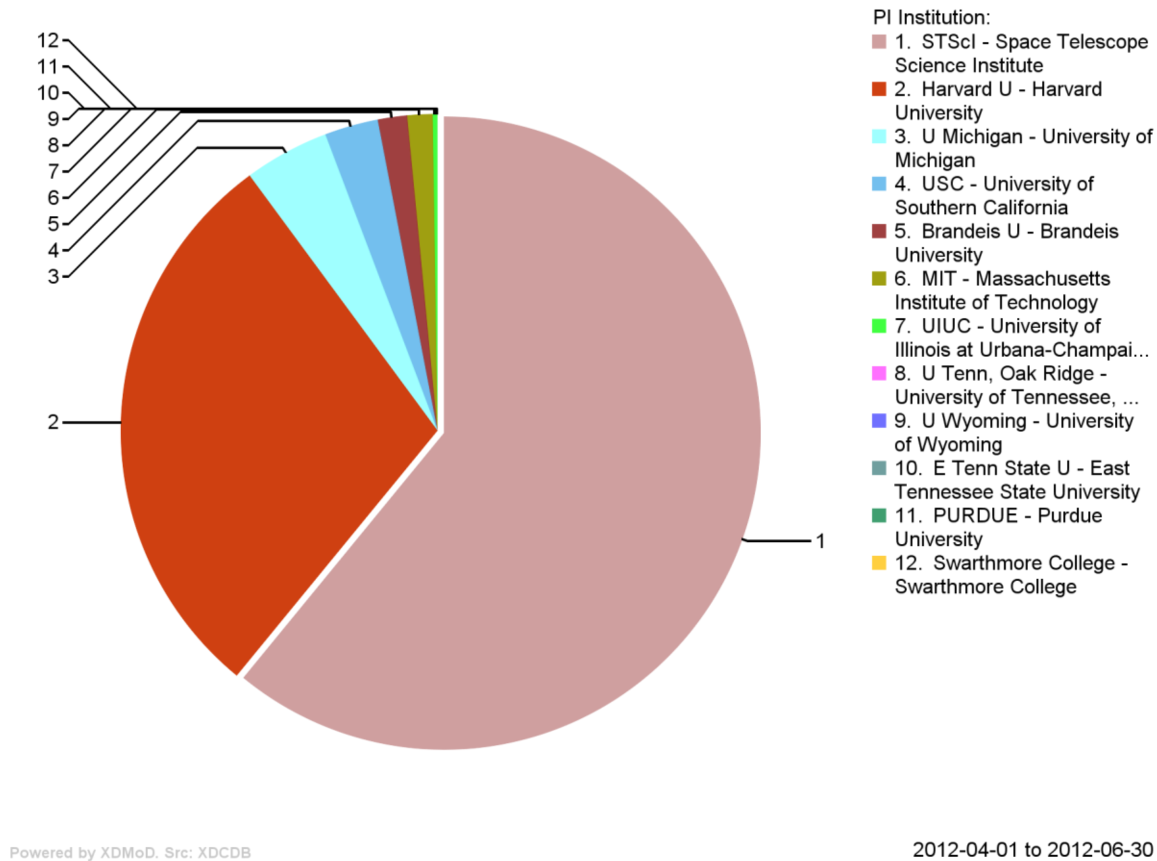
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = PURDUE-CONDOR

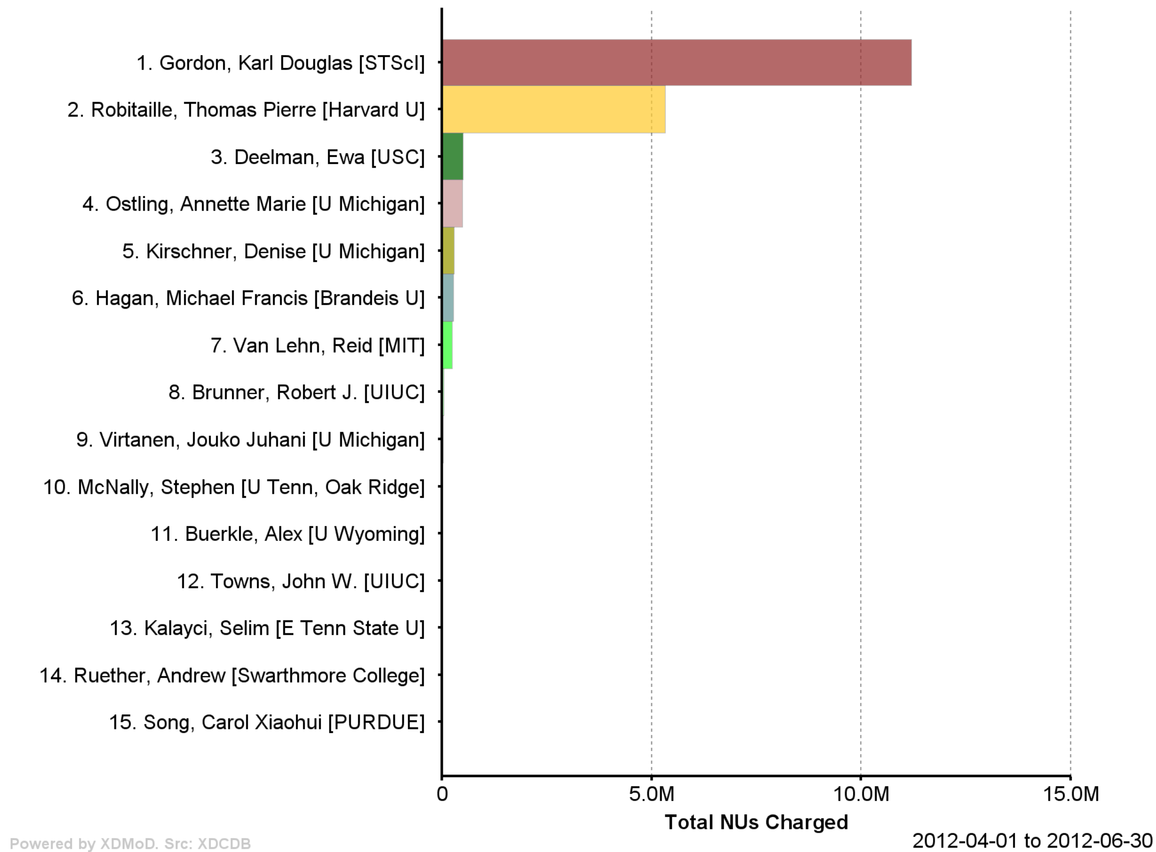
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = PURDUE-CONDOR

2012-04-01 to 2012-06-30



18 San Diego Supercomputer Center (SDSC) Service Provider Quarterly Report (Q2 2012)

18.1 Executive Summary

SDSC successfully deployed the innovative *Gordon* system in early 2012, the first NSF resource dedicated specifically to data-intensive computing, and has continued to effectively serve the gateway and modest-scale user community with the high-productivity *Trestles* cluster.

Gordon is an Appro-integrated cluster that was co-designed by SDSC and Appro to address the challenges of data intensive computing, with a focus on *moving data efficiently through the memory hierarchy*. The beginning of the reporting period focused on risk mitigation and testing, using early Gordon hardware (e.g. I/O nodes) and *Dash*, the risk-reduction system for *Gordon*. The *Gordon* compute nodes were delivered to SDSC in November 2011, shortly after Intel released the Sandy Bridge processors. After passing a rigorous set of acceptance tests and a 20-day reliability test, *Gordon* was officially accepted on February 24, 2012 and has now been added to the XSEDE resource portfolio as the first system designed specifically for data intensive computing. Since entering production status, *Gordon* has been stable and performing very well.

The interest in *Gordon* is high, and we are encouraging allocation requests that make use of its unique architecture features such as flash memory and large memory vSMP "supernodes." In addition, several users have been allocated dedicated use of *Gordon's* I/O nodes; this is a new prototype allocations model for data mining and database projects.

Trestles is now in its second year of production and is successfully supporting the modest-scale/gateway user community, with a focus on user productivity and fast turnaround. The system is well over-requested in the allocation process. The system utilization has recently ramped up from the targeted 70% level to now well over 90%; queue wait times and expansion factors, while still good relative to most XSEDE systems, are increasing beyond desired levels. We are analyzing usage data and user job patterns to optimize utilization while maintaining low expansion factors (results presented at XSEDE'12).

SDSC completed deployment of the *Data Oasis* Lustre parallel file system, which supports *Gordon, Trestles* and SDSC's *Triton,* with 4PB capacity and demonstrated 100 GB/s aggregate bandwidth. *Gordo*n users have access to a ~2PB/50GBps scratch file system, while *Trestles* users access a 400TB/10GBps scratch system. We recently deployed 400 TB of *Data Oasis* "project storage" for allocated storage accessible by both *Gordon* and *Trestles* users; an additional 2 PB of project storage will be deployed next year. This represents a new paradigm of intermediate-term allocated storage, with the hope of significantly reducing long-term archival requirements.

Four new *data mover* nodes were put into production to support high bandwidth transfers between *Data Oasis* and other XSEDE providers. Transfer rates to and from the new nodes using GridFTP have been exceptional, and a large improvement over our previous capabilities. The rates to and from the *Data Oasis* nodes exceeded 1GB/s when transferring to and from NCSA and were over 600 MB/s when moving files between SDSC and NICS.

SDSC provided other XSEDE production resources during the reporting period. The *Dash* system was used effectively as an allocated resource for porting of applications and initial production science in a *Gordon*-like architecture; it was decommissioned as an XSEDE resource after *Gordon* was deployed. The *GPFS-WAN* storage system served as an allocated resource by users and also used by XSEDE for evaluation of GPFS for its future wide-area storage designs; this resource was decommissioned in June 2012 after all its users were assisted in migrating necessary data to other storage resources. Legacy archival data from TeraGrid/XSEDE-allocated users was

retained in the HPSS and SAM-QFS tape archives; all users have been notified and assisted in migrating necessary data, and these systems are being decommissioned in the summer 2012.

SDSC conducted a large number of effective education, outreach and training activities that engaged users, students and new user communities with conferences, summer workshops and education programs.

The resources delivered by SDSC to XSEDE users have resulted in a number of important scientific advances, including:

- The successful CIPRES gateway for analysis of phylogenetic data sets has utilized *Trestles* (and recently *Gordon*) for its entire ~10M SU allocation this past year. This single gateway uses <1% of XSEDE compute resources, but now accounts for 29% of all XSEDE users. In addition to the broad adoption of this gateway, the scientific impact is enormous with gateway users producing nearly 200 scientific publications this past year. Please see "SDSC's CIPRES Science Gateway Clarifies Branches in Evolution's 'Tree of Life'" at http://www.sdsc.edu/News%20Items/PR070312_cipres.html .
- Harvard researchers use *Trestles* and now *Gordon* to perform computations for their Clean Energy Project (CEP), with the goal of creating the next generation of organic solar cells as an inexpensive and efficient source of energy. Please see "SDSC's Trestles Supercomputer Speeds Clean Energy Research" at http://www.sdsc.edu/News%20Items/PR042312_solar.html for the full story.
- Researchers from the University of California campuses at Irvine and San Diego have found a new approach to the creation of customized therapies for virulent flu strains that resist current antiviral drugs. The researchers utilized simulations on *Trestles* to predict how pocket structures on the surface of influenza proteins promoting viral replication can be identified as these proteins evolve, allowing for possible pharmaceutical exploitation. The findings, published in Nature Communications, could aid development of new drugs that exploit so-called flu protein 'pockets.'

18.2 Science Highlights

During the reporting period SDSC completed the deployment of its Lustre-based *Data Oasis* parallel file system, with four petabytes (PB) of capacity and 100 gigabytes per second (GB/s), to handle the data-intensive needs of *Gordon* and *Trestles* as well as SDSC's *Triton* cluster. **Please see SDSC Supercharges its 'Data Oasis' Storage System** at http://www.sdsc.edu/News%20Items/PR060412_dataoasis.html for the full story.

SDSC also announced plans to host a Summer Institute, expanding upon its successful *Gordon* Summer Institute program to include both *Gordon* and *Trestles*. The 'SDSC Summer Institute' will be held August 6-10 at SDSC. Please see **SDSC to Host "Summer Institute" Supercomputer Workshop August 6-10** at http://www.sdsc.edu/News%20Items/PR052412_summerinstitute.html for the full story.

One research project highlighted during the reporting period involved a team of Harvard University researchers that was allocated time on *Trestles* to perform computational calculations with the goal of creating the next generation of organic solar cells as an inexpensive and efficient source of energy. The allocation is a key part of the team's efforts to conduct larger, data-intensive computations related to its Clean Energy Project (CEP), which combines the group's computational chemistry expertise with the large, distributed computing power of IBM's World Community Grid (WCG). The group was allocated an additional 4 million SUs on *Gordon* and ~658,000 SUs on *Trestles* starting July 1 to further their research. Please see **SDSC's Trestles**

Supercomputer Speeds Clean Energy Research at http://www.sdsc.edu/News%20Items/PR042312_solar.html for the full story.

18.3 User-facing Activities

18.3.1 System Activities

Key systems activities in this reporting period included:

- Continued support and enhancement of the *Gordon* production environment
- Developing a test environment for a planned CentOS 6 upgrade of *Trestles* and *Gordon*
- Deployed a non-purged "projects" file system for *Trestles* and *Gordon* which will become the basis of SDSC's allocable storage
- Deployment of 4 new "data mover" nodes to facilitate high speed transfers to/from SDSC and other sites
- Decommissioning older file systems, including GPFS-WAN, Albedo and the legacy tape archive

Gordon systems support focused primarily on hardening the production environment following *Gordon*'s production launch in February of this year. This included: significant interaction with User Services for user-requested application installations; implementation of systems monitoring scripts; providing environments for *Gordon* dedicated I/O node projects; working with the Networking group on upgrades and bug fixes to the Data Oasis core switch infrastructure; and miscellaneous day-to-day activities related to ensuring the smooth operations of the queuing environment, and job accounting. Additionally, the Systems group has been working with the Applications and User Services staff to look at additional options for deploying flash based on user requests which is tied to the system's strategic goals of supporting novel applications and usage models.

The single biggest impact to production operations was a campus wide power outage in May that was caused by a transformer fire in a campus power substation. This brought down the entire campus, including SDSC's data center. While both *Trestles* and *Gordon* came up quickly following the outage, there was a severe issue with the Lustre metadata server for the *Gordon* scratch file system that required a file system rebuild. This delayed making *Gordon* available to users by an additional two days as Aeon Computing, the hardware vendor for Data Oasis, came onsite to perform a very cautious and complex recovery of the server. No data was lost as a result of the power outage. The outage also resulted in the loss of one of *Gordon*'s InfiniBand switches. An onsite spare was put into service and the disruption to users was minor. Lessons-learned follow-up meeting were conducted and new processes monitoring were put in place to better respond to events of this magnitude.

The *Trestles* management rack experienced several power failures due to a faulty circuit in the Data Center, *Trestles* was removed from the circuit in question so that the circuit and its UPS could be repaired.

Two new Lustre file systems in *Data Oasis*, internally code-named Puma and Meerkat, were put into production this quarter. Puma provides scratch space for *Trestles* and Meerkat is cross-mounted project storage for *Gordon* and *Trestles* NSF users at SDSC. Puma entered production in April to replace "phase1", and is part of SDSC's *Triton* cluster. Meerkat became available to users in May. June marked the planned end of GPFS-WAN wide-area file system. Also in June, after Oracle declined to extend the warranty, the decision was made to migrate the data from SDSC's portion of the Albedo file system to other XSEDE Service Providers and decommission that hardware.

In June four new data mover nodes were put into production to support high bandwidth transfers between *Data Oasis* and other XSEDE providers. Transfer rates to and from the new nodes using GridFTP have been exceptional, and a large improvement over our current capabilities. The rates to and from the *Data Oasis* nodes exceeded 600 MB/s when moving files between SDSC and NICS, and were over 1GB/s when transferring to and from NCSA.

Finally, SDSC streamlined its local accounting system by automating AMIE packet generation. When a refund or credit is granted to a national allocation at SDSC, the associated AMIE packet is now automatically created and sent to the XSCDB. Previously this was a two-step manual process. Work has begun on developing a method for tracking and reporting storage allocations with the goal of having a viable mechanism in place by the end of 2012.

18.3.2 Services Activities

Between the two tickets systems used to support *Trestles* and *Gordon* (NCSA's XSEDE ticket system and SDSC's local ticket system) a total of 806 tickets were created between April 1 and June 30, 2012. These tickets included account questions, file system issues, software requests, Globus support, code support, password resets, code optimizations and debugging, allocation refunds/problems, project space requests, software support, licensing queries, and resource availability. 758 of those tickets were closed, leaving 48 tickets that we are still working to resolve. The average time to close the tickets across the two systems was 8.1 days, with a median time of 2 days. The XSEDE ticket system statistics are provided as an Appendix to this report.

Gordon is now in its second quarter of operation and supporting users from 3 XRAC allocation cycles. Typical *Gordon* support included fielding questions and providing information on Gordon's system architecture, updating documentation, and software install requests. Installation support was provided for GROMACS, IDL, VisIt, R, and NEMO software on *Gordon*. Additionally, installation and testing of performance tools (PAPI, TAU, IPM), a newer version of AMBER (v.12), and Hadoop is in progress. Large memory and I/O cases for Abaqus and *Gordon* were tested and benchmarked using the vSMP node. This involved memory usage of ~750GB and ~800GB of I/O to the aggregated SSD files ystem on the vSMP node. Both these packages are now available in production on *Gordon* and being used by allocated users. User Services staff members were also involved in testing and performance evaluation of gridftp services to the *Gordon's* scratch file system using the new data mover nodes mentioned above.

In support of *Trestles* users, User Services staff worked to help users transition from the retired /phase1 filesystem to the two new *Data Oasis* Lustre file systems (/oasis/scratch and /oasis/projects). New software installs included the FASTQ tools, OSC mpiexec launcher, NRMOL, Weka, and GUILE. User Services staff also provided support for bundling IDL program runs on *Trestles*. They also developed and tested dynamic scripts (modified at run time) to enable use of SSD scratch on multiple nodes simultaneously. This was a use case which enabled bundling of multi-node R software runs with the use of SSDs for I/O.

18.4 Security

No security incidents occurred during this reporting period.

The security assessment of the software upgrade to *Trestles*, mentioned in last quarter's report, has been rescheduled until next quarter. In its stead, we focused on the configuration and deployment of the four new Data Oasis data-mover nodes, as well as the maintenance of the grid-authentication PKI used by those nodes.

In the coming quarter, we expect to complete the aforementioned security assessment and port any necessary security-related software, as well as the authentication software used on *Gordon*, to the new platform.

18.5 Education, Outreach, and Training Activities

Types: Workshop, Seminar, Tutorial, Conference talk/presentation/panel, synchronous course, online course, web-based training, campus visit (e.g. CI Days, Campus Champions, etc.)

Audiences: Faculty (F); Graduate students (G), Undergraduates (U), Post-Docs (D), Industry (I), K-12 Teachers (T), K-12 Students (S), Public (P)

Type	Audiences	Title	Location	Date(s)	Hrs	# Participants	# Under-represented	Method
Tutorial	F,D, G,U	Introduction to Gordon and XSEDE's Big Data Computing Services	Santa Barbara, CA	April, 2012	4	15-18	n. a.	Synchronous
Tutorial	F,D, G,U	Gordon: A Data-Intensive Supercomputer	Richmond, VA	April, 2012	4	100	n.a.	Synch
Wkshp	F,D, G,U	Gordon 101: Introduction to Data-Intensive Computing	San Diego, CA	May, 2012	4	51		Synch.
Tutorial	F,D, G,U	Introduction to "R" programming	San Diego, CA	May, 2012	4	24		Synch.
On-line	F,D, G,U,I ,P	Intel Chip Chat Episode 188: Supercomputing with the Intel Xeon Processor E5 Family	http://soundcloud.com/intelchipchat/supercomputing-and-xeon-e5	May 2, 2012	1			Async.
Wkshp	F,D, G,U	Data Intensive Computing Introduction	Dublin, Ireland	June, 2012	6		n. a.	Synch.
Tutorial	F,D, G,U	Computational Neuroscience	UCSD	June 2012	30	20	n.a.	Synch
Conf.	F,D, G,U	SDSC's Next Generation Cyberinfrastructure for Big Data Applications;	San Diego, CA	March 5-8, 2012	1	350	n. a.	Synch.
Conf	F,D, G,U	Supercomputing Goes Data-Intensive	Omaha, NE	June, 2012	1	400	n. a.	Synch..
Conf	F,D, G,U	Implementations of Urgent Computing on Production HPC Systems	Omaha, NE	June, 2012	1	400	n. a.	Synch.

Seminar	F,D, G,U	Gordon and other HPC Resources for the Research Community	San Diego, CA	June, 2012	2		n. a.	Synch..
Keynote Pres.	F,D, G,I	Gordon – Design and Performance of a 3# Torus Interconnect for Data-Intensive Computing	Hamburg, Germany	June 2012	1	100	n. a.	Synch..
Invited Spkr; Conf	F,D, G,I	Gordon - Design, Performance, & Experiences Deploying & Supporting a Data Intensive Supercomputer	Hamburg, Germany	June 2012	1	200	n. a.	Synch..
Event	T	TeacherTECH Science Olympiad Coaching	SDSC	April 1	8	22	3	Sync
Conf	S	Girl Scout Imagine Conference	SDSC	April 13	8	18	18	Sync
Workshop	T	A Geology Toolkit for Educators	SDSC/ Torrey Pines	April 14	4	18	14	Sync
Event	S, T, U, F, G	16th Annual ASBMB Undergraduate Student Poster Competition for SMART Teams	La Jolla, CA	April 20	4	25	11	Sync
Workshop	T	ASBMB Teacher Workshop Presentation	San Diego, CA	April 21	4	50	33	Sync
Course	T	Introduction to CS Principles through Alice for Educators	UCSD	April – May	20	12	5	Sync
Course	S	Learn MIT's Scratch! An Intermediate Class	SDSC	April 28	6	21	4	Sync
Workshop	T	Project Astro: Galileo Educator Workshop- Grades 4-12	SDSC	May 19	6	18	11	Sync
Course	S	To Infinity and Beyond: Physics of Space Flight for Grades 6-9	SD Air & Space Museum	June 2	6	45	22	Sync
Workshop	T	SMART Summer Teacher Meeting	SDSC	June 25-29	20	18	4	Sync
Course	S	Alice: Beginning Programming in a 3D	UCSD	June 25-29	30	22	6	Sync

		Environment						
Workshop	T	Teaching CS Principles through Alice: Pedagogy	San Diego, CA	June 6-8; 27-29	18	12	5	Sync
Course	S	Introduction to Object Oriented Programming Using Java	SDSC	June 25-29	30	21	2	Sync

18.5.1 Training

Training activities during the second quarter of 2012 reflected the strong community interest in SDSC's newest XSEDE resources. In April, Mahidhar Tatineni introduced researchers at UC Santa Barbara to XSEDE's newest "Big Data" services, and in Richmond, VA, SDSC staff introduced attendees at the 45^{th} Annual HPC User Forum to "Gordon: A Data-Intensive Supercomputer." May included two types of training. The first was individual extended support for data-intensive computing for a non-traditional user in a visit to USC to meet with Virginia Kuhn regarding digital video data applications. The second training format was the more traditional workshop for big-data researchers, but with a topic twist. The first workshop, "Gordon 101," introduced users to XSEDE's new data-intensive supercomputer and how to get started using it. An "Introduction to R Programming Language" workshop presented by a faculty member from UCSD's Political Science department opened the door to potential new users from that social science domain. In June, Bob Sinkovits taught *Data Intensive Computing Introduction* at the European-US Summer School on HPC Challenges in Computational Sciences in Dublin, Ireland. A *Computational Neuroscience* tutorial (June, 23-27, 2012) at UCSD drew twenty (20) attendees (grad students, postdocs, scientists and faculties) from various US universities, NIH and Europe. Tutorial participants learned how to use the widely used neuronal simulation software NEURON (http://www.neuron.yale.edu/neuron/). The presenters included Ted Carnevale, Michael Hines (both from the Neurobiology Dept, Yale) and Terry Sejnowski (SALK, UCSD). As a part of this workshop the NEURON software was installed on both *Gordon* and *Trestles* and the attendees performed hands-on exercises, using NEURON, on the *Gordon* machine. Additional SDSC presentations this quarter spanned the globe, and are covered below under "Outreach."

18.5.2 Education

SDSC's NSF-funded Computing Education for a 21st Century Workforce (CE 21) project, entitled "Computing Principles for ALL Students' Success" (ComPASS) supports teacher professional development to implement a new computer science principles course at the college and pre-college levels. The course is being introduced to pave the way for a new AP (Advanced Placement) test in computer science; one that will focus on the foundation principles of computational thinking. During the second quarter of 2012, SDSC hosted three professional development workshops introducing computer science principles to high school teachers. One workshop focused on content of a new college freshman-level curriculum, with an eye to introducing it into the pre-college classrooms. Two subsequent workshops introduced pedagogical approaches to stimulate computational thinking practices and support students' ability to communicate around computing.

Late in June, SDSC received notice of an REU award supplement for *Gordon*. Due to the late date of notification, students were recruited to start working in September 2012 or June 2013.

SDSC hosted 17 workshops, courses, and events for students, K-14 educators, and the general public. The programs varied from a two-hour class to multiple-meeting workshops, to an all-day family activity for 50,000 event attendees. Participant demographics are in the table below.

Education Program Audience Demographic Breakdown	***Female***	***Male***	***Teachers***	***Students***
	127	***150***	***125***	***152***

The SDSC Research Experience for High School Students (REHS) program grew to 32 interns this year. Two of the 2012 REHS students applied to attend and present posters at XSEDE 12.

18.5.3 Outreach

XSEDE TEOS Blog: SDSC's Ange Mason strives to keep the XSEDE TEOS pages' content fresh and timely. During this quarter, she posted 54 articles to the XSEDE TEOS blog. Posts to the blog consisted of a combination of single news posts, the XSEDE weekly Newsroom posts and the monthly XSEDE Education Blog Spot post. The "Monthly Education Blog Spot" is being retired as the Newsroom's weekly posts will include that content.

SDSC's Ange Mason began working with the website redesign committee to revise and update the HPC University site. She will be adding content to the HPCU calendar regularly with a new "HPC News You Can Use" weekly newsletter.

SDSC staff extended outreach to national and international colleagues through professional conference presentations and tutorials related to XSEDE Data-Intensive computational services:

XGN Congress and Expo Meeting; San Diego, CA *SDSC's Next Generation Cyberinfrastructure for Big Data Applications*; Bob Sinkovits, presenter	March, 2012
International Conference on Computational Science; Omaha, NE *Supercomputing Goes Data-Intensive*; Michael Norman, presenter	June 2012
International Conference on Computational Science; Omaha, NE *Implementations of Urgent Computing on Production HPC Systems*; Kenneth Yoshimoto, presenter	June 2012
UC Cloud/Grid/HPC group; San Diego, CA *Introducing Gordon and UC's HPC Resources*; Shawn Strande, presenter	June 2012
HPC Advisory Council meeting; Hamburg, Germany Keynote presentation: *Gordon - Design and Performance of a 3D Torus Interconnect for Data Intensive Computing*; Shawn Strande, presenter	June 2012
International Supercomputing Conference 2012; Hamburg, Germany Invited Presentation: Gordon - *Design, Performance, & Experiences Deploying & Supporting a Data Intensive Supercomputer*; Shawn Strande, Presenter	June 2012
European-US Summer School on HPC Challenges in Computational Sciences; Dublin, Ireland *Data Intensive Computing Introduction*; Instructor Bob Sinkovits	June 2012

18.6 SP Collaborations

18.6.1 Collaborations with SP XSEDE Users

- Wayne Pfeiffer used RAxML and MrBayes to do expedited phylogenetic analyses of large insect data sets for Duane McKenna (University of Memphis and Harvard University Museum of Comparative Zoology). The RAxML analysis was done on *Trestles* for a data set with 14 taxa and 6M DNA characters, which was too large to submit via the CIPRES gateway. The MrBayes analysis was done on *Gordon* for a smaller data set with 14 taxa and 389k characters, because MrBayes would not work with the larger data set. *Gordon* was used for the latter analysis because of the long queue-wait times on *Trestles*. The phylogenetic trees obtained by the two analyses were the same.
- Wayne Pfeiffer expedited a problematic RAxML analysis being done on the CIPRES gateway for Sarah Mathews (Harvard). After more than a week of execution on 60 cores of *Trestles*, her analysis was not progressing. Pfeiffer found that the parameter optimization of the specified amino acid model was converging very slowly and might, in fact, never converge. He contacted Alexis Stamatakis, the developer of the code, who suggested two work-arounds. A variation of one of those was adopted and allowed the analysis to run in 36 hours using 160 cores of *Trestles*.
- Robert Sinkovits worked with Ramon Huerta (UCSD BioCircuits Laboratory) to optimize the performance of a time-series classifications algorithm based on autoregressive kernels. Ported to *Gordon* and reduced the wall clock time by factor of 100x. This was achieved through a combination of linking threaded libraries, compiler options, eliminating of redundant calculations and use of OpenMP directives.
- Pietro Cicotti started working with Vahe Peroomian (UCLA IGPP) to optimize performance of 3D magneto-hydrodynamics application.
- Pietro Cicotti initiated collaboration with David Lake and Danielle Jung (UCSD Political Science) and has ported STATA applications to *Gordon*.
- Pietro Cicotti and Robert Sinkovits met with Virginia Kuhn (USC Cinema) to determine requirements for *Gordon* dedicated I/O node project.
- Robert Sinkovits, in collaboration with Mao Ye (U. Illinois), improved the performance of existing threaded parallel application for construction of limit order books by a factor of 70x. This was accomplished through elimination of redundant calculations, removal or critical sections within parallel regions, OpenMP parallelization of serial loops, and identification of opportunities to terminate calculations early. A joint proposal with SDSC and UIUC was submitted to the NSF to continue this work.
- The *Gordon* team met with Kelly Frazer (UCSD Moores Cancer Center) to discuss running computationally intensive genomics analyses on *Gordon*, *Trestles* and *Triton*. Two members of the Frazer lab are planning on attending the SDSC Summer Institute
- David O'Neal (PSC) joined *Gordon* applications meetings to develop expertise that can be applied to ECSS projects.
- Mahidhar Tatineni worked with Matt Goff (Cornell) to perform large Abaqus simulations of the response of bone to stress. These simulations could not be run successfully on any other available resources. Using a single *Gordon* compute node and 700 GB of flash storage, the calculation could be completed in 50 hours, while running completely in core reduced the run time to 10 hours. Additional benchmarks are planned using an upgraded Abaqus license that is valid for larger core counts.
- *Gordon* team met with Weizhong Li to arrange for the use of director's discretionary time on *Gordon* for the metagenomic analysis of Larry Smarr's gut biome.

- *Gordon* team discussed use of *Gordon* for very large sorting problems with George Porter (UCSD).
- Shawn Strande worked with John Helly (UCSD/Center for Multiscale Modeling of Atmospheric Processes (CMMAP)) on an R analysis of CMMAP cloud simulation data. This was a 160GB data set that was analyzed using R in a vSMP node. Additional work is underway to explore the multicore capabilities of R within vSMP.
- Dongju Choi, Amit Majumdar, Mahidhar Tatineni and Amit Chourasia worked with Darcy Ogden to improve the performance and stability of the solid-fluid interaction code CFDlib. Amit Chourasia also provided visualization support using VisIt on *Gordon* to gain insight into volcanic eruptions.

18.6.2 Collaborations with External Partners

Visualization Collaborations

Amit Chourasia participated in the following collaborative activities:

- New collaboration with Tannishtha Reya at School of Medicine at UCSD on visualization of stem cells in bone marrow. The visualization results are available at http://visservices.sdsc.edu/projects/stemcells/
- New collaboration with Claudio Silva and Huy Vo on *Gordon* for a project to try visualization using Hadoop.
- Continued collaboration with SCEC on visualization of *w2w1hz80SN* earthquake simulation as well as crated visualization for diagnosis and validation of results for porting of AWP simulation from CPU to GPU architecture. .
- Continued support for John Helly with visualization of Climate simulation.
- Continued collaboration with Darcy Ogden at SIO/UCSD on visualization of volcano simulations. Visualization results are available at the following website http://visservices.sdsc.edu/projects/volcano

Rick Wagner continued his collaboration with Argonne National Laboratory, working with Joseph Insley to visualize cosmology simulations, and preparing videos for the XSEDE12 and SC12 Visualization Showcases.

SDSC-VAO Letter of Collaboration

In June, SDSC signed a Letter of Collaboration with the Virtual Astronomical Observatory (VAO) project with the goal of improving virtual observatory-based access to simulations and share expertise in management of large data collections, including digital libraries.

18.7 SP-Specific Activities

CIPRES Gateway

To lessen the load on *Trestles*, it was decided to move all MrBayes jobs run via the CIPRES gateway from *Trestles* to *Gordon* (part of an approve XRAC reviewed allocation request). To support this, Wayne Pfeiffer installed the hybrid parallel version of Mrbayes 3.1.2 on *Gordon* and

ran benchmarks comparing its performance on *Gordon* and *Trestles*. For data sets that are not partitioned, the speedup from *Trestles* to *Gordon* increases significantly with the number of characters, e.g., from 1.7x for 2.3k characters to 3.3x for 389k characters. For data sets that are partitioned, as is often the case, the speedup is more modest, typically about 1.6x.

Based on the benchmarks, Pfeiffer devised rules for efficiently running MrBayes on *Gordon* via the CIPRES gateway. Jobs with the default parameters of nruns=2 and nchains=4 run on 16 cores of a single Gordon node, whereas previously they ran on 32 cores of a *Trestles* node. Because the speedup going from 16 to 32 cores is typically modest, such jobs should run faster and more cost-effectively on *Gordon*.

To facilitate the move, Jerry Greenberg made MrBayes 3.1.2h available via modules on *Gordon*. Mark Miller and Terri Schwartz subsequently moved all MrBayes jobs submitted via the gateway from *Trestles* to *Gordon*.

Scheduling and Resource Optimization

A new Quality-Of-Service policy allowing large numbers of small jobs by certain users was implemented. This allows higher utilization by those users.

On *Trestles*, the reservations system was used in a new process for separating parallel files system availability from compute node capability. Users may submit jobs that are labeled as not requiring the parallel file system, and those jobs will be run even if the parallel files system is not available.

For *Gordon* VSMP nodes with mixed CPUs (compute and I/O node cores) the VSMP-specific prologue was modified to exclude I/O node cores.

Documentation on topology-aware scheduling on *Gordon* was written. The Torque resource manager was configured to accommodate heterogeneous I/O node flash configurations on *Gordon*: some nodes with no flash; some nodes with the standard proportional amount; and some with all flash from the associated I/O node.

Trestles on-demand access was described in a paper presented at ICCS 2012: Implementations of Urgent Computing on Production HPC Systems K.K. Yoshimoto, D.J. Choi , R.L. Moore , A. Majumdar , E. Hocks.

Gridftp was configured for *Gordon* data mover nodes.

We also had the release of AMBER 12 and AMBERTools 12 in April 2012; this was extensively tested and tuned for release using *Trestles.*

18.8 Publications

A. Chourasia (2012). “Memory Layout Experiments for Common Visualization Tasks” Presented at XSEDE 12 Conference, Chicago, Jul 2012.

A. Chourasia, D. Ogden, D.J. Choi and K. Wohletz. (2012) “Snapshots of a Volcano Eruption Simulation". Presented at Visualization Showcase at XSEDE 12 conference, Chicago, July 2012.

A. Chourasia, J. Zhou, Y. Cui, D.J. Choi and K. B. Olsen (2012)"Role of visualization in porting a seismic simulation from CPU to GPU architecture". Presented at Visualization Showcase at XSEDE 12 conference, Chicago, July 2012.

Y. Bazilevs, A. L. Marsden, F. Lanza di Scalea, A. Majumdar, M. Tatineni, “Towards a Computational Steering Framework for Large-Scale Composite Structures Based on Continually and Dynamically Injected Sensor Data,” Procedia Computer Science, Vol. 9, 2012, Pages 1149-1158.Salomon-Ferrer, R.; Case, D.A.; Walker, R.C., “**An Overview of the Amber Biomolecular Simulation Package**”, Wiley Interdisciplinary Reviews: Computational Molecular Science, **2012**, *in press*.

Park, K.; Goetz, A.W.; Walker, R.C.; Paesani, F., “**Adaptive QM/MM molecular dynamics simulations of aqueous systems.**”, Journal of Chemical Theory and Computation, **2012**, DOI: 10.1021/ct300331f.

Pierce, L.C.T.; Salomon-Ferrer, R.; Augusto F. de Oliveira, C.; McCammon, J.A.; Walker, R.C., “**Routine access to millisecond timescale events with accelerated molecular dynamics.**”, Journal of Chemical Theory and Computation, **2012**, *in press* DOI: 10.1021/ct300284c.

Dickson, C.J.; Rosso, L.; Betz, R.M.; Walker, R.C.; Gould, I.R., “**GAFFlipid: a General Amber Force Field for the accurate molecular dynamics simulation of phospholipids**”, Soft Matter, **2012**, DOI: 10.1039/c2sm26007g, *in press*.

Votapka, L., Ozlem, D., Swift, R.V., Walker, R.C., Amaro, R., “**Variable ligand- and receptor-binding hot spots in key strains of influenza neuraminidase.**”, J Mol Genet Med, **2012**, 6, 293-300.

Case, D.A., Walker, R.C. *et al*. **AMBER 12**, (2012), University of California, San Francisco.

R. L. Moore, L. Carson, A. Ghadersohi, A. Jundt, K. Yoshimoto, W Young, “Optimization of Throughput and Utilization on Trestles.” *XSEDE’12,* July 2012, Chicago, IL, USA. ACM TBD.

K.K. Yoshimoto, D.J. Choi, R.L. Moore, A. Majumdar, E. Hocks, “Implementations of Urgent Computing on Production HPC Systems,” June 2012, International Conference on Computational Science, ICCS 2012.

18.9 Metrics

Appendices 1.9A-D includes the following metrics:

- 1.9-A XSEDE-generated user ticket statistics
- 1.9-B *Trestles* and *Gordon* Quarterly stats from XDMoD (April – June 2012)
- 1.9-C *Trestles* and *Gordon* annual stats from XDMoD (July 2011 – June 2012)
- 1.9-D Local *Trestles* stats related to achieving user productivity objectives

Appendix 18.9A Standard User Assistance Metrics

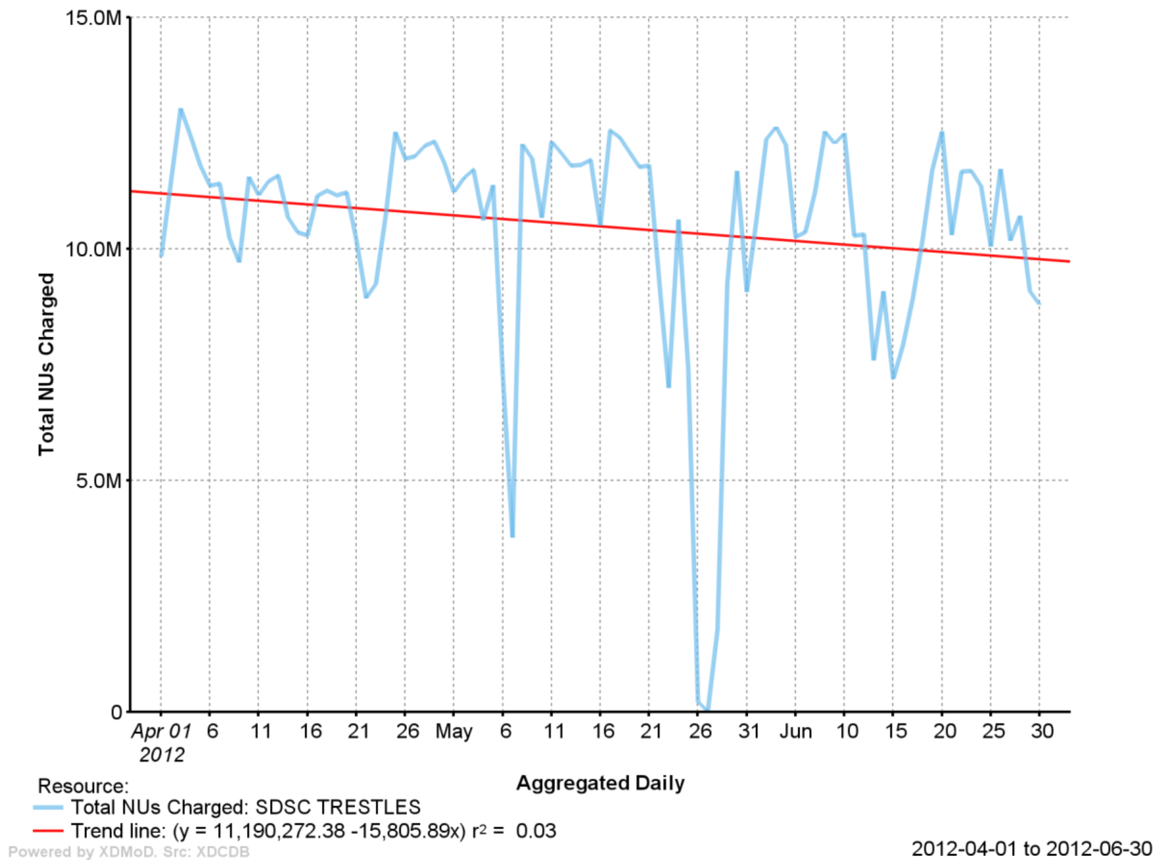
Time to Resolution	account issues	file systems	grid software	jobs/batch queues	login/access issues	mss/data issues	network issues	software/apps	system issues	other
0-1 hr				4	1				1	1
1-24 hr	2	10		33	3	4		11	2	3
1-7 d	7	22	3	46	16	9		21	2	9
1-2 wk	4	7	1	25	5	2		7	1	4
> 2 wk	1	3		33	4	5		20		3
Still Open	2		2	21	2	3		14	1	

Appendix 18.9B Trestles and Gordon Quarterly Usage Statistics

Total NUs Charged by Resource

Resource = SDSC-TRESTLES

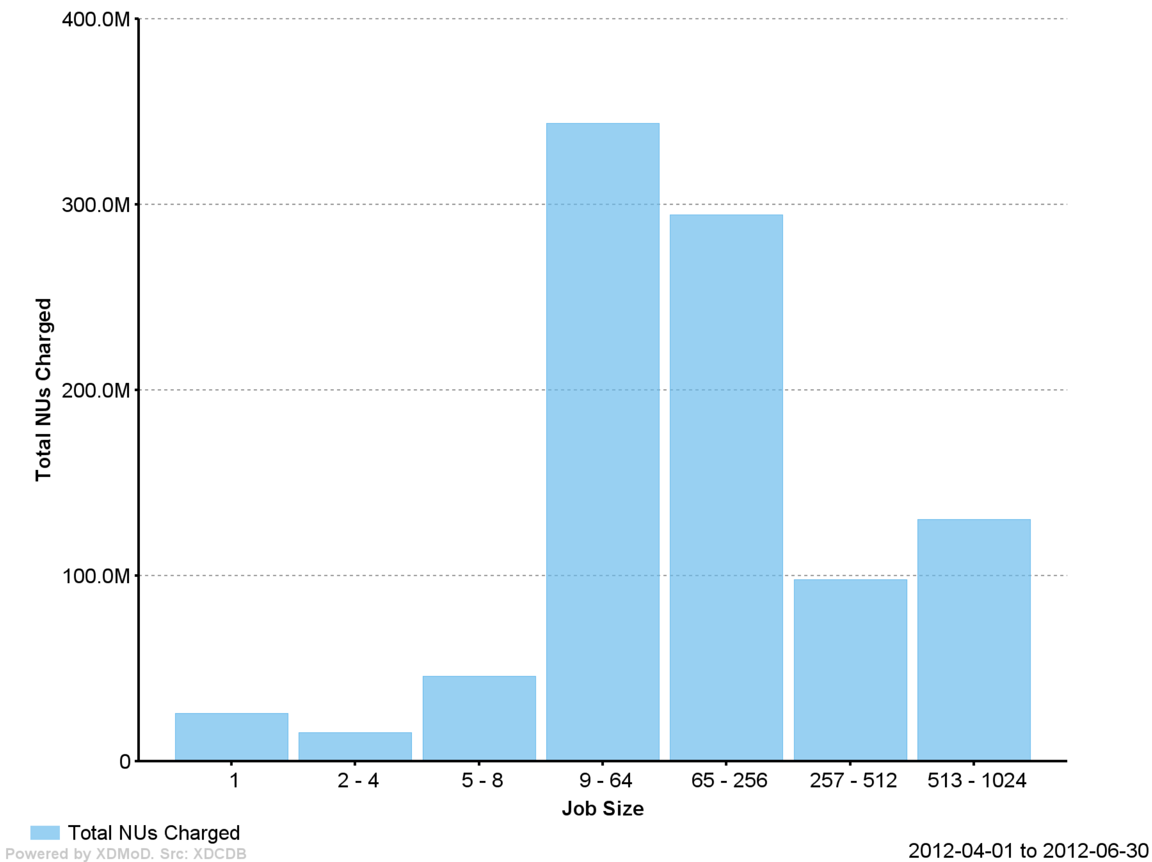
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = SDSC-TRESTLES

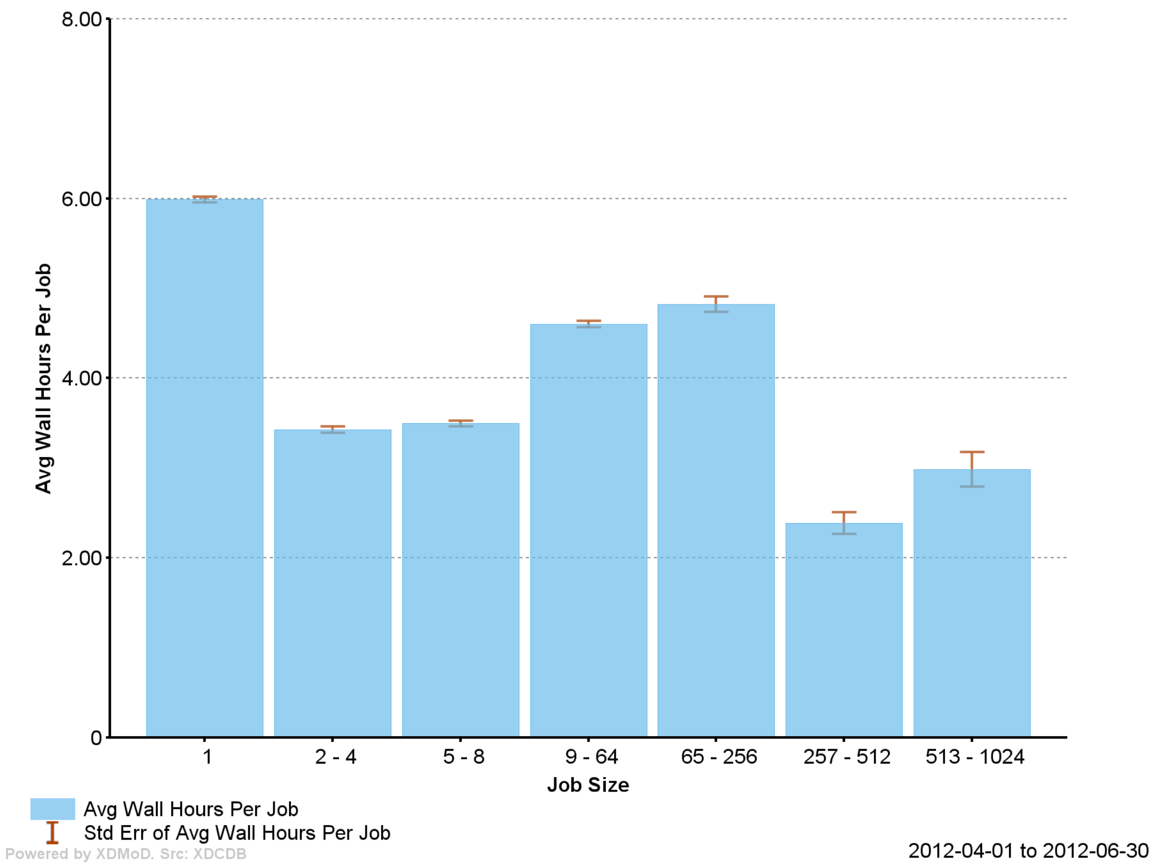
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = SDSC-TRESTLES

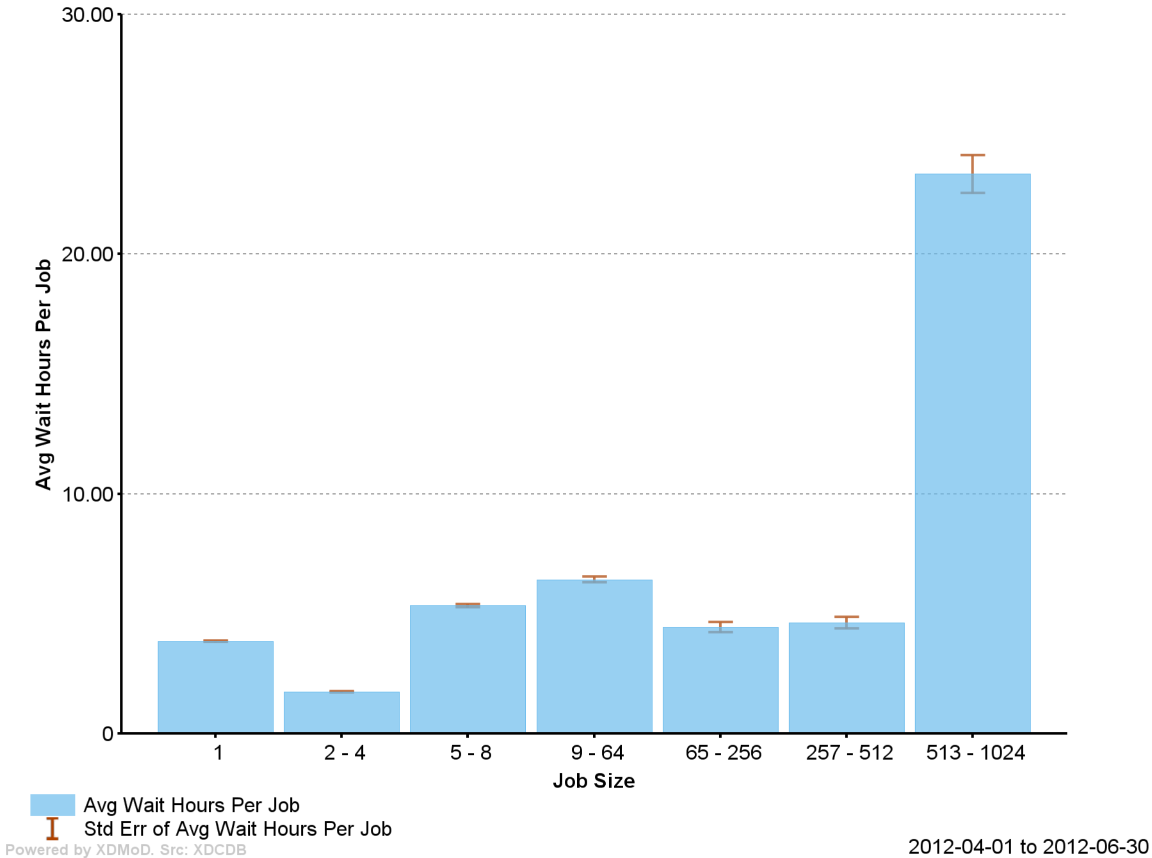
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = SDSC-TRESTLES

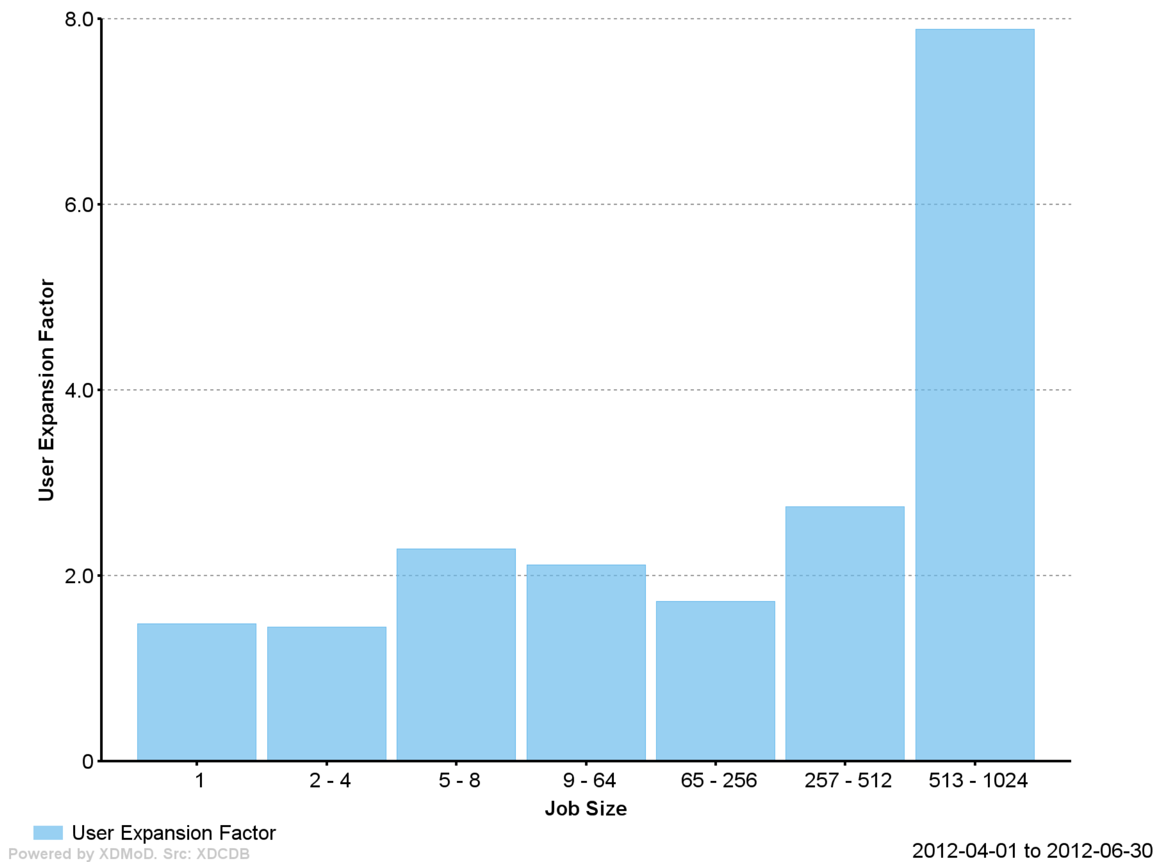
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = SDSC-TRESTLES

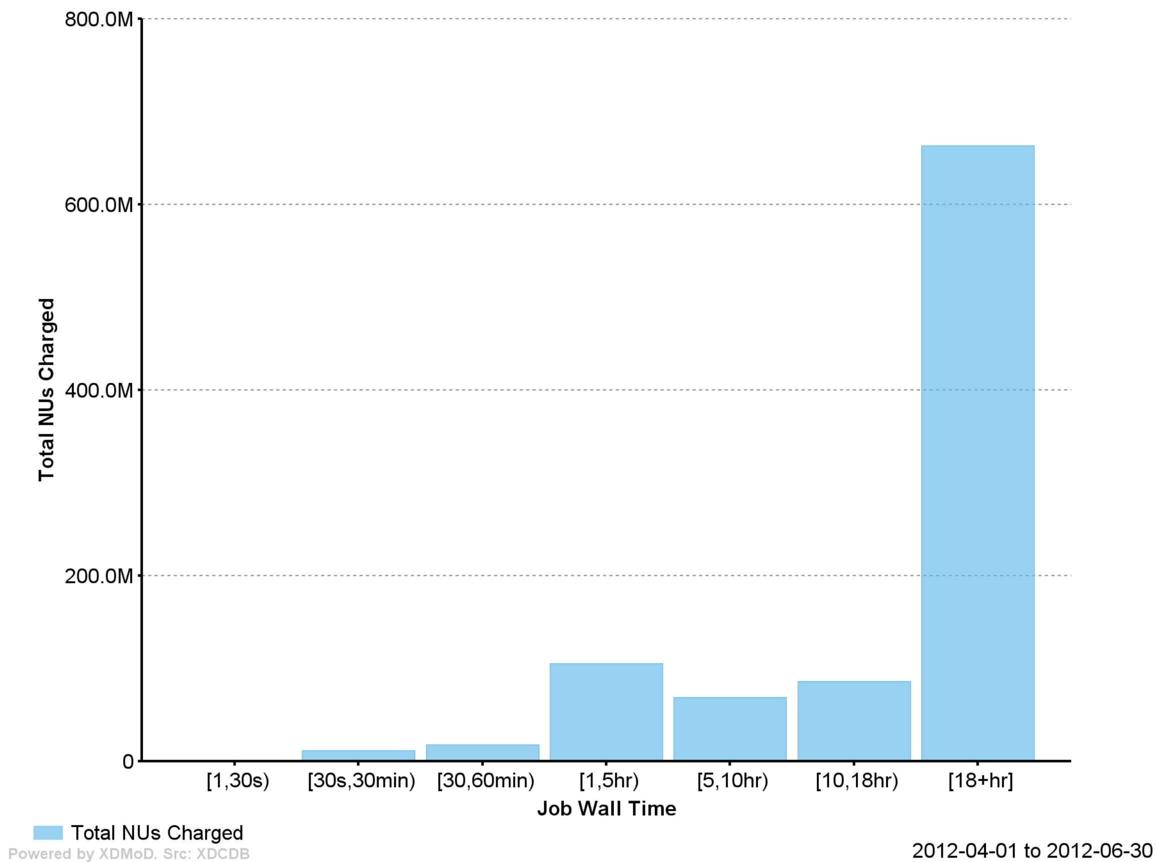
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = SDSC-TRESTLES

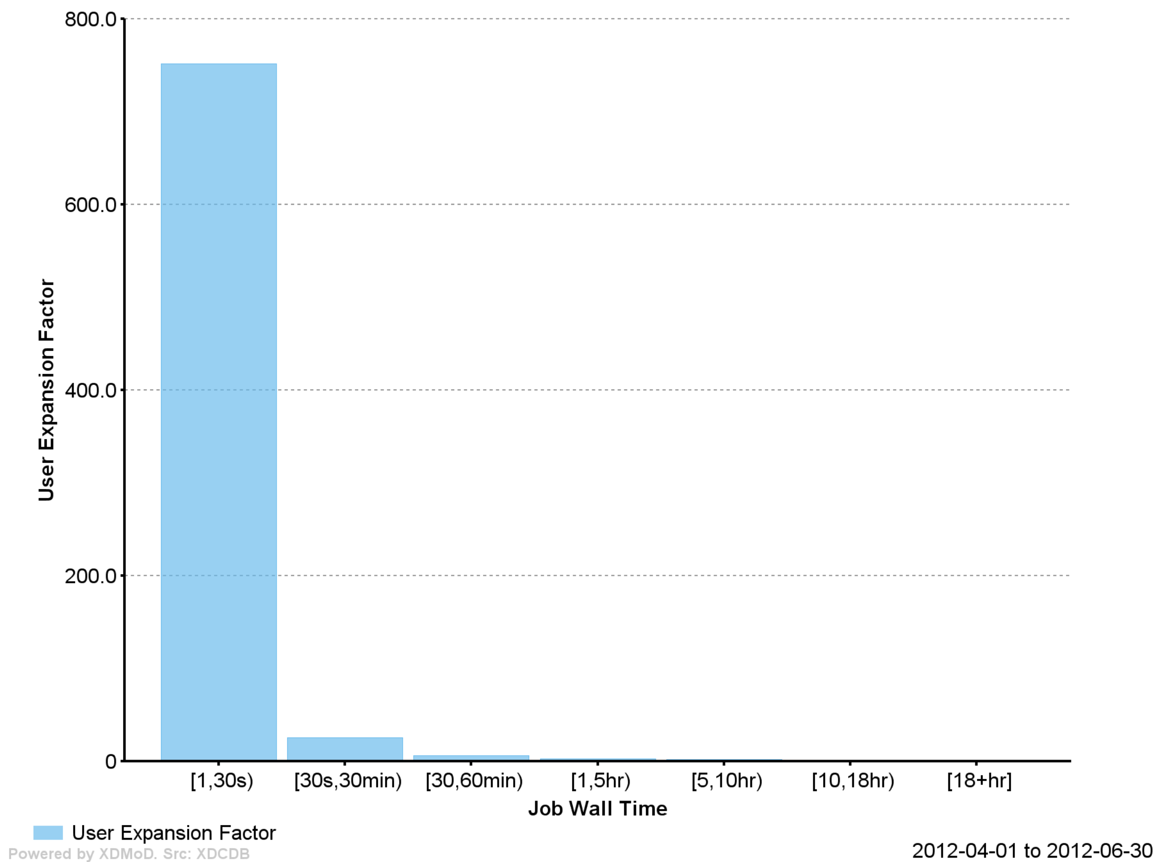
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = SDSC-TRESTLES

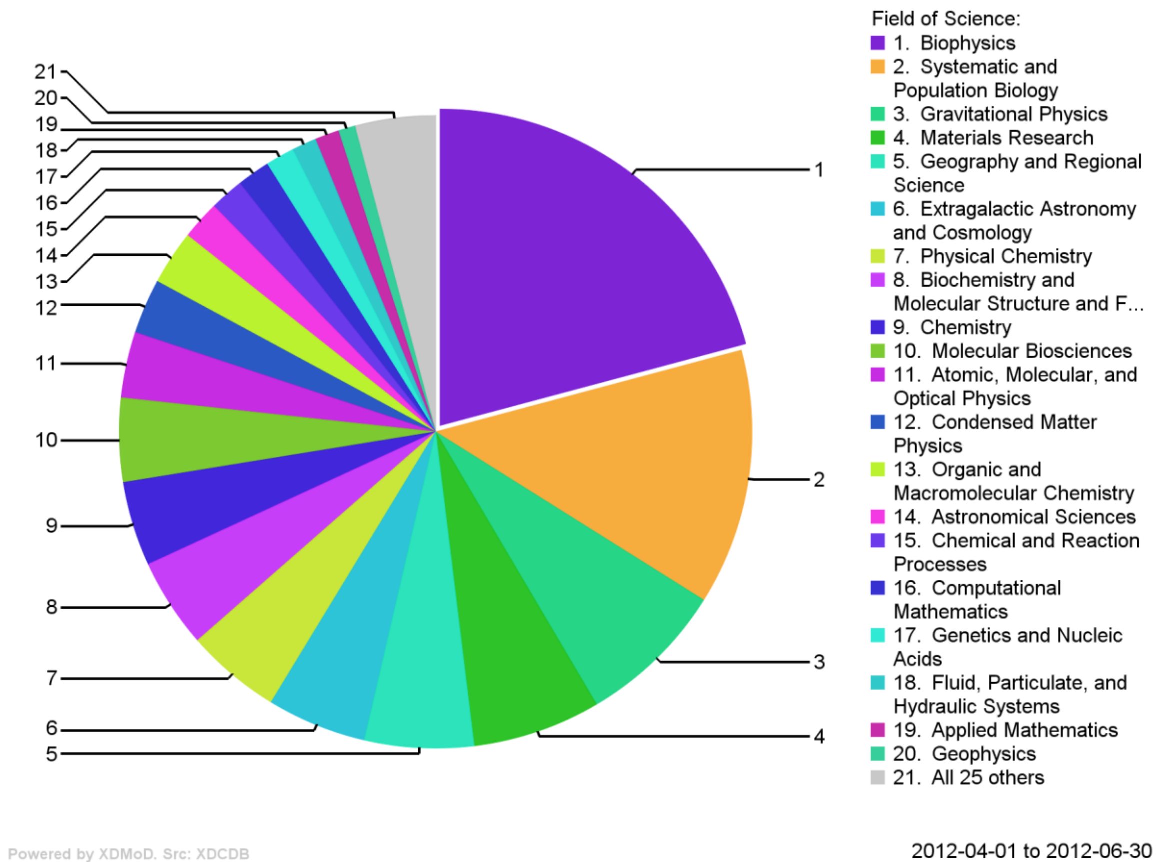
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = SDSC-TRESTLES

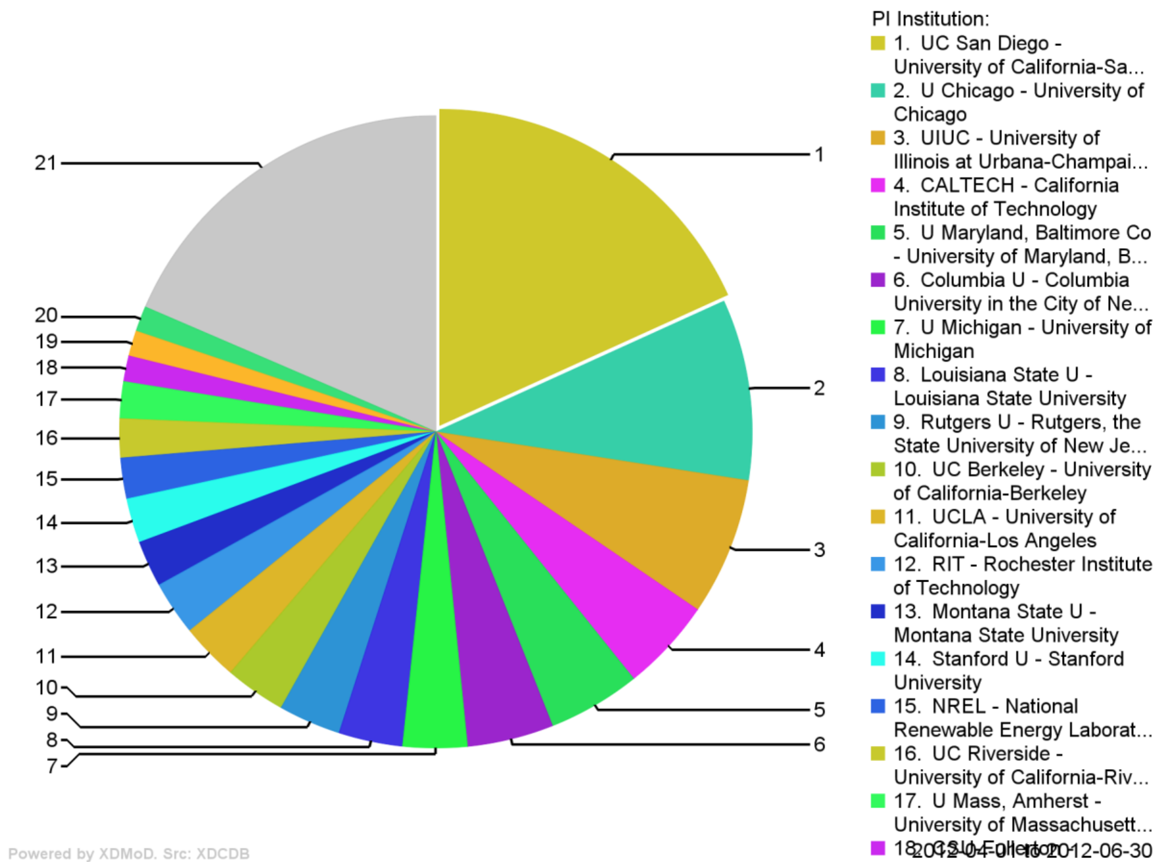
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = SDSC-TRESTLES

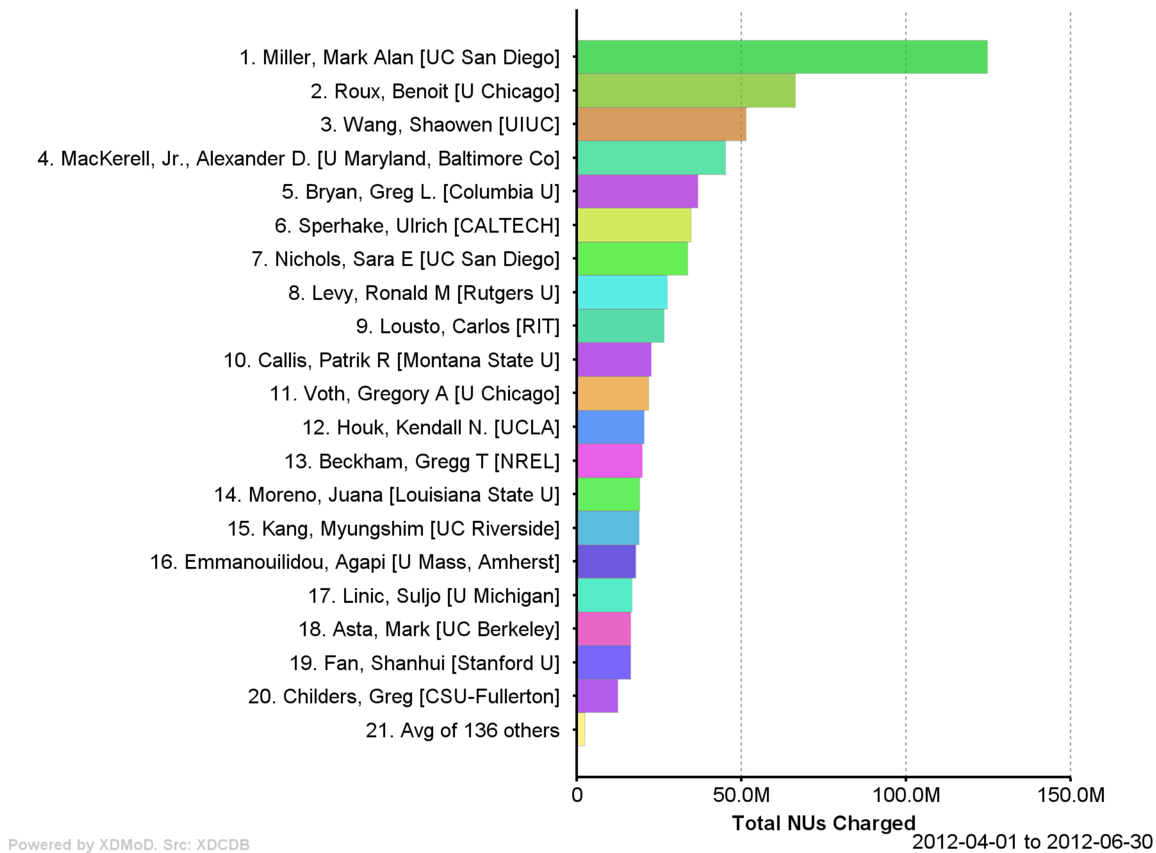
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = SDSC-TRESTLES

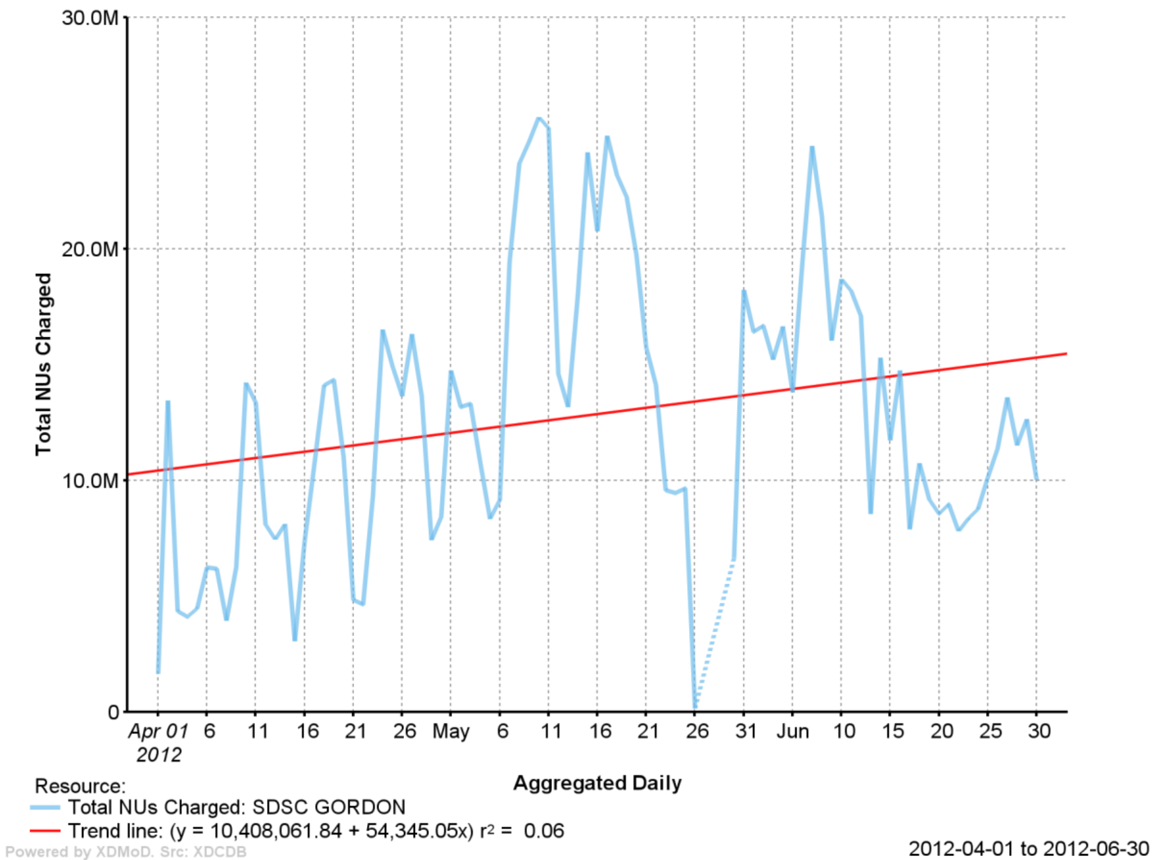
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = SDSC-GORDON

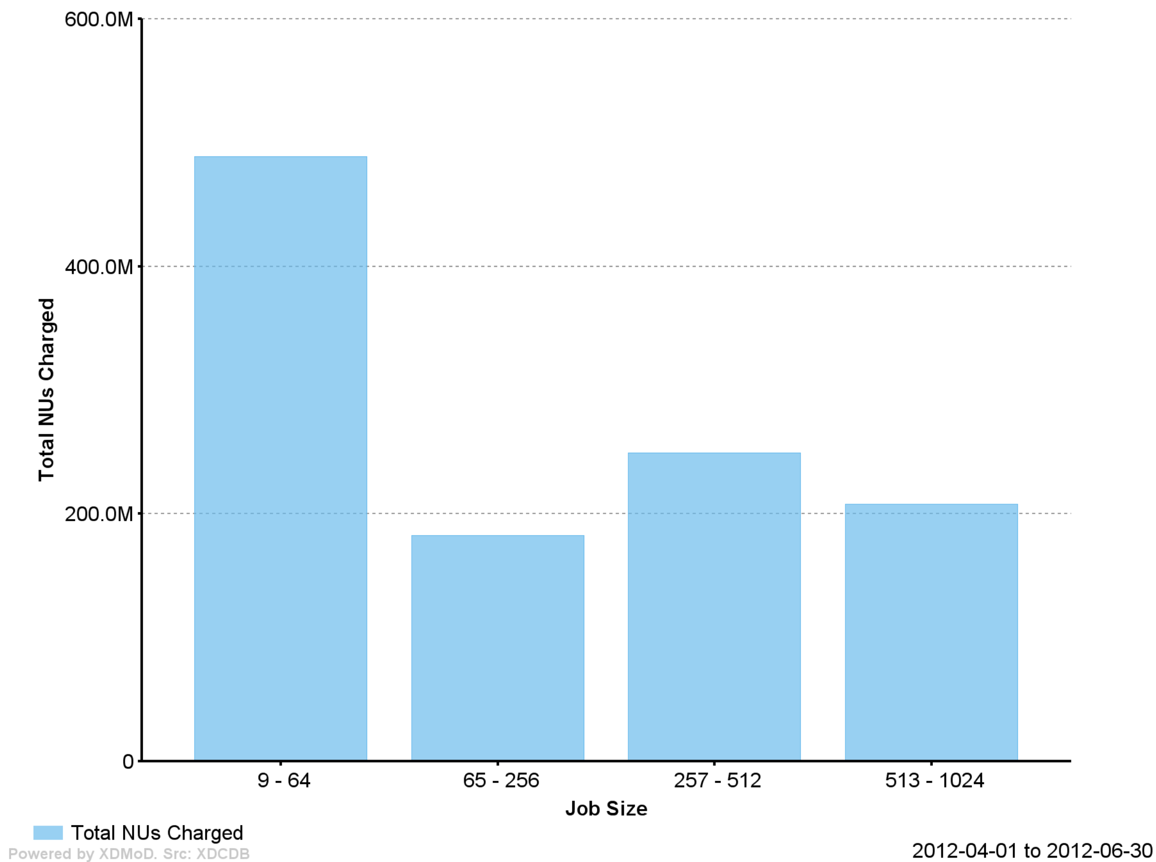
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = SDSC-GORDON

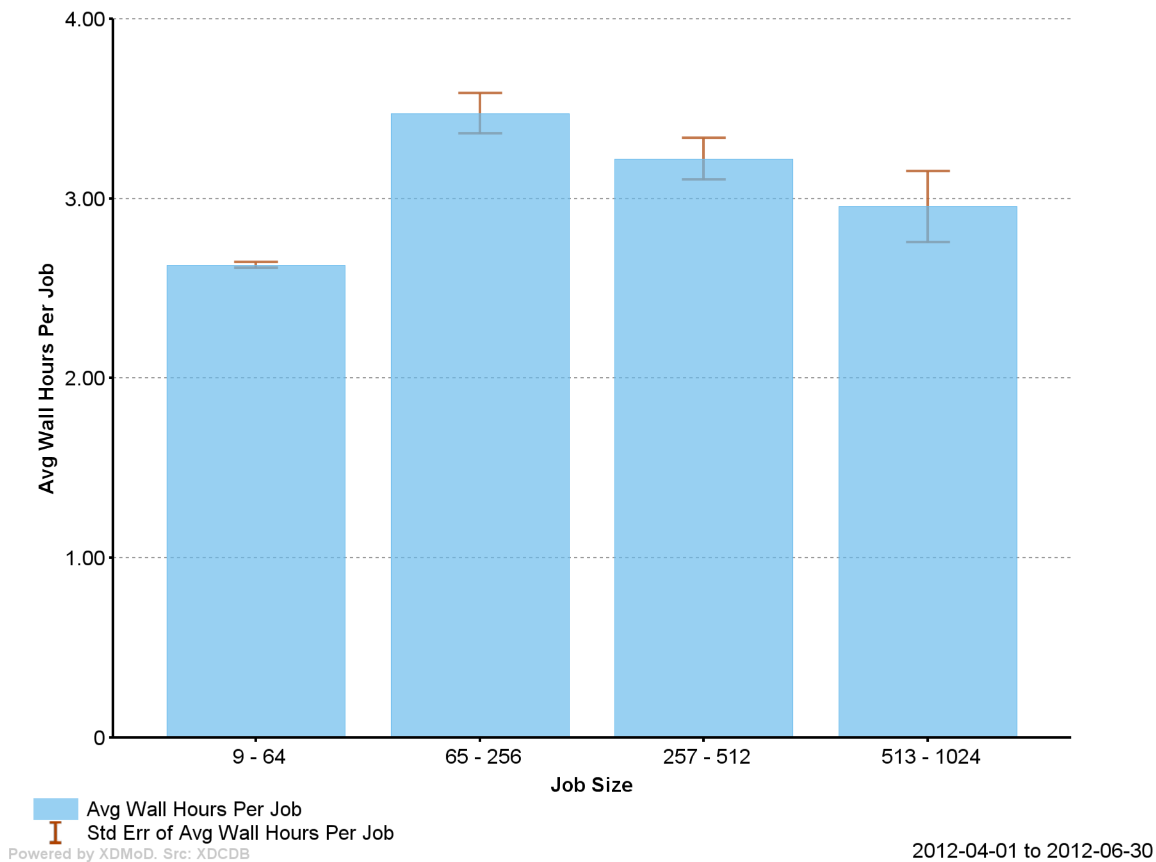
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = SDSC-GORDON

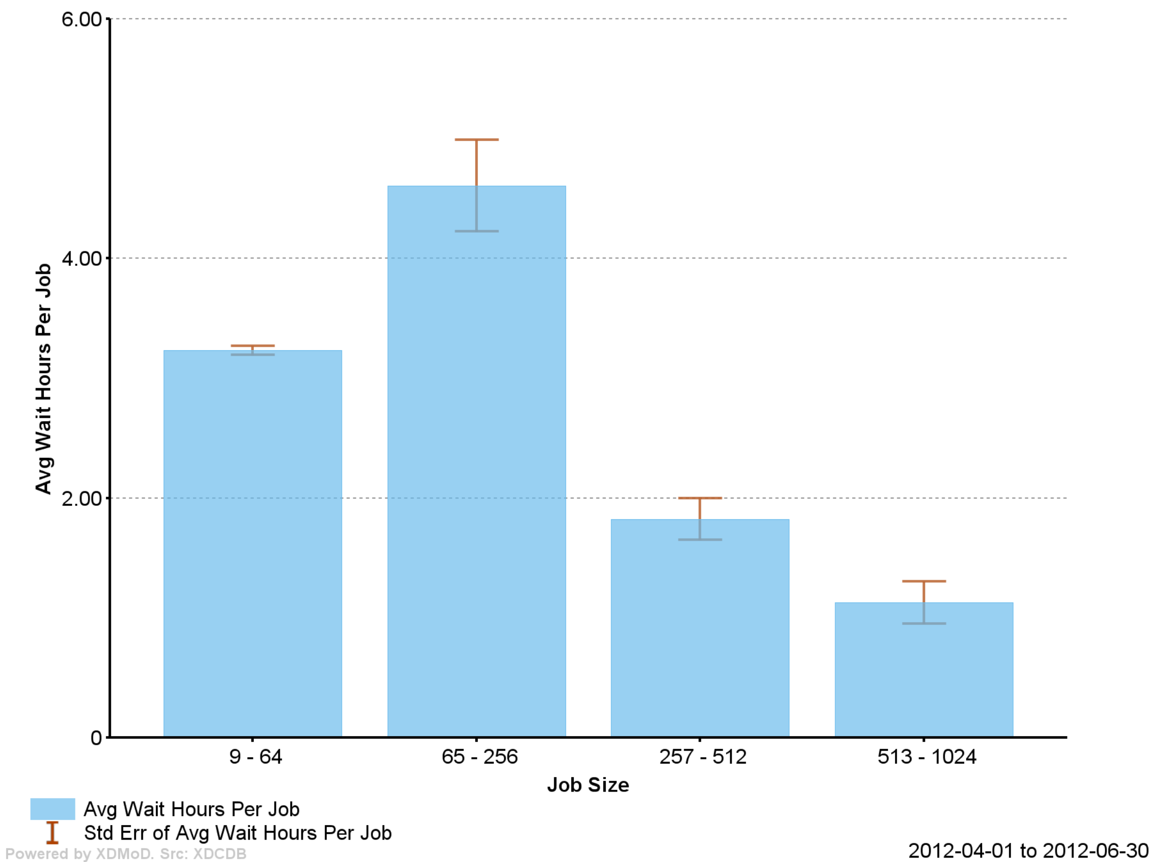
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = SDSC-GORDON

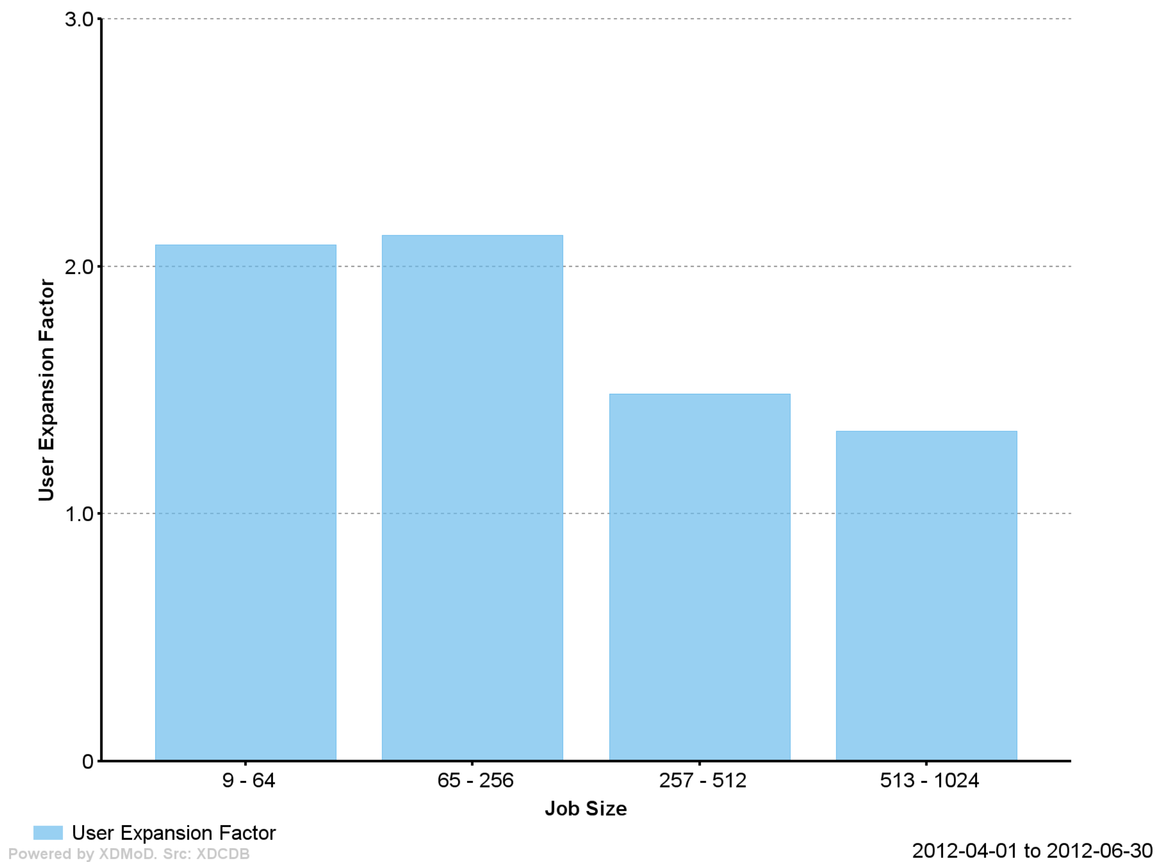
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = SDSC-GORDON

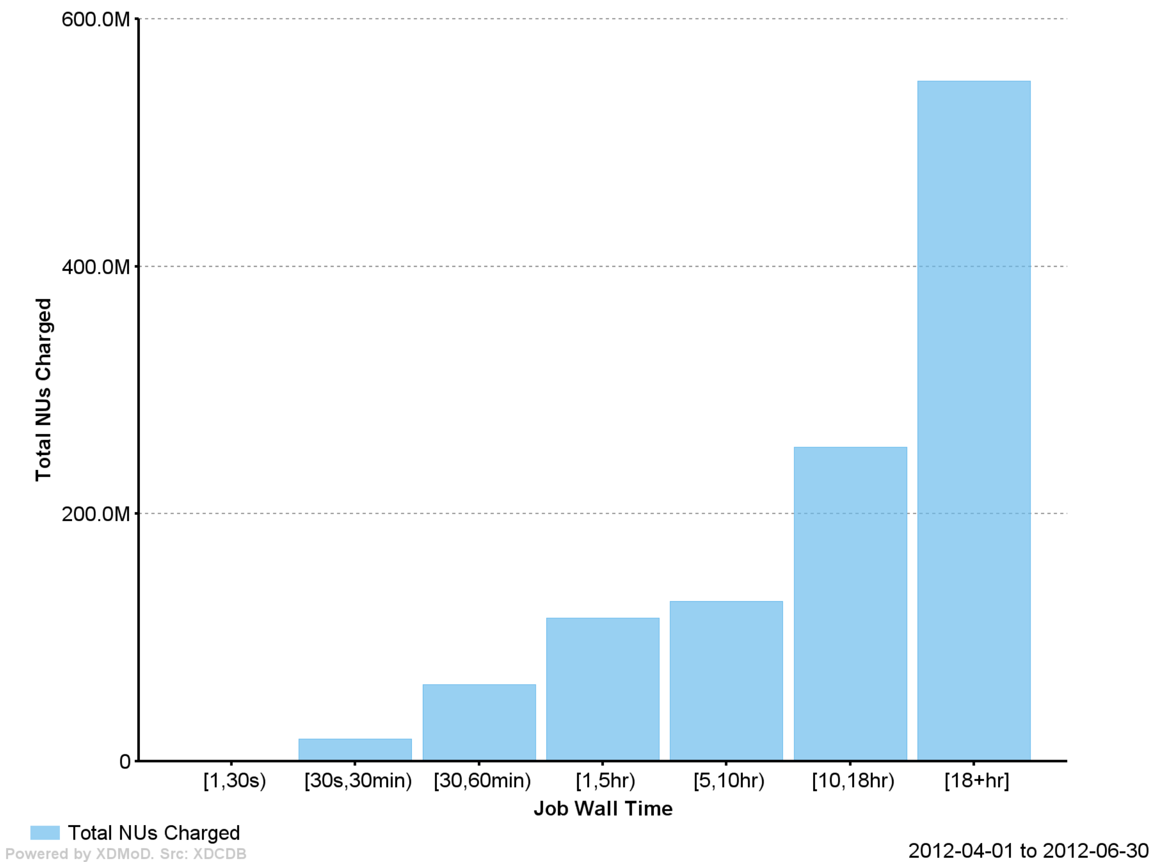
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = SDSC-GORDON

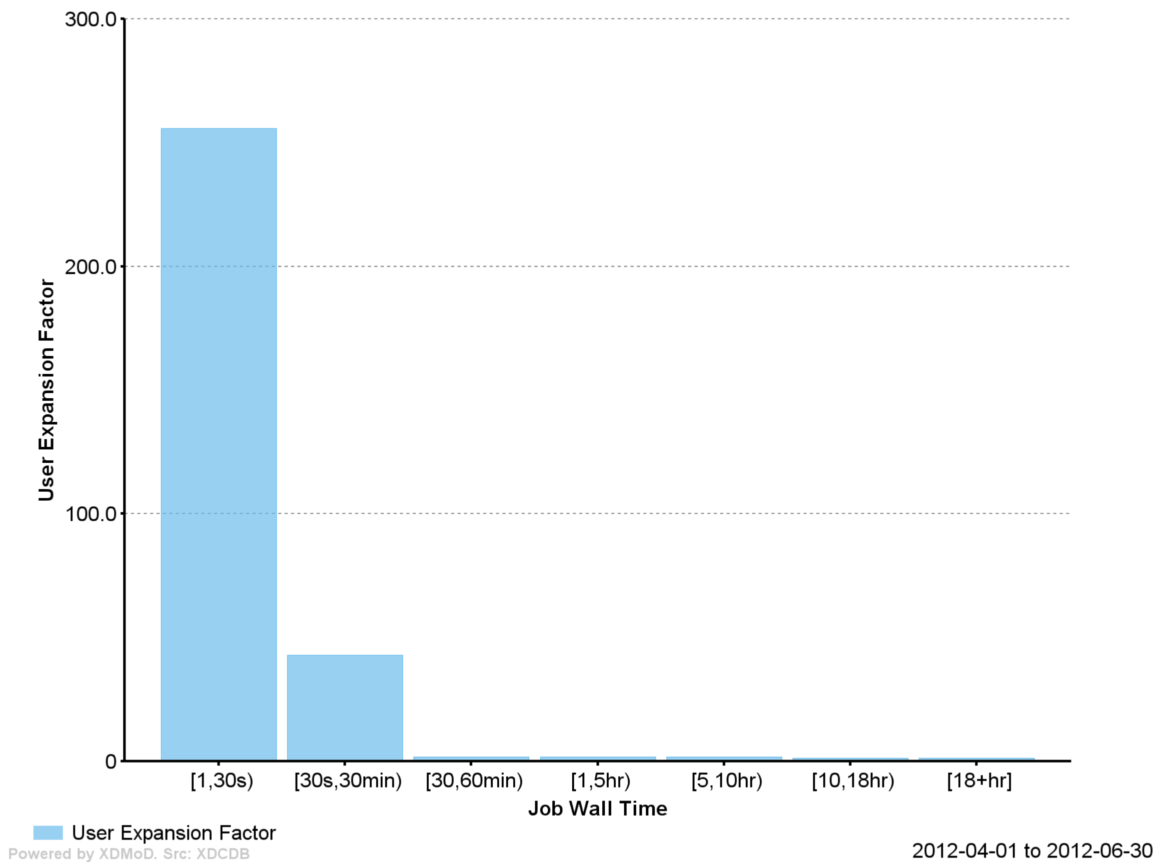
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = SDSC-GORDON

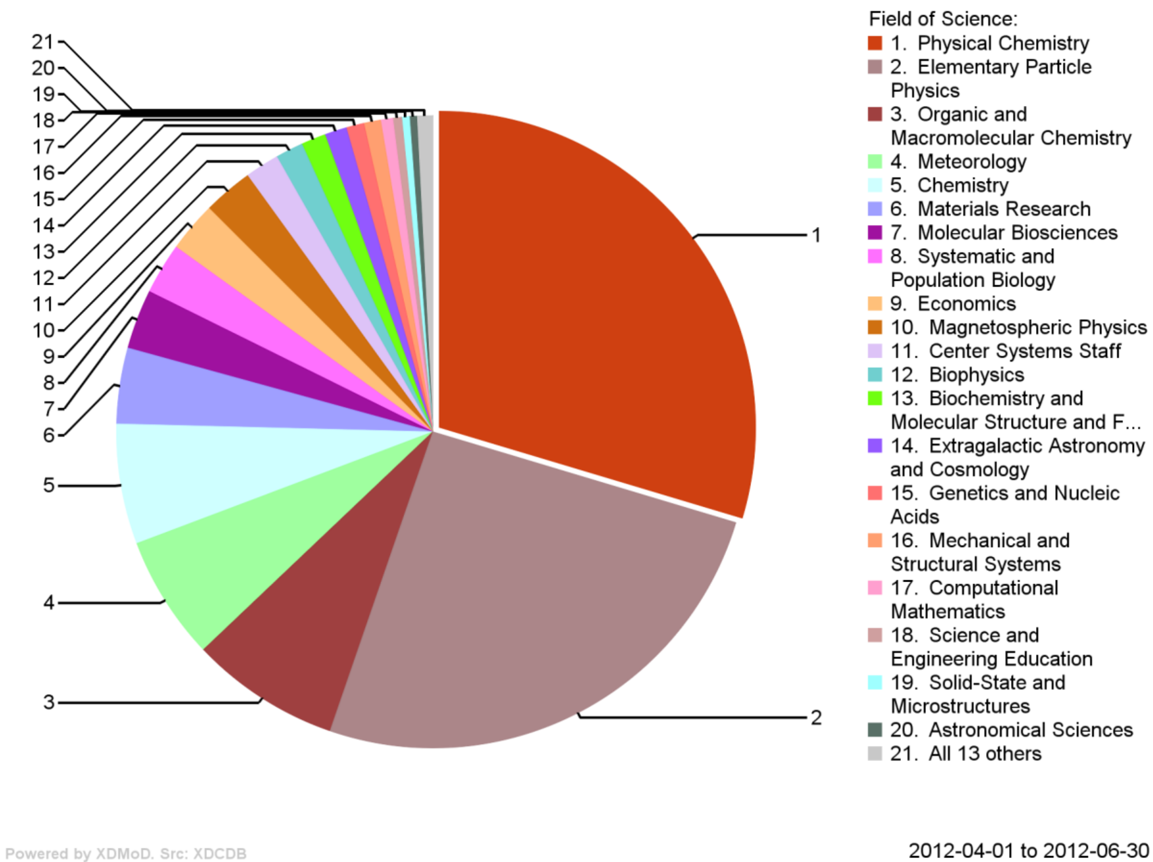
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = SDSC-GORDON

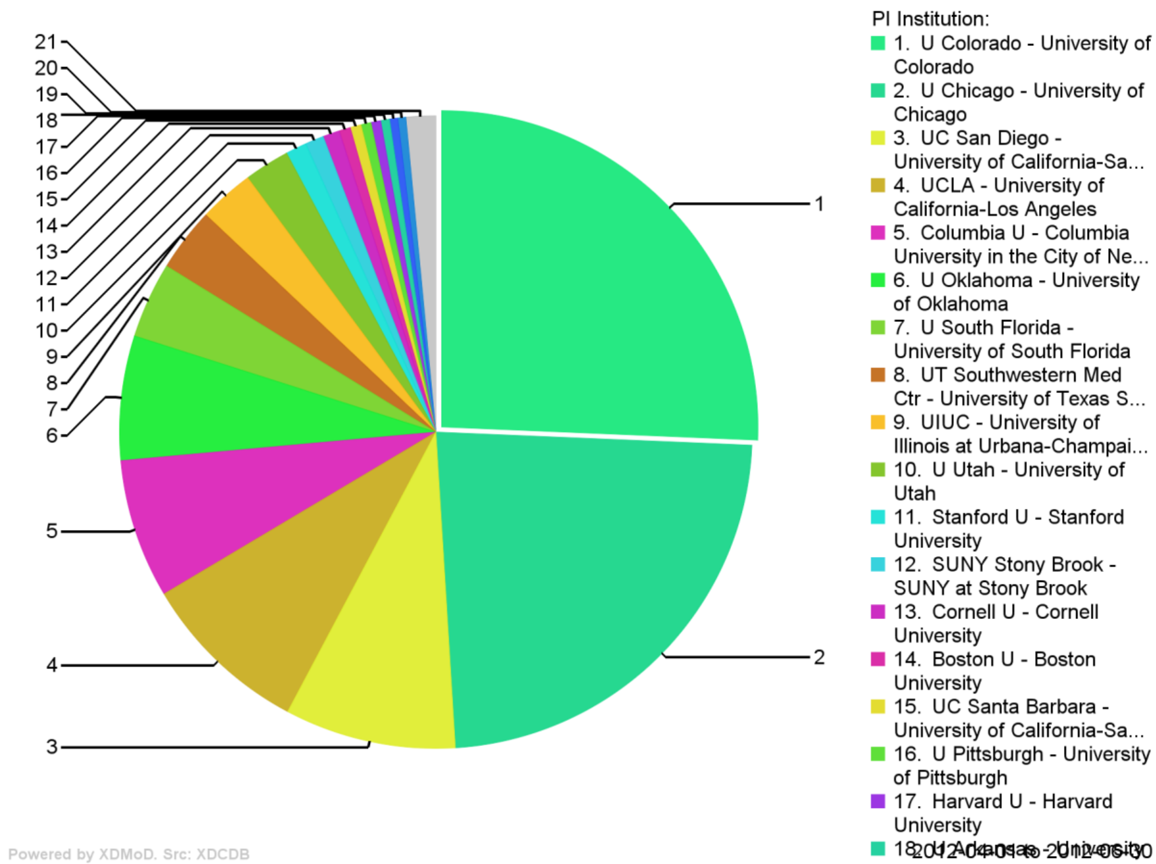
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = SDSC-GORDON

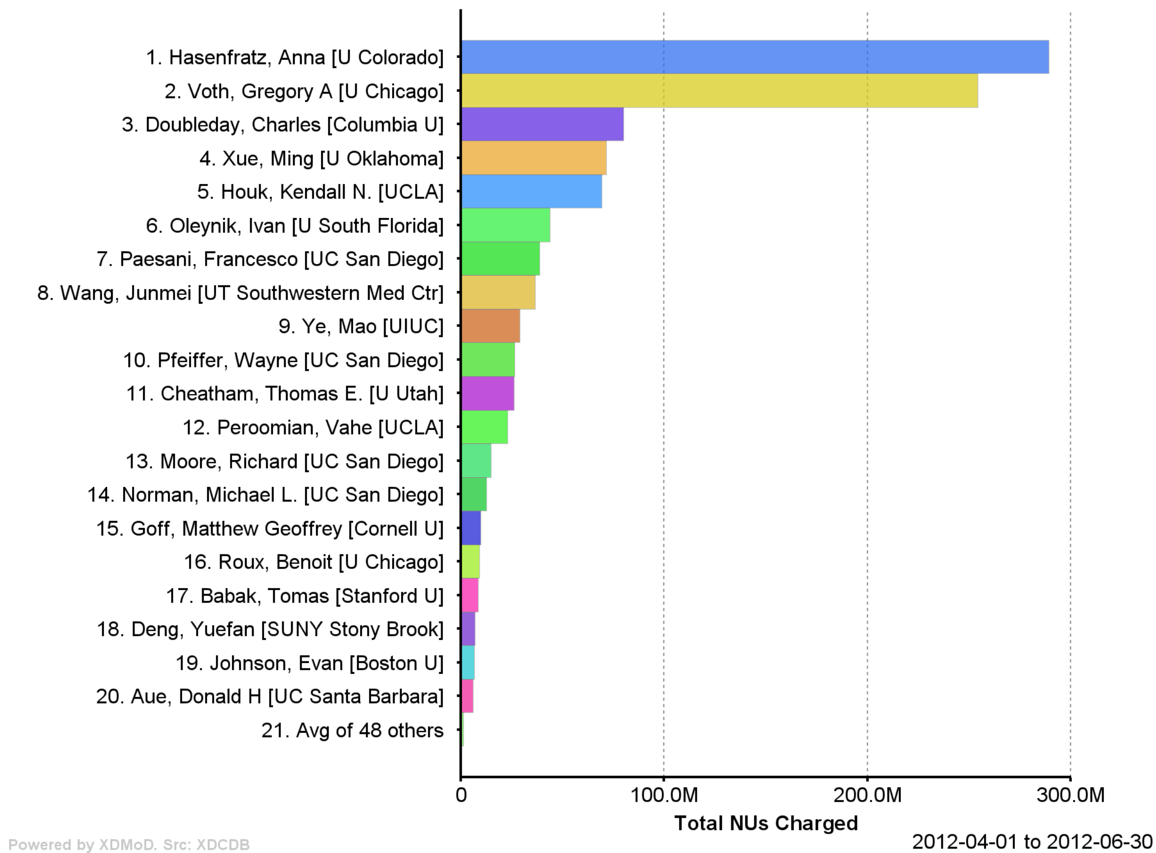
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = SDSC-GORDON

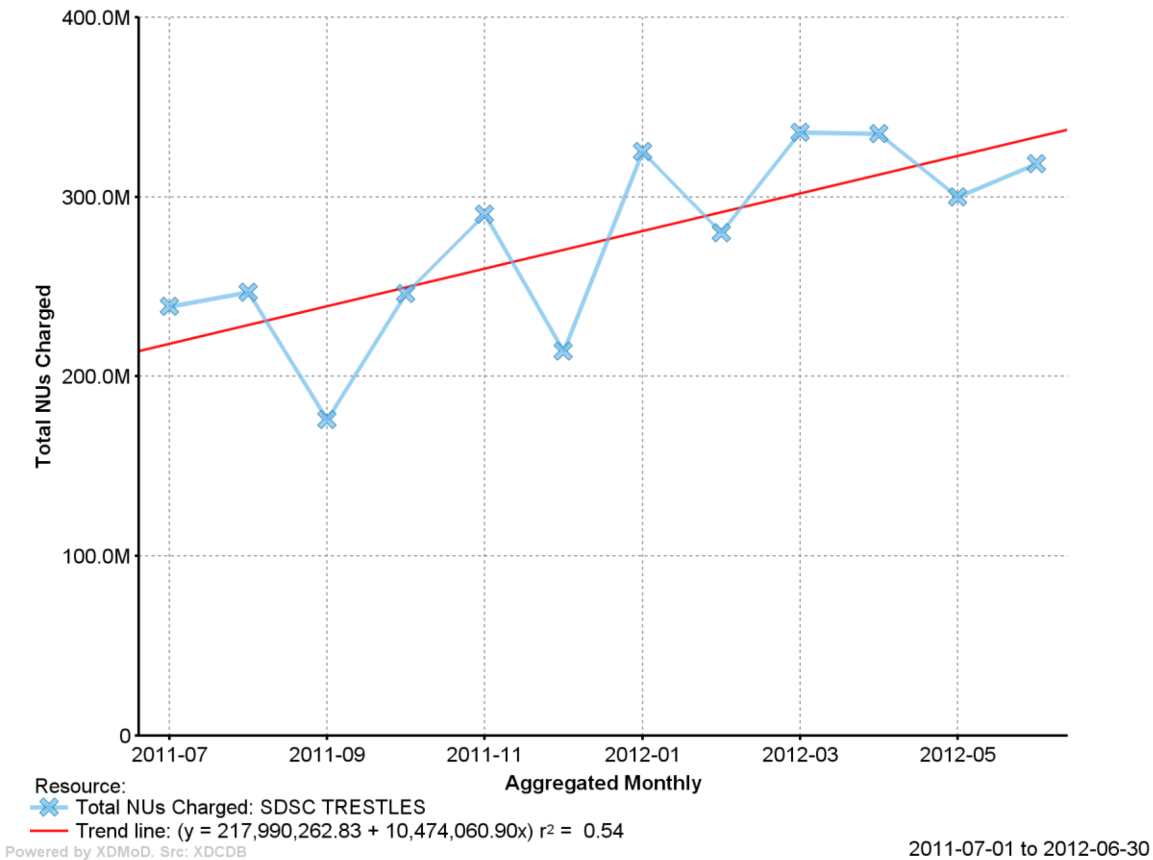
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = SDSC-TRESTLES

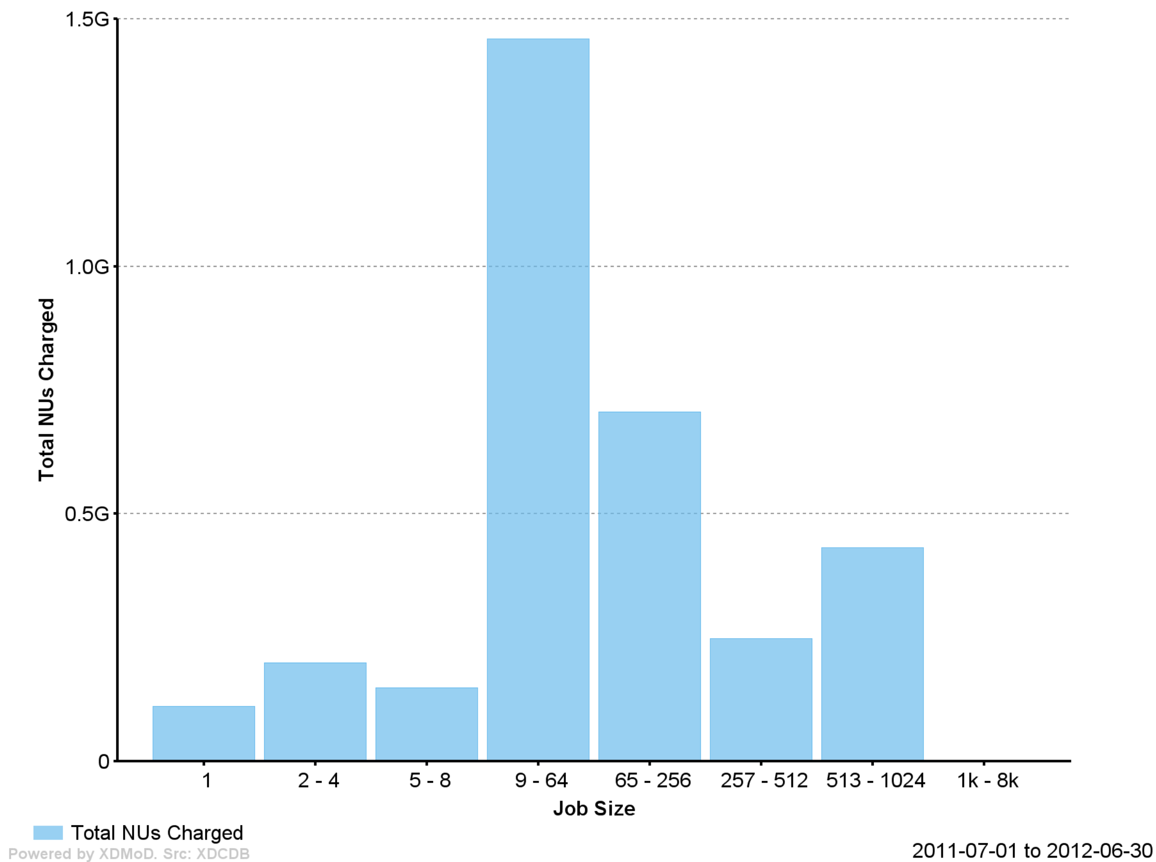
2011-07-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = SDSC-TRESTLES

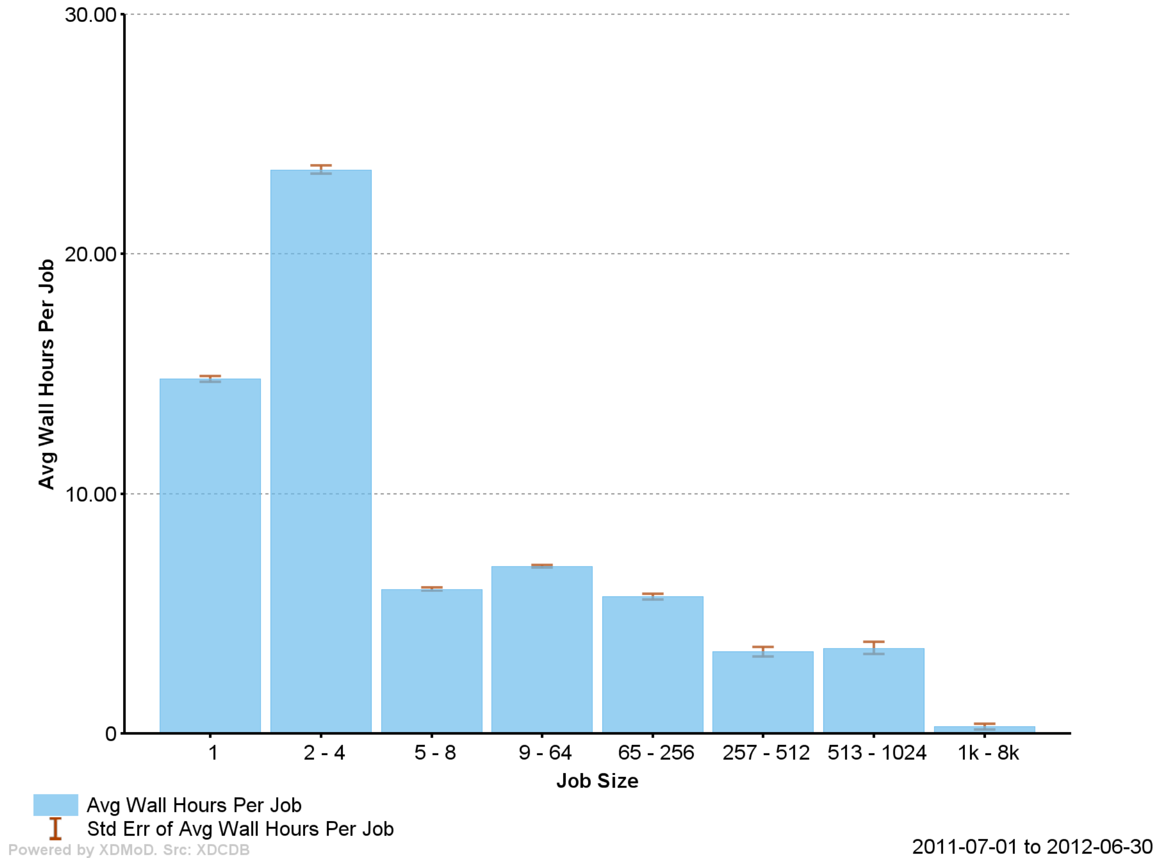
2011-07-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = SDSC-TRESTLES

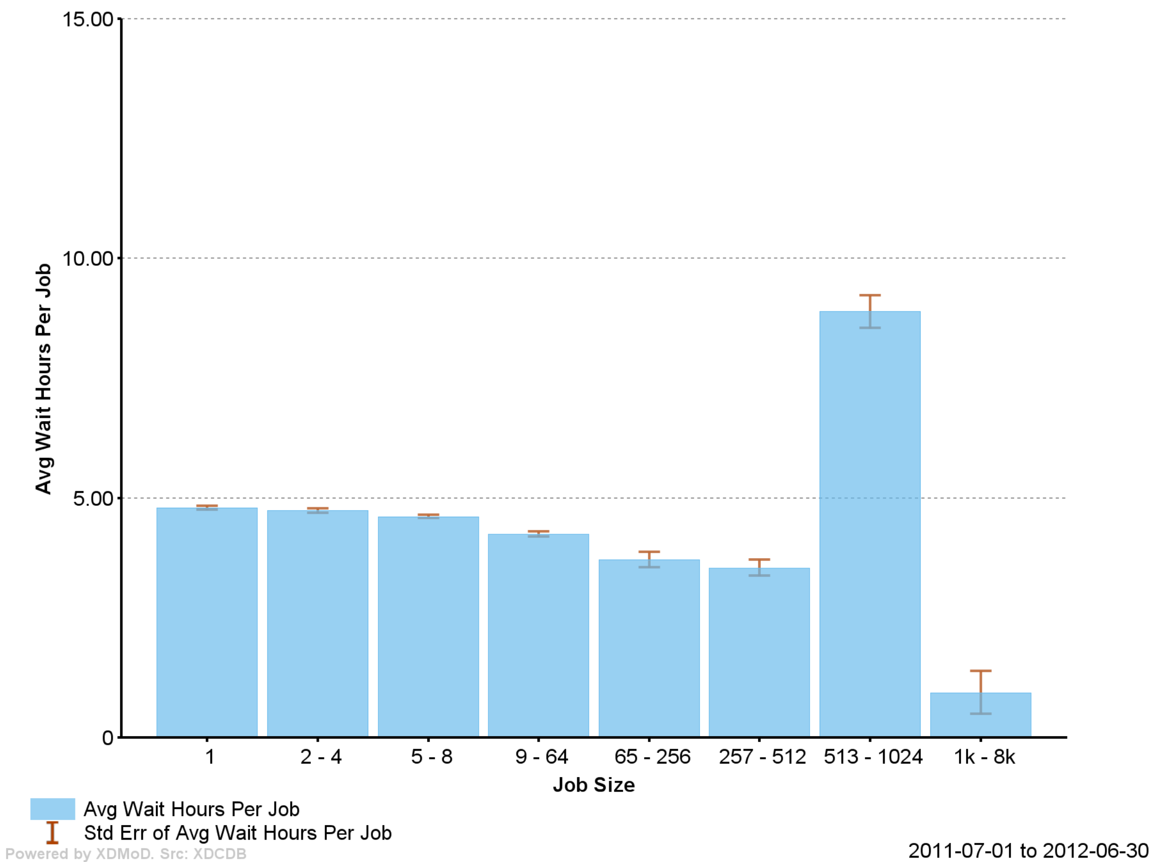
2011-07-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = SDSC-TRESTLES

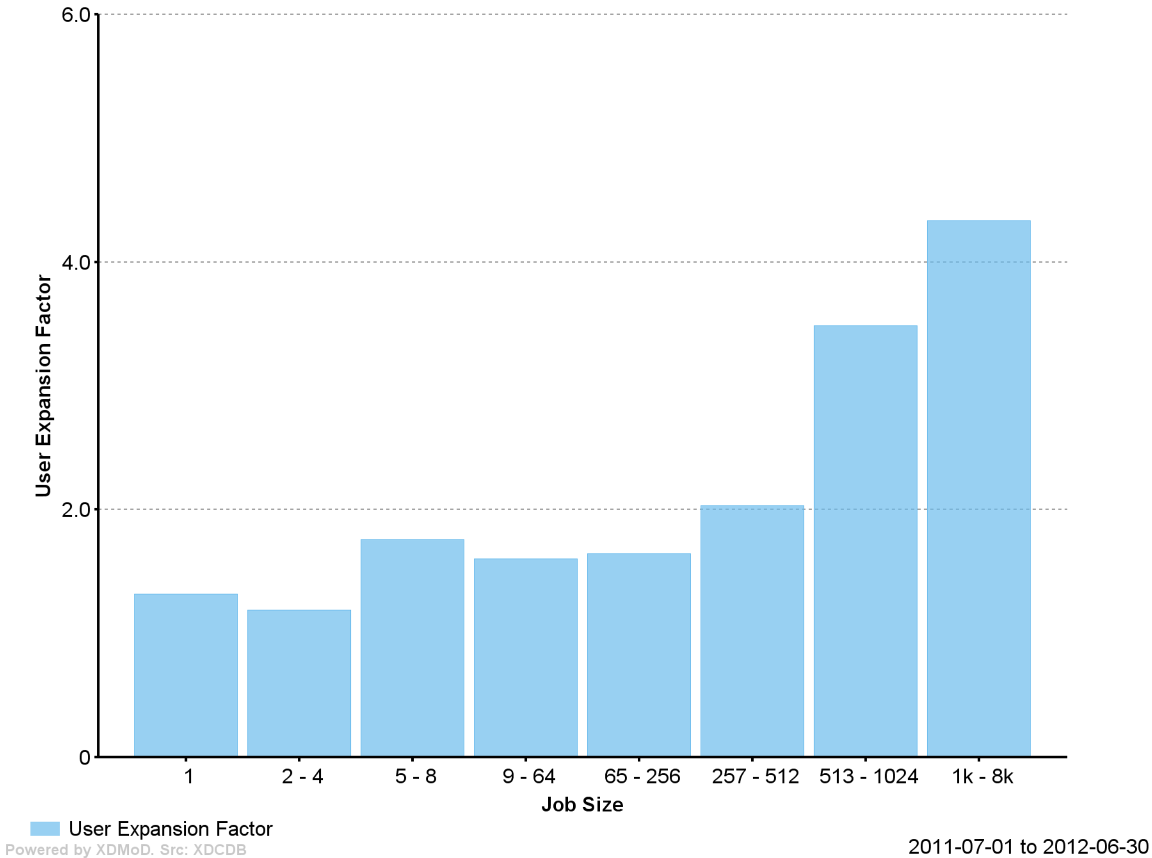
2011-07-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = SDSC-TRESTLES

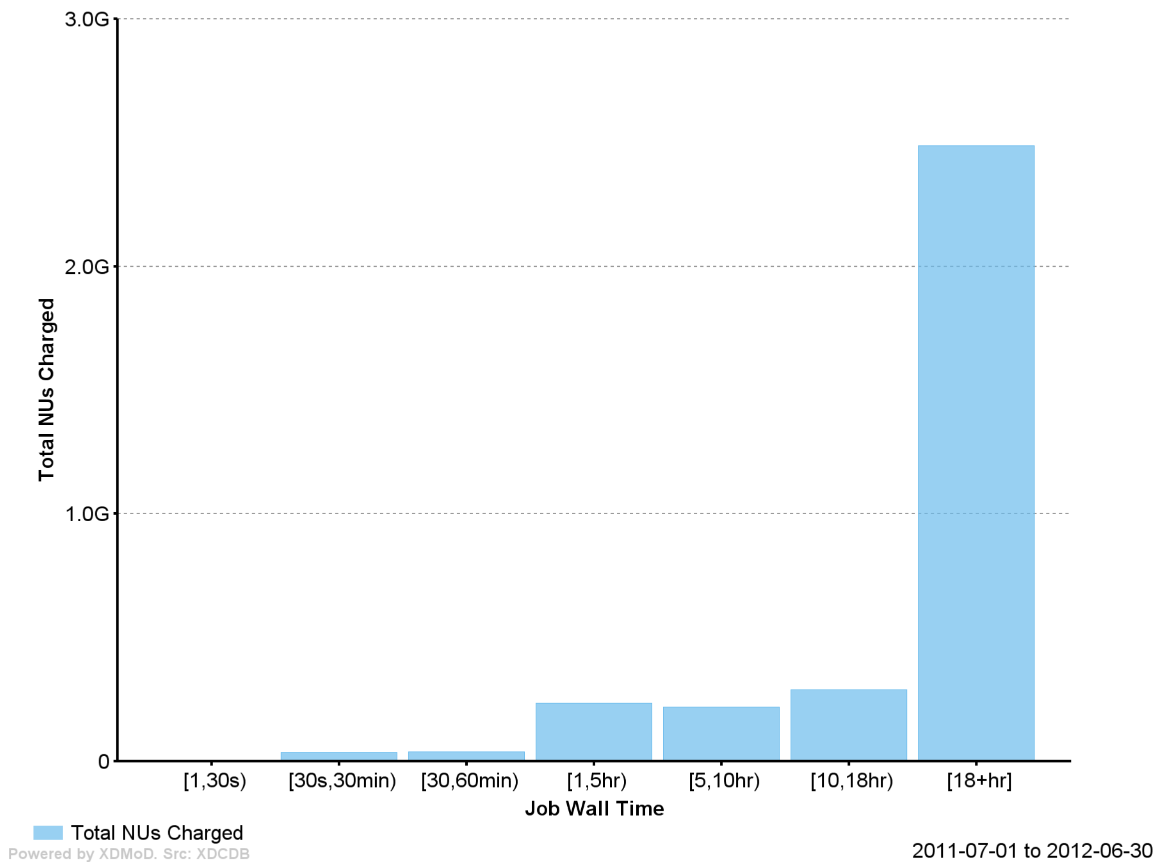
2011-07-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = SDSC-TRESTLES

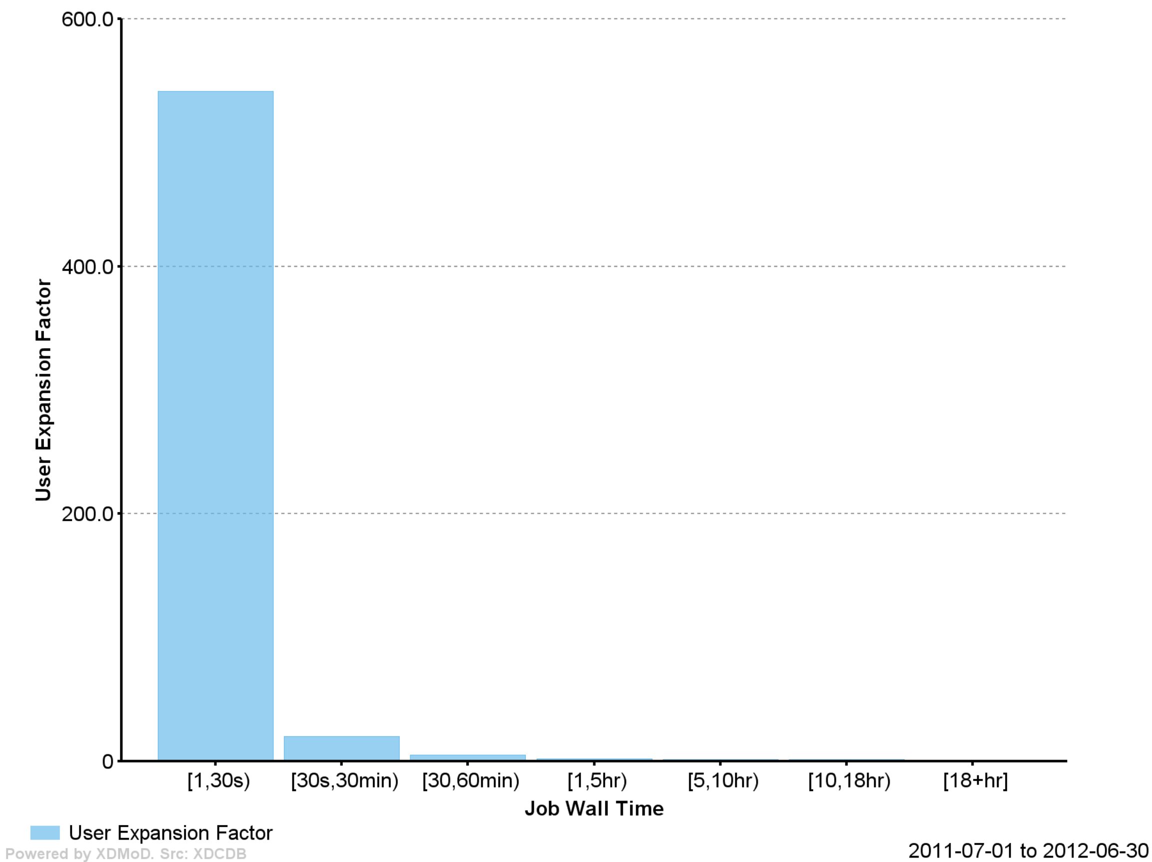
2011-07-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = SDSC-TRESTLES

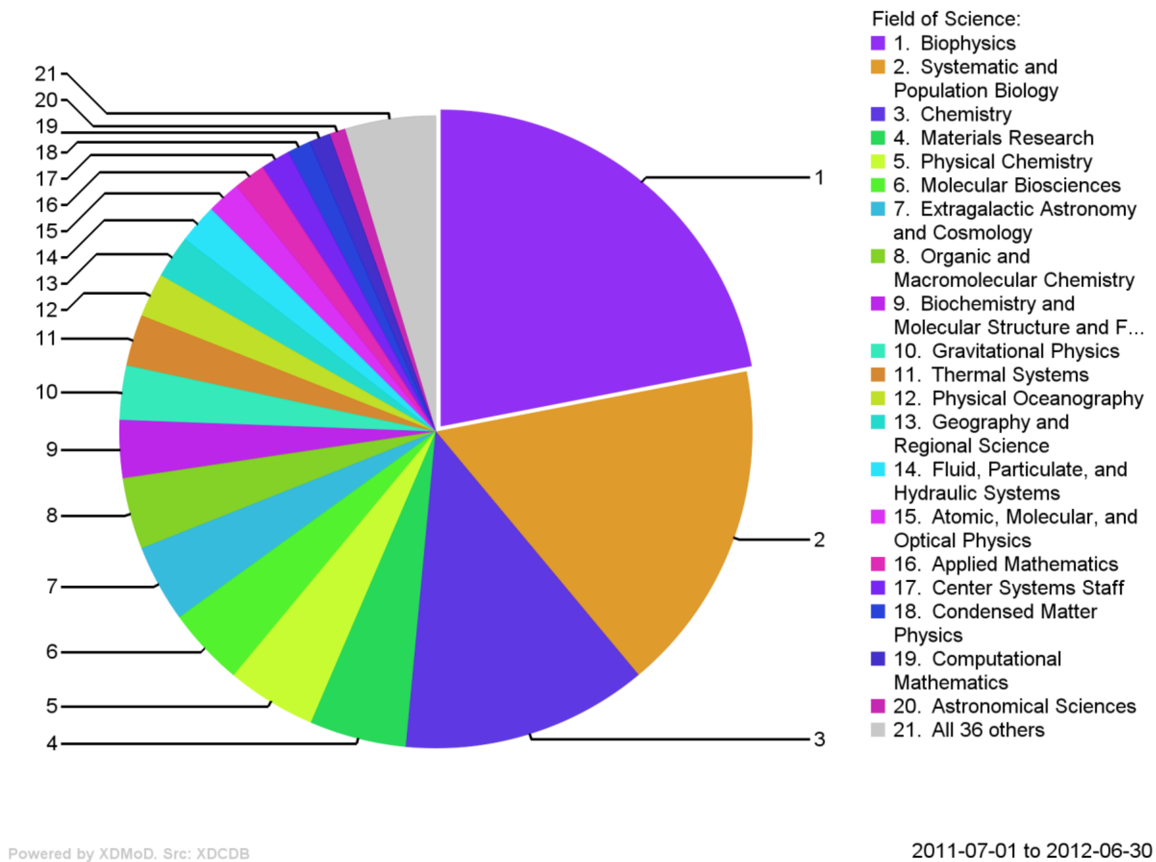
2011-07-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = SDSC-TRESTLES

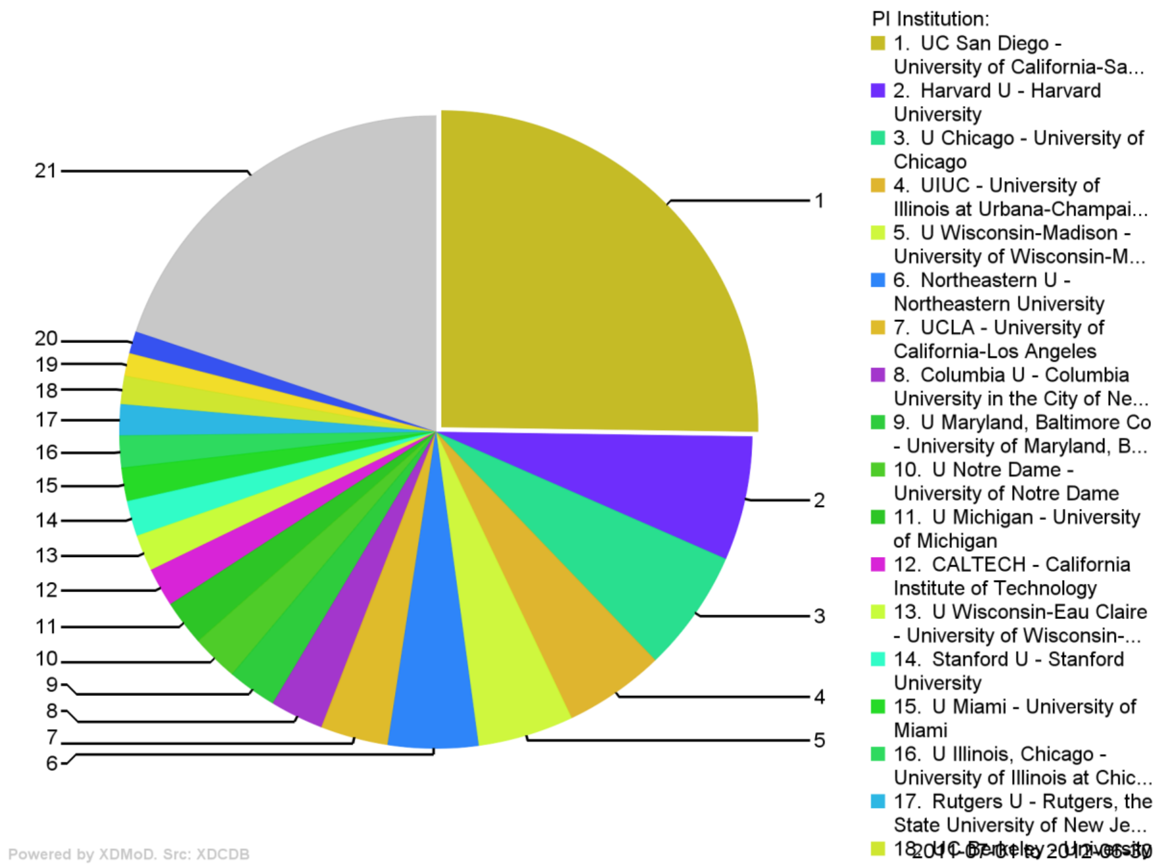
2011-07-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = SDSC-TRESTLES

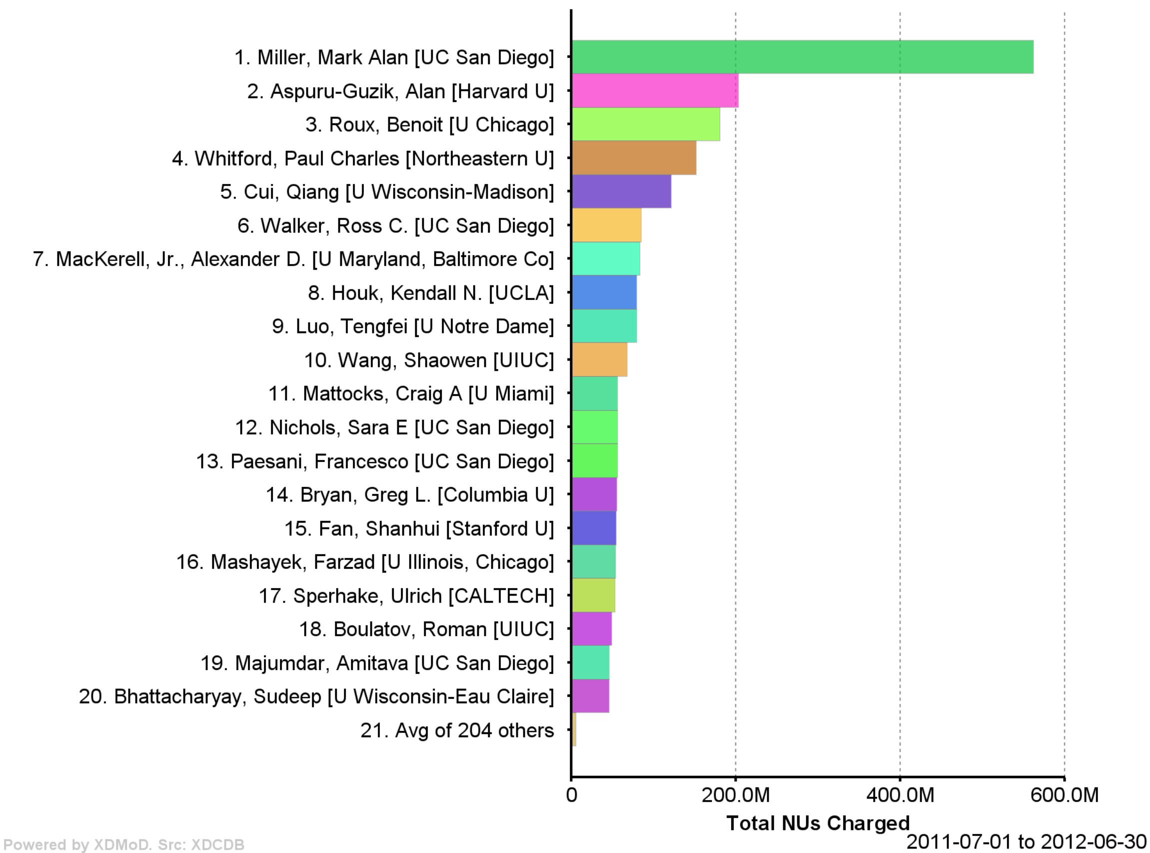
2011-07-01 to 2012-06-30



Total NUs Charged by PI

Resource = SDSC-TRESTLES

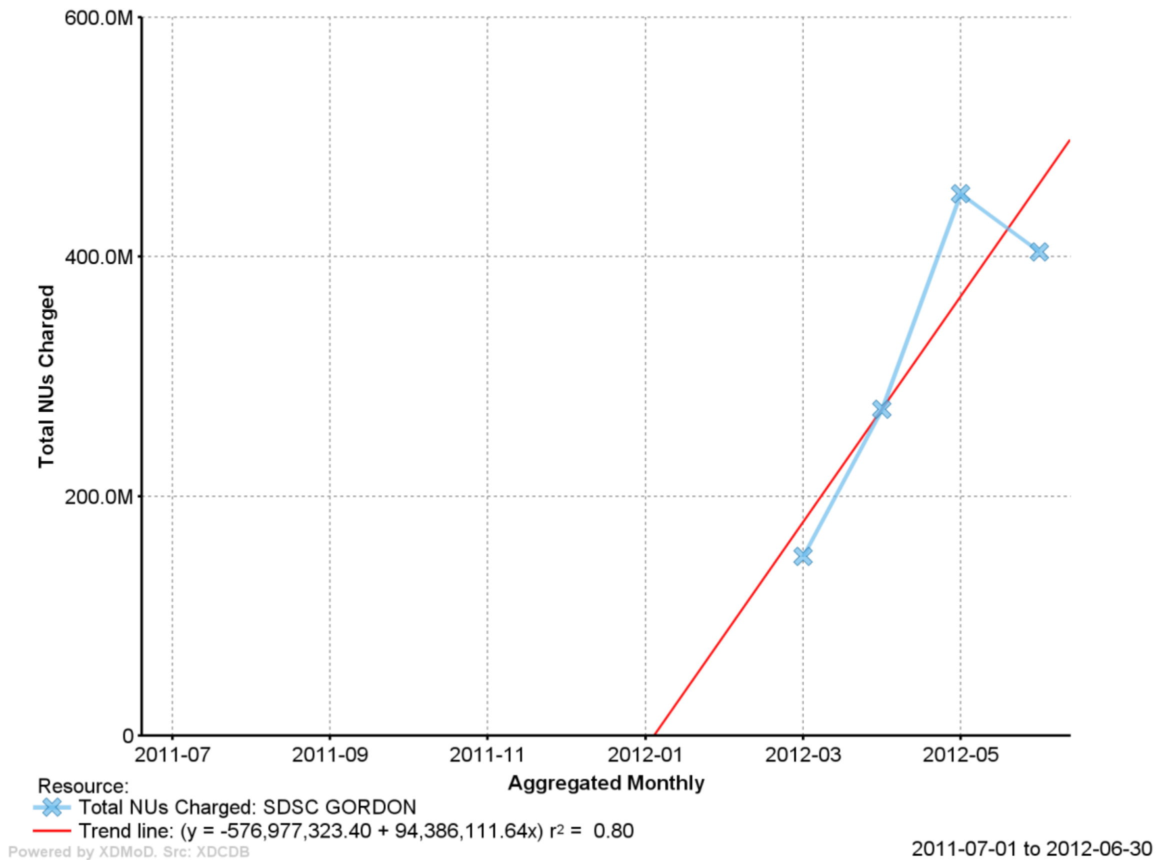
2011-07-01 to 2012-06-30



Total NUs Charged by Resource

Resource = SDSC-GORDON

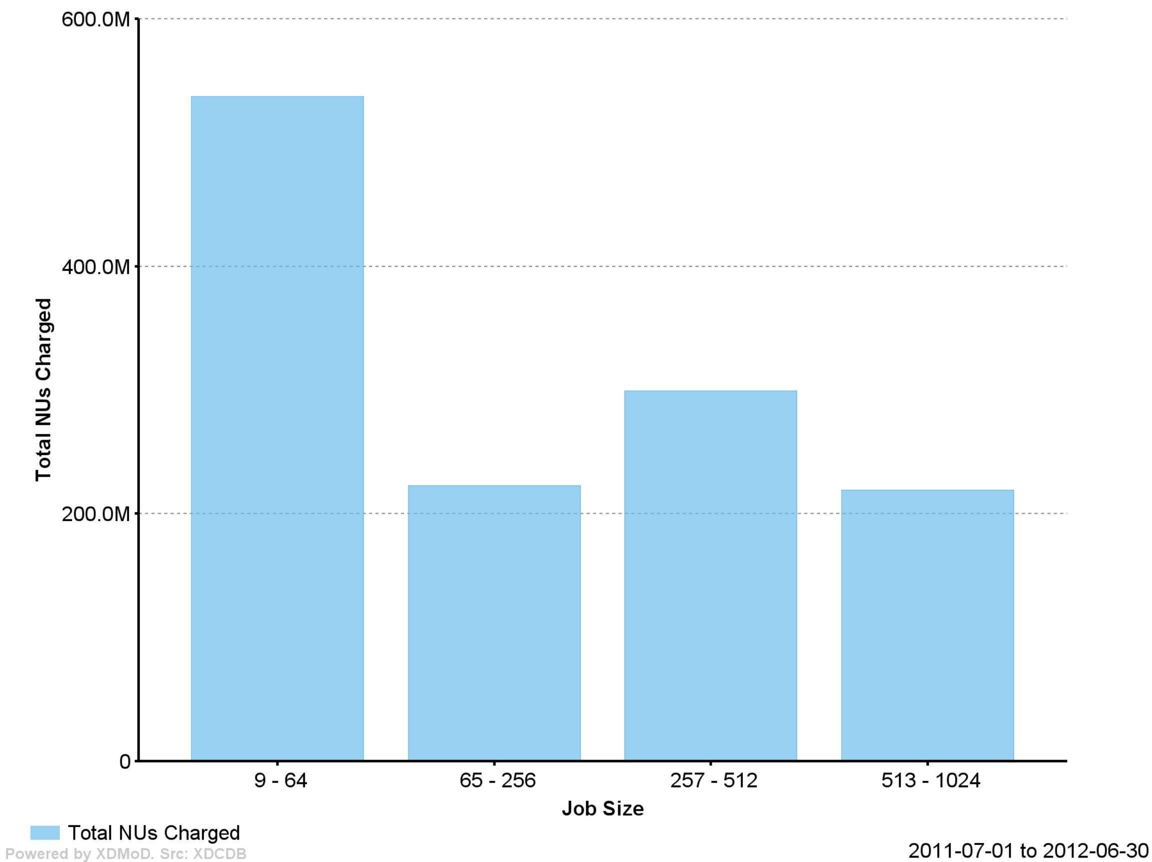
2011-07-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = SDSC-GORDON

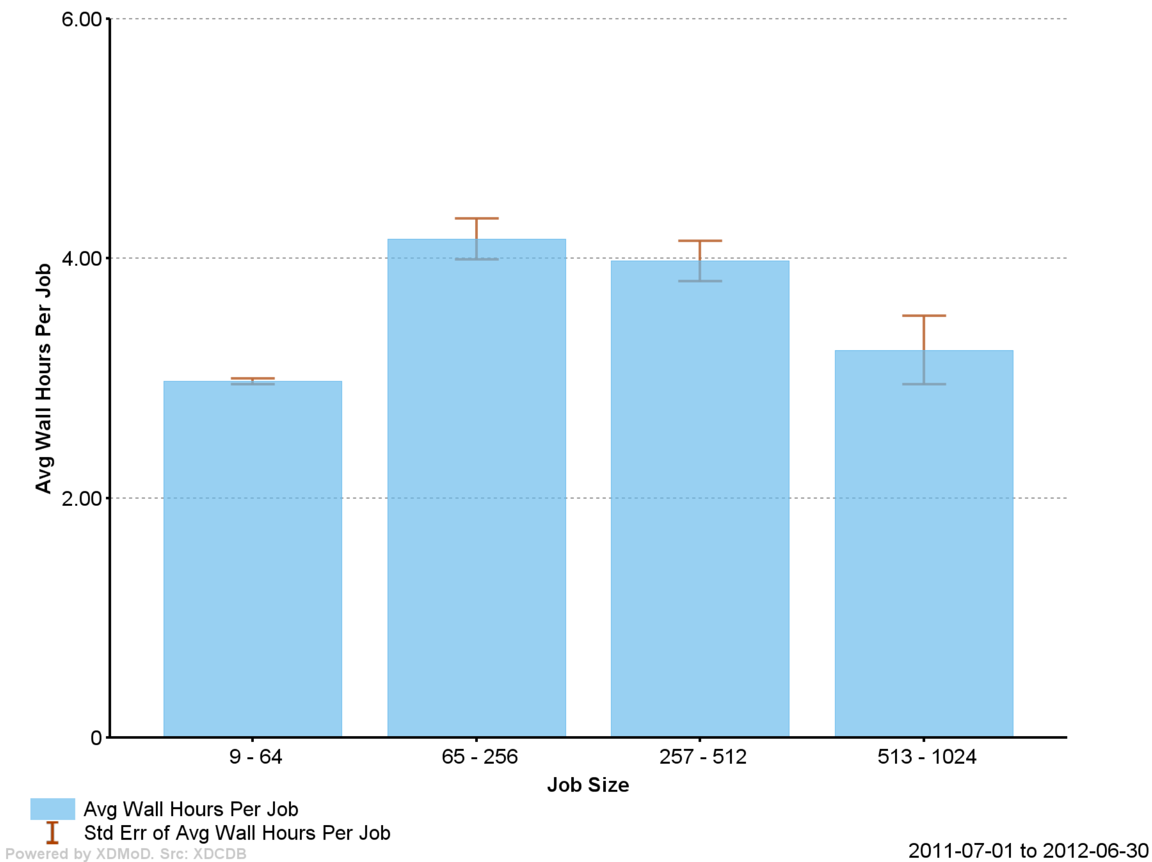
2011-07-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = SDSC-GORDON

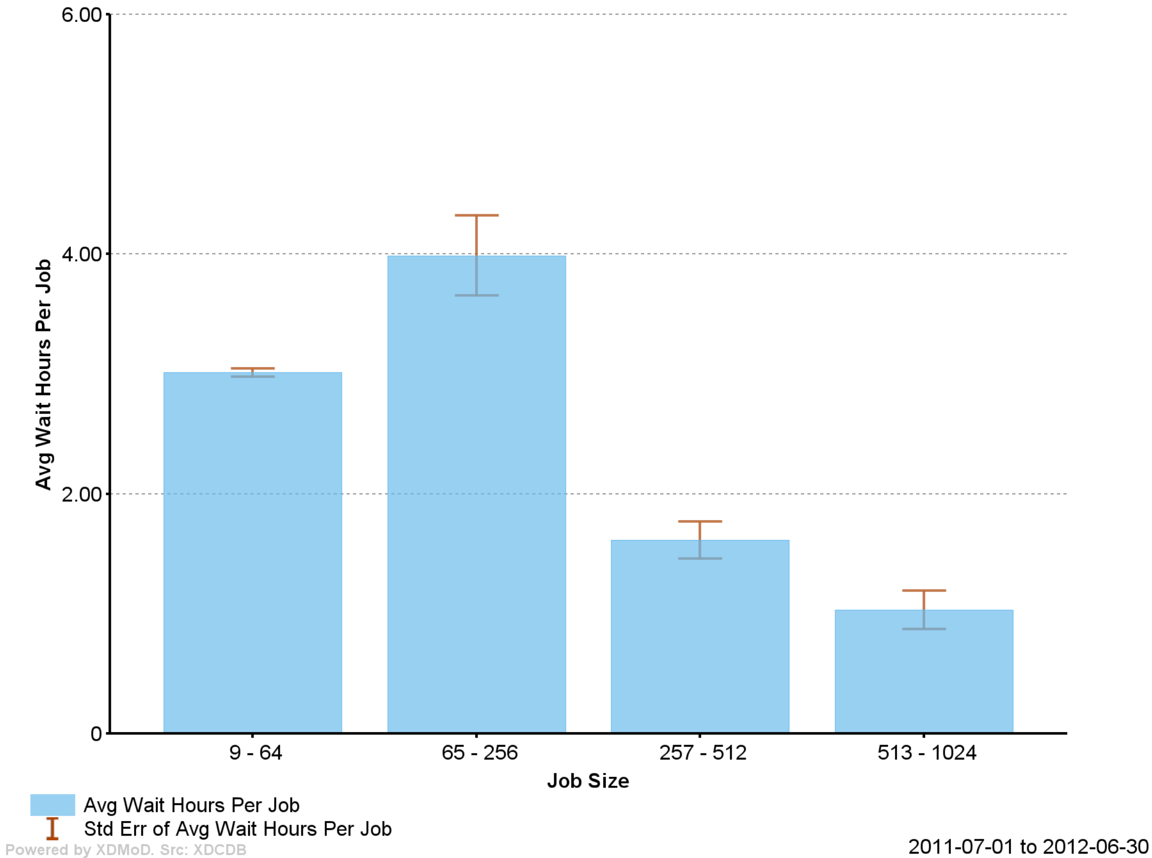
2011-07-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = SDSC-GORDON

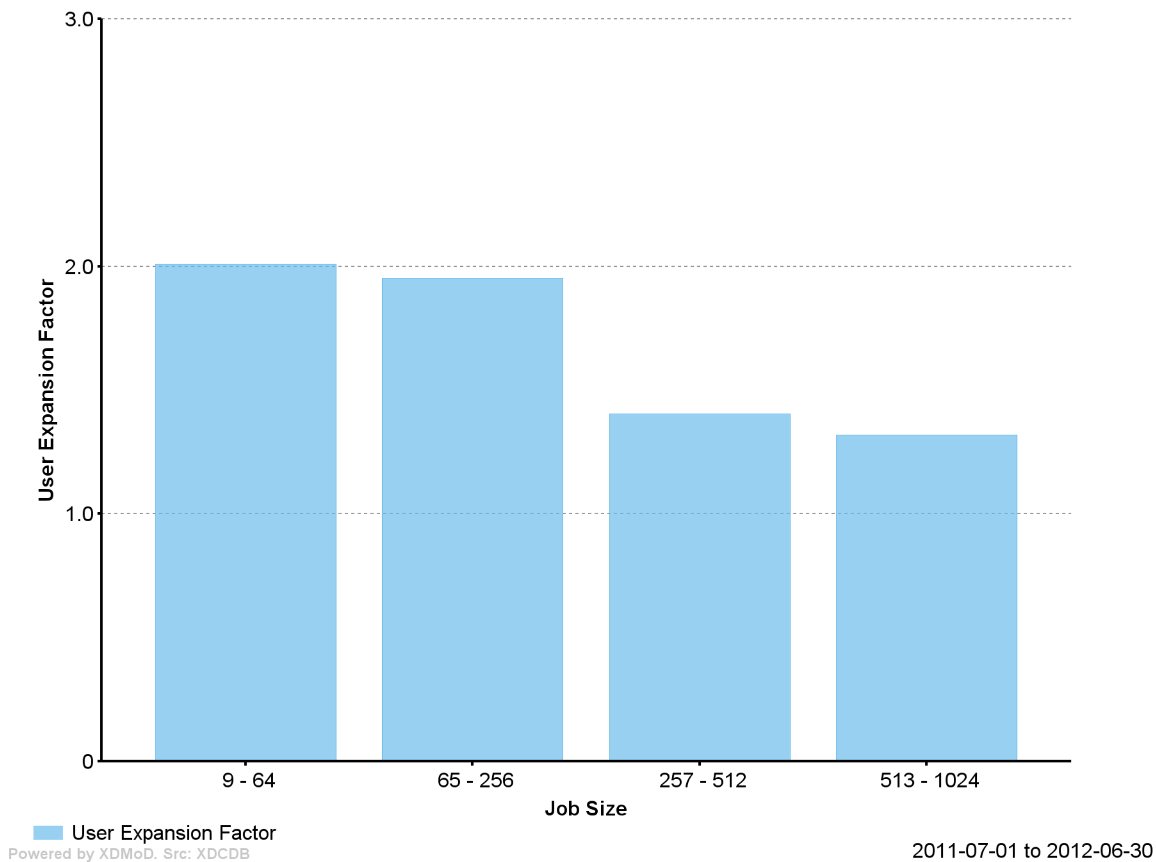
2011-07-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = SDSC-GORDON

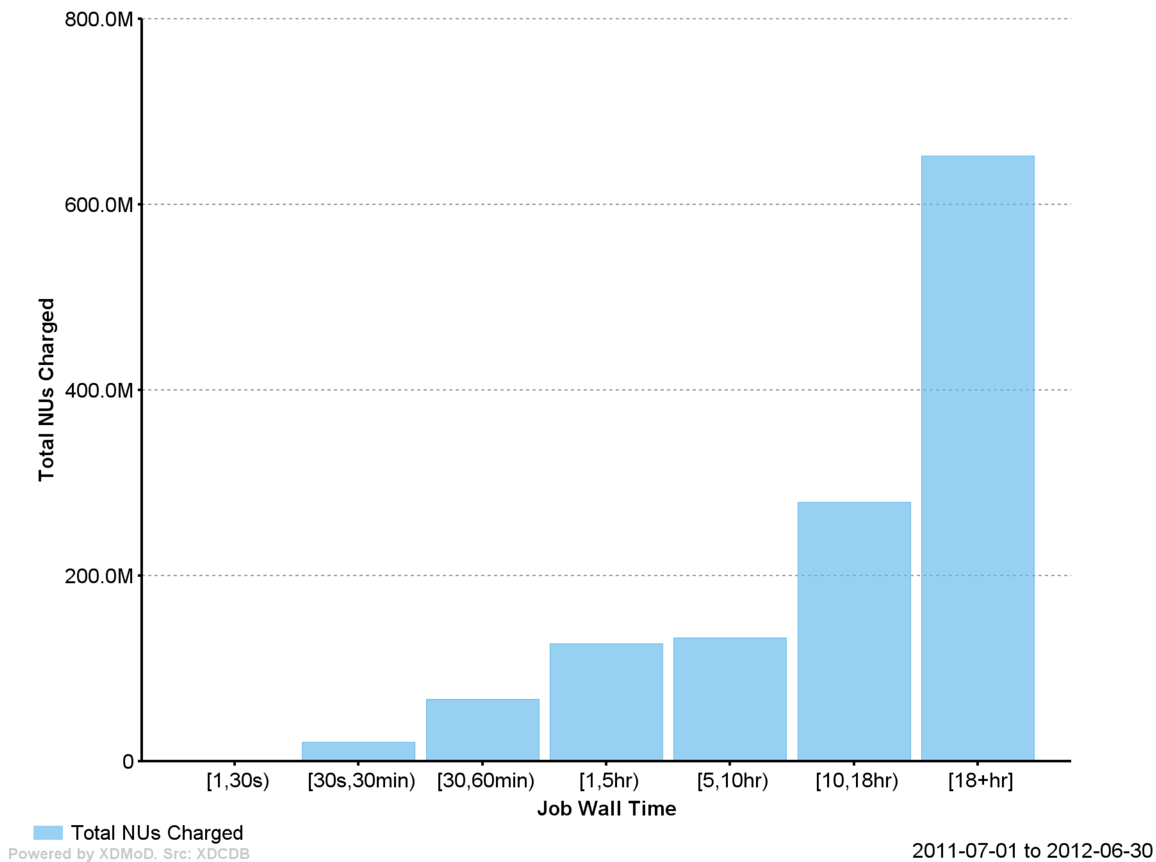
2011-07-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = SDSC-GORDON

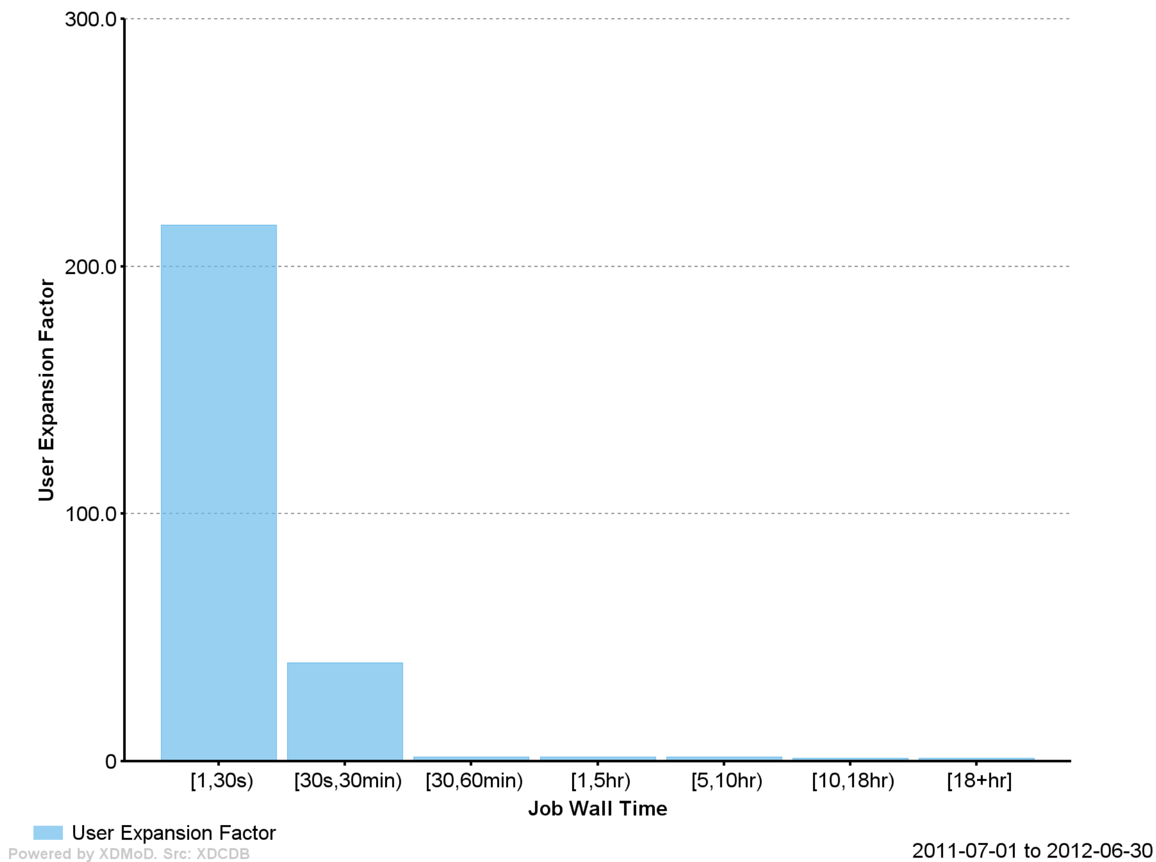
2011-07-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = SDSC-GORDON

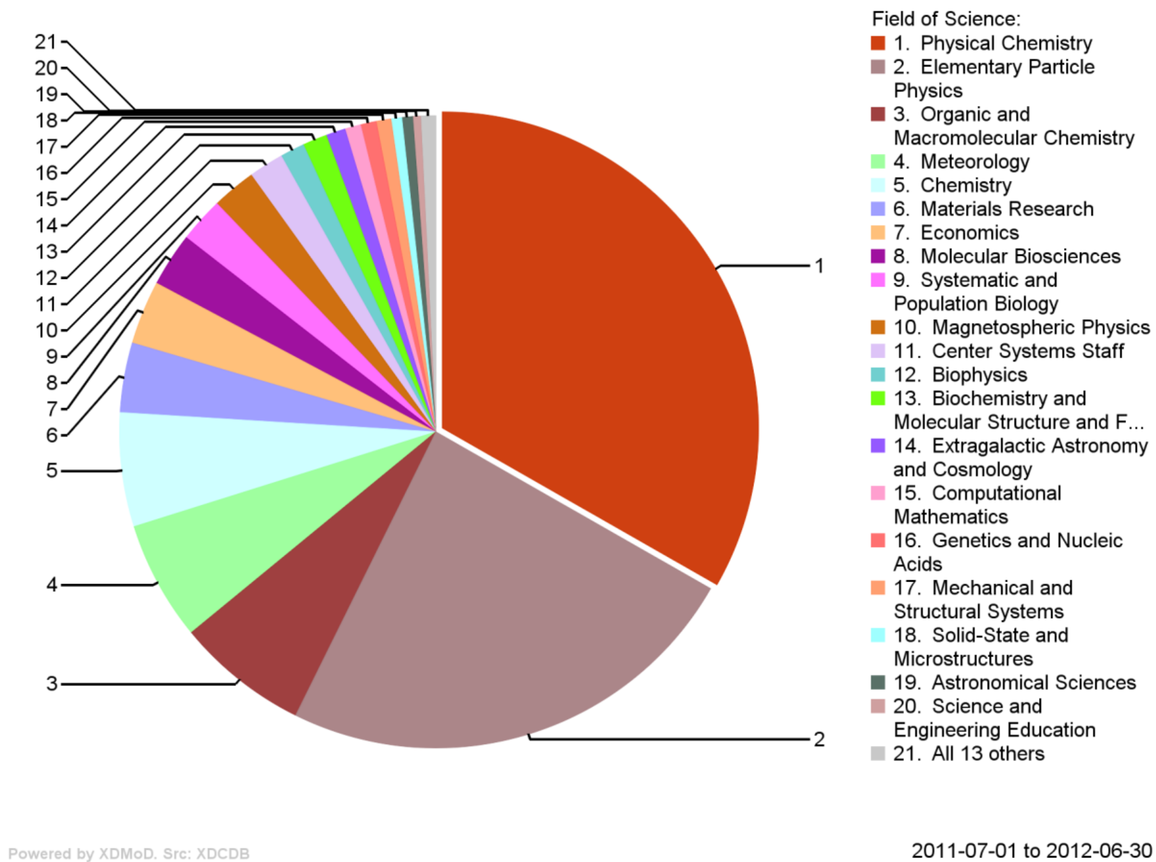
2011-07-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = SDSC-GORDON

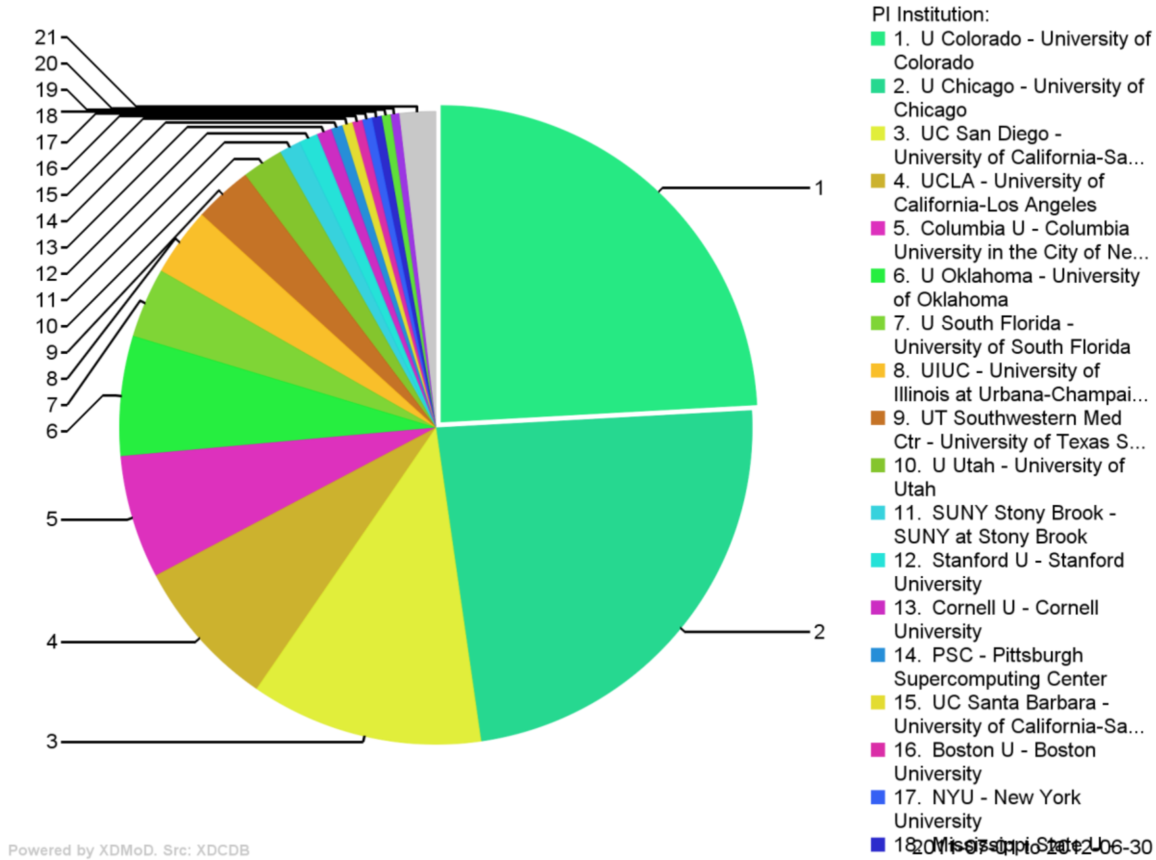
2011-07-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = SDSC-GORDON

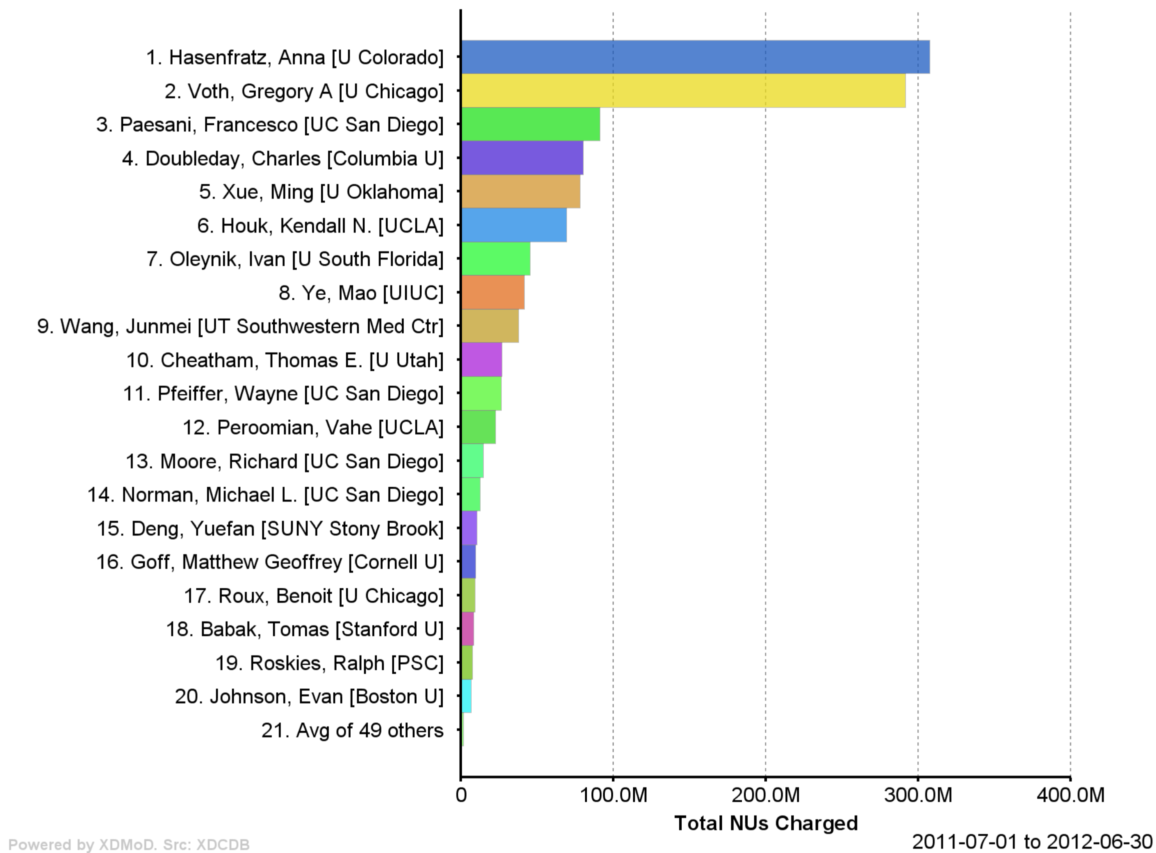
2011-07-01 to 2012-06-30



Total NUs Charged by PI

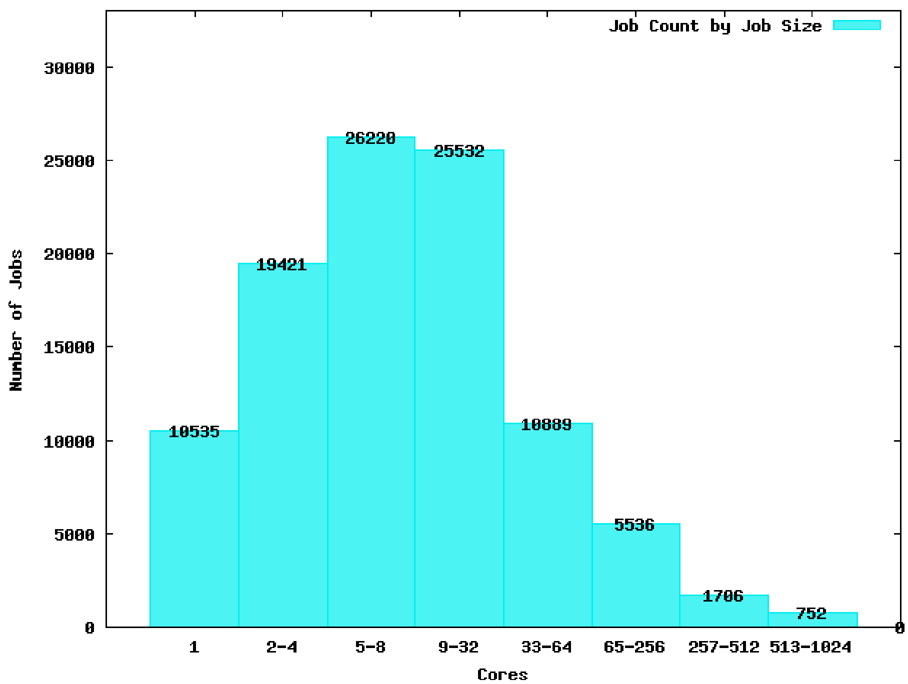
Resource = SDSC-GORDON

2011-07-01 to 2012-06-30

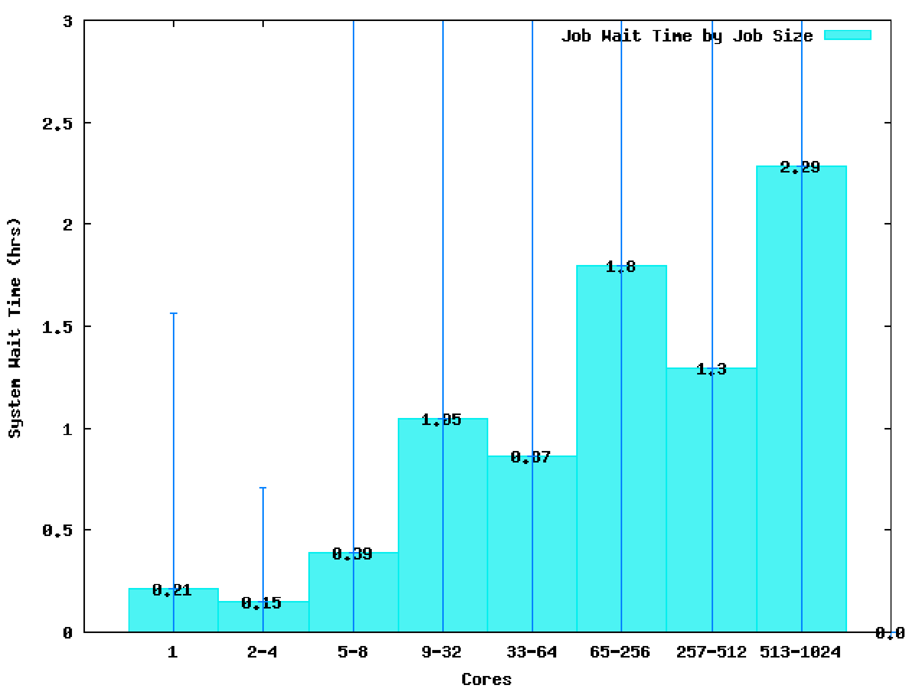


Appendix 1.9D Trestles SP-specific Metrics

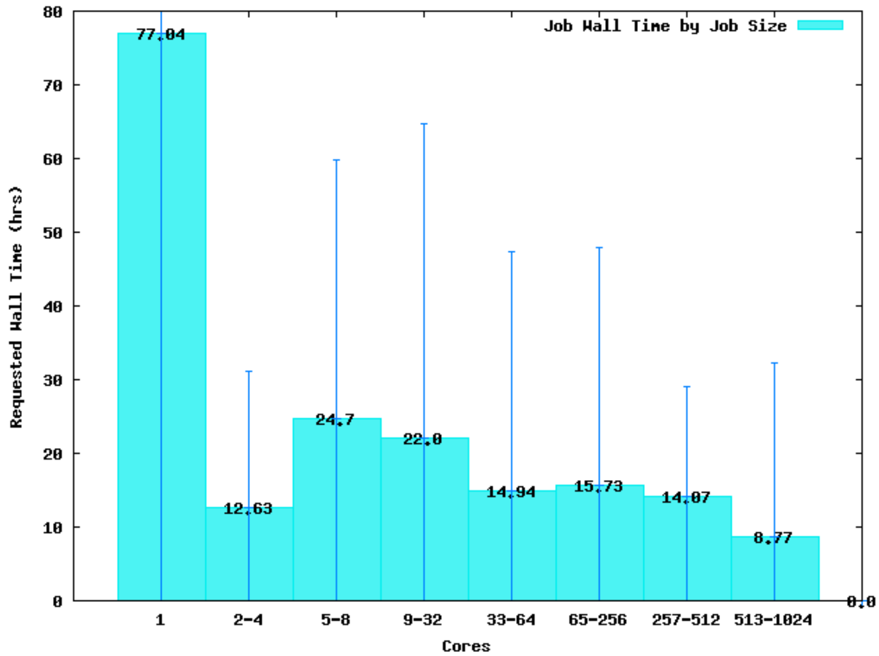
Number of Jobs run by Job Size (Cores)



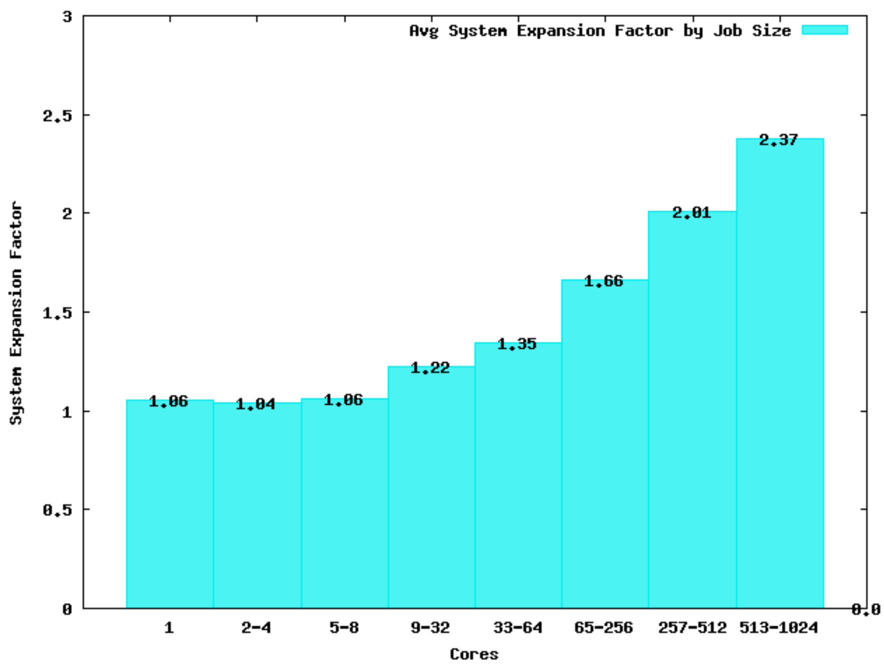
Average System Wait Time (Hours) by Job Size (Cores)



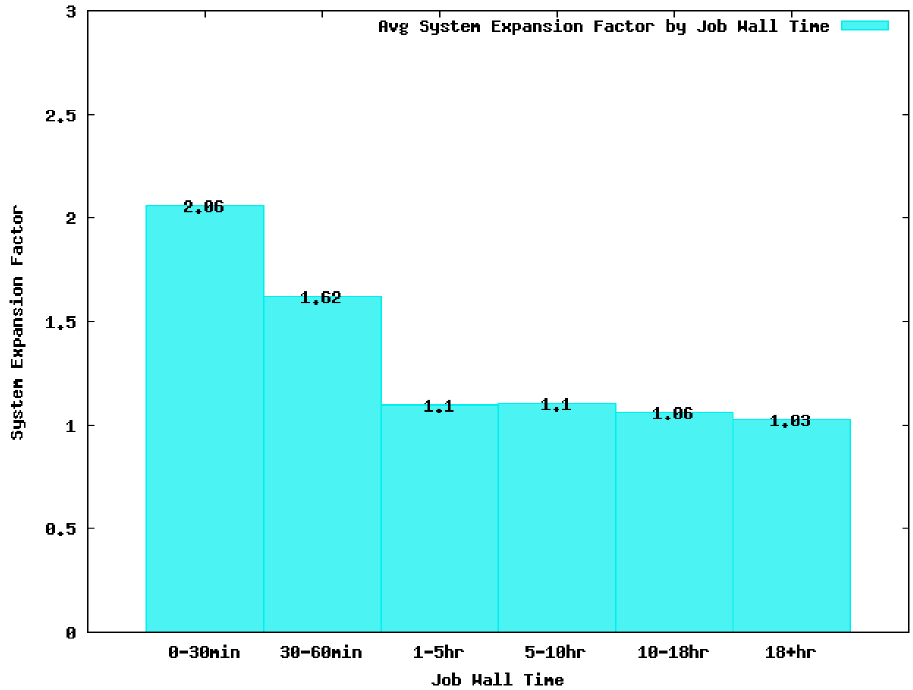
Average Requested Wall Time (Hours) by Job Size (Cores)



Average "Scheduler" Expansion Factor by Job Size (Cores)
(Requested Wall Time + System Wait Time)/Requested Wall Time



Average Scheduler Expansion Factor by Requested Wall Time (Hours)



19 TACC - Service Provider Quarterly Report

19.1 Executive Summaries

Ranger Executive Summary

This reporting period marks the fourth full year of operations of the Ranger resource. This system remains in high demand, and continues to exceed metrics for performance and stability. Within the reporting period the system has been used by 1,920 users to run 478,420 jobs consuming 483.5 million service units (SUs). A total of 1,088 principal investigators used Ranger within the reporting period in support of 1,306 research projects with 88.5% of the SUs being utilized by XSEDE PIs. Researchers from institutions within the state of Texas also continued to use Ranger via the pool of allocation designated for that community. Details of this usage, as well as highlights of some of the scientific achievements will be provided in the annual report (submitted separately). The outstanding reliability of Ranger has resulted in great demand among the open science community. For the current quarter, researchers requested more than 100M SUs on Lonestar.

Overall, the Lonestar project continues to be a success. Users continue to flock to the resource in large numbers, and the effort to continue to expand the system with external funding continues.

Lonestar Executive Summary

This reporting period marks the first full year of operations of the Lonestar-4 resource. This system remains in high demand, and continues to be the most over requested system in the XSEDE ecosystem. During the past year, the system has continued to exceed metrics for performance and stability, delivering 2-4x the performance of Ranger on most user applications. Within the reporting period the system has been used by more than 1,600 users to run nearly 425,000 jobs consuming 184 million service units (SUs). Of this total, more than 46% of delivered cycles went to the XSEDE community, well over the committed 40% of the system promised in the original grant agreement (XSEDE users received nearly 13M more SUs than TACC promised in the period). Details of this usage, as well as highlights of some of the scientific achievements are provide in the annual report. The outstanding performance of Lonestar has put it in great demand among the open science community. For the current quarter, researchers requested more than 100M SUs on Lonestar. Lonestar is the first non-Track 2 system in XSEDE to have more than 100M SUs requested in a quarter, and with requests exceeding available cycles by more than 600%, is the most over requested resource in XSEDE/TeraGrid history.

As detailed in the original proposal, the Lonestar project involved leveraging the funding provided by NSF with funding from other sources to build a much larger and capable system. The proposed 302 TeraFlop system represented the $2.86M project budget, and an additional $6.14M in funds from other partners to deploy the $9M system. Fundraising has exceeded that target, and currently $11.4M has been made available for investment in the system. The original 302TF system has been augmented with additional compute nodes (taking the total to 311TF), large shared memory nodes, and GPU nodes for remote visualization and GPGPU computing to provide a comprehensive environment for scientific computing.

Overall, the Lonestar project continues to be a success. Users continue to flock to the resource in large numbers, and the effort to continue to expand the system with external funding continues.

The third year of the Longhorn XD Vis project has been successful in continuing to recruit new users, developing effective visualizations, developing tools to respond to evolving user needs, and training current and next generation scientists and engineers. Our efforts over the last year have focused on increasing visualization and data analysis usage on Longhorn. Additionally, we have focused on the enhancement and improvement of the user experience by improving the remote visualization capabilities, particularly through the Longhorn Visualization Portal. Longhorn usage has continued to increase with a diverse portfolio of users wanting Visualization and Data Analysis. There are currently 423 persons running jobs from 607 projects with 513 PIs burning an average of 600K SUs over the third year.

More details can be found in the Longhorn annual report (to be submitted separately.)

19.1.1 *Resource Description*

19.1.1.1 Sun Constellation Linux Cluster (Ranger)

The TACC Sun Constellation Cluster contains 62,976 cores (2.3 GHz) within 3,936 Sun Constellation blades (nodes), an X4600 Rocks master node, 4 X4600 user login nodes, 4 X4600 user gridftp nodes, 4 X4600 data movers, 2 X4600 nodes dedicated to supporting the SGE batch system, 2 X4600 external management service nodes, 2 X4600 InfiniBand subnet management nodes, an X4100 software build node, and 6 X4600 metadata server nodes to support the Lustre parallel file systems. Multiple work and home file systems are configured from 1.7 PB of storage managed by the Lustre parallel file system management software. Two Sun Data Center 3456 switches are the core of an InfiniBand fabric through which all components are connected. The basic configuration is as follows:

- 3936 Sun Constellation Blade Servers, each with
    - four quad-core 2.0 GHz processors
    - 32 GB of Memory
    - 8 GB flash drive
- 1.7 PB of storage managed by the Lustre Parallel File System software
- InfiniBand Interconnect

19.1.1.2 Dell Westmere Linux Cluster (Lonestar)[2]

The TACC Dell Westmere Cluster contains 22,656 compute cores (3.33 GHz) within 1,888 Dell PowerEdge M610 compute blades (nodes), 15 PowerEdge R610 compute-I/O server-nodes, and 2 PowerEdge M160 login/management nodes. Each compute node has 24 GB of memory, and the login/development nodes each have 24 GB. 14 large memory (1TB) nodes are available for high-throughput computing and applications that require access to a shared-memory architecture and 8 GPU nodes are configured for visualization and applications that can take advantage of the computational speed of the GPUs. The system storage includes a 421 TB parallel WORK Lustre file system, a 841 TB parallel SCRATCH Lustre file system, and 275 TB of local compute-node disk space (146GB/node). A QDR InfiniBand switch fabric interconnects the nodes (I/O and compute) via a fat-tree topology, with a point-to-point bandwidth of 40Gb/sec. The basic configuration is as follows.

[2] Began production on February 1, 2011.

- 1888 Dell PowerEdge M610 Blade Servers, each with
  - Dual Intel Westmere 6-core, 3.33 GHz processors
  - 24 GB of Memory
  - 146 GB of Local Disk
- 14 Dell PowerEdge R910 servers, each with
  - Four 6-core, 2.0 GHz Intel Xeon processors
  - 1 TB of Memory
  - 292 GB of Local Disk
- 8 Dell PowerEdge C6100 servers, each with
  - Two NVIDIA M2070 GPUs
  - Two 6-core, Intel Xeon X5670 2.93 GHz processors
  - 24 GB of Memory
  - 146 GB of Local Disk
  - 16-lane PCI Express to Dell C410x PCI expansion box housing the NVIDIA GPUs
- 421 TB Lustre Parallel File System (WORK)
- 841 TB Lustre Parallel File System (SCRATCH)
- QDR InfiniBand Interconnect

19.1.1.3 DELL/NVIDIA Visualization and Data Analysis Cluster (Longhorn)

The TACC DELL/NVIDIA Visualization & Data Analysis Cluster, Longhorn, is a hybrid CPU/GPU system designed for remote, interactive visualization and data analysis. In addition, Longhorn supports production, compute-intensive calculations on both the CPUs and GPUs via off-hour queues. The large, per-node memory is intended to support serial and parallel visualization and analysis applications that take advantage of large memories, multiple computing cores, and multiple graphics processors. Longhorn is an ideal companion resource for working with large data sets created on Ranger, since Longhorn can directly access Ranger's Lustre parallel file system through a 10 GigE network link.

The system consists of 256 dual-socket nodes, each with significant computing and graphics capability. Total system resources include 2048 compute cores (Nehalem quad-core), 512 GPUs (128 NVIDIA Quadro Plex S4s, each containing 4 NVIDIA FX 5800s), 13.5 TB of distributed memory and a 210 TB global file system. Longhorn configuration details can be found below.

128 NVIDIA Quadro Plex S4s, each with

- 4 NVIDIA FX 5800 GPUs
- 16GB Graphics Memory (4GB per GPU)
- 2 independent graphics busses, one per GPU pair

240 Dell R610 Compute Nodes, each with

- 2 Intel Nehalem quad-core processors (8 cores) @ 2.53 GHz
- 48GB RAM
- 73GB local disk
- connected to 2 dedicated NVIDIA FX 5800 GPUs via Quadro Plex graphics bus

16 Dell R710 Compute Nodes, each with

- 2 Intel Nehalem quad-core processors (8 cores) @ 2.53 GHz
- 144GB RAM
- 73GB local disk
- connected to 2 dedicated NVIDIA FX 5800 GPUs via Quadro Plex graphics bus

Mellanox QDR InfiniBand Interconnect

14 Dell PowerVault MD1000 Direct Attached Storage Arrays (210TB global file system, managed by the Lustre Parallel File System)

19.1.1.4 Terascale Sun Visualization Cluster (Spur)

TACC's Terascale Sun Visualization Cluster contains 128 compute cores, 1 TB aggregate memory and 32 GPUs. Spur acts not only as a powerful stand-alone visualization system: it also enables researchers to perform visualization tasks on Ranger-produced data without migrating to another file system and to integrate simulations and rendering tasks on a single network fabric. The cluster consists of the following hardware:

- 1 Sun Fire X4600 server with 2 NVIDIA Quadro Plex model 4. The X4600 contains 8 dual-core CPUs (16 cores total) and 256GB of RAM. Each Quadro Plex model 4 contains 2 NVIDIA Quadro FX5600 GPUs;
- 1 Sun X4400 servers, with 4 quad-core CPUs (16 cores total) and 128GB of RAM, connected to 2 NVIDIA Quadro Plex model 4. Each Quadro Plex model 4 contains 2 NVIDIA Quadro FX 5600 GPUs;
- 6 Sun X4400 servers, each with 4 quad-core CPUs (16 cores total) and 128GB of RAM, and each connected to an NVIDIA Quadro Plex S4. Each Quadro Plex S4 contains 4 NVIDIA Quadro FX 5600 GPUs; and
- Total system capability: 128 cores, 1TB aggregate memory, 32 GPUs.

Because Spur shares Ranger's interconnect fabric and file systems, researchers will be able to easily transition between HPC runs to generate and visualize data. Furthermore, visualization software is able to harness both the rendering power of the graphics hardware and the compute power of Ranger to enable the analysis of terascale and larger data sets.

19.2 Science Highlights

Provided in annual report.

19.3 User-facing Activities

Provided in annual report.

19.3.1 <u>System Activities</u>

Provided in annual report.

19.3.2 *Services Activities*

Provided in annual report.

19.4 Security

There are no changes in security procedures or security incidents to report within this reporting period.

19.5 Education, Outreach, and Training Activities

19.5.1 *Outreach*

XSEDE activity in QY2012 continued with participation in monthly All-TEOS meetings led by Scott Lathrop (NCSA), bi-weekly Outreach meetings led by Laura McGinnis (PSC), and bi-weekly Underrepresented Outreach meetings led by Linda Akli (SURA). The meetings continued to provide an opportunity for information and clarification of activities across TEOS and a chance to identify opportunities for collaboration within the existing planned activities.

The first cohort of the XSEDE Student Engagement Program had 16 participants. PSC sponsored 6 students with 4 working remotely. NICS sponsored 4 students, with 3 working remotely. TACC sponsored 6 students with 1 working at SDSC, 3 working at TACC and 2 working remotely. TACC based participants included graduate students from the University of Texas, Clemson University, and University of Colorado. Students attended orientation that was held synchronously via conference call at TACC, NICS, and PSC on May 30 and 31. Work on their projects remotely or on-site began the first week of June, with various starting dates based on the students' schedules. The students worked towards presenting their works-in-progress at XSEDE12.

Underrepresented outreach continued with a highly productive assortment of networking activities at the National Science Foundation Joint Administrative Meeting (JAM), June 12-15 in Washington, D.C. for Principal Investigators in the EHR Directorate. Conference participants included PIs from ADVANCE (Advancement of Women), RDE (Research in Disabilities Education), HBCU (Historically Black Colleges and Universities), MSI (Minority Serving Institutions), TCUP (Tribal College and Universities Program), and LSAMP (Louis Stokes Alliance for Minority Participation). Presentations about XSEDE's cyberinfrastructure and user services were made to the PIs for HBCU/TCUP and LSAMPs and were received well, resulting in numerous contacts for XSEDE training and collaborations.

From the JAM contacts, Xavier University in New Orleans expressed a strong interest in working with XSEDE to establish a computational biology program through its computer science department due to the large number of students who come to Xavier for pre-med. UTEP expressed an interest in collaborating on an NSF AGEP for visualization. The Minorities Striving and Pursuing Higher Degrees of Success (MS PHD'S) in Earth Systems Science program is seeking to engage in a collaboration to offer scientific computation training for their participants. Elizabeth City State University, Tennessee State University, and Virginia State University were interested in visualization and parallel programming training, with Virginia State making an immediate follow up to SURA. Lastly, the Southern Regional Educational Board (SREB) invited Samuel Moore and Linda Akli to attend the invitation only Conference on Mentoring and Teaching in Tampa, Florida in October to network with the 1200 invited minority PhDs about XSEDE cyberinfrastructure and user services.

The campus champion in the works for Austin Community College (ACC) a two-year college in Austin, Texas was delayed until further discussion in Fall 2012. A May meeting with the ACC Computer Science department chair included suggesting the idea of Lorraine Brill as one of two Campus Champions, one for outreach and one for curriculum. The curriculum Champion will

collaborate with TACC to develop a scientific computing certificate for ACC. Steve Gordon at OSC has offered to assist in the development of the certificate through the XSEDE Education group. Dr. Brill would serve as a liaison between TACC and ACC to increase the awareness of and participation in scientific computation by ACC students.

The Austin Forum on Science, Technology & Society

In this quarter, TACC hosted a total of 3 monthly Austin Forum events with invited speakers from areas of interest focused on science and technology. The goal of The Austin Forum on Science, Technology & Society is to engage and educate the local community about the numerous ways in which science and technology enhance the quality of their everyday life, as well as the health, prosperity and security of the nation. One hour is devoted to a presentation and Q&A discussion between the speaker and guests. Ample time for networking is offered, both preceding and following the speaker presentation. A total of 418 people attended The Austin Forum.

TACC Facility Tours

From K-12 and higher education groups, TACC conducted facility tours impacting 366 people, 40% of whom were under-represented. An overview of XSEDE and TACC were given at each event.

Additionally, this quarter, TACC led tours for diverse audiences, including participants from two unique programs:

1) Shadow a Scientist Program - Coordinated by the College of Natural Sciences, this summer program paired a group of two Middle School students at a time with Research Scientists at TACC for a two-hour tour of scientists' experiments in progress. This provided the opportunity for young students to cultivate their interest in science and allowed them to experience first-hand what a scientist does on a daily basis.

2) TEDGlobal- TACC hosted TEDxAustin's live simulcast of the entire second day of TEDGlobal 2012 taking place in Edinburgh, Scotland. Participants toured the Vislab and saw the simulcast on "Stallion" (307 megapixel resolution).

The following table lists TACC outreach activities during the reporting period.

Type	Title	Location	Date(s)	Number of Participants	Number of Under-represented people
The Austin Forum on Science, Technology and Society	Product Design: The New Interplay of People, Objects and Information" w/Dean Dillon, Dr. Diane Bailey, and Dr. Randolph Bias	AT&T Conference Center	4/3/12	183	Not tracked.
Machine Room Tour	VIP: Dell	PRC	4/4/12	4	4
Machine Room Tour	VIP: Dell4	PRC	4/9/12	5	1
Machine Room Tour	SSC students	PRC	4/24/12	17	5
Vislab Tour	VIP: Dell Peru	Vislab	4/25/12	2	2
Vislab Tour	ischool undergraduate students	Vislab	4/26/12	18	7
Vislab Tour	eForum conference participants	Vislab	4/26/12	50	Not tracked.
Vislab Tour	VIP: United Technologies	Vislab	4/30/12	4	2
Vislab Tour	UT SEEK	Vislab	5/1/12	40	Not tracked.
The Austin Forum on Science, Technology and Society	"Using Technology to Refine Physical Education in the 21st Century" w/ Jen Ohlson	AT&T Conference Center	5/3/12	85	Not tracked.
Vislab Tour	VIP: Dell- Electromagnetic Norway	Vislab	5/3/12	6	Not tracked.
Machine Room Tour	Prairiland High School	PRC	5/22/12	6	1
The Austin Forum on Science, Technology and Society	"The Smart Grid Unveiled: How Pecan Street's Big Data Research is Uncovering the Promise of a Customer-Driven Energy Revolution" w/Brewster McCracken	AT&T Conference Center	6/5/12	150	Not tracked.

Vislab Tour	UT: EOE MITE	Vislab	6/7/12	49	49
Machine Room Tour	Shadow a Scientist- Doug James	PRC	6/13/12	2	0
Machine Room Tour	Shadow a Scientist- Ritu Arora	PRC	6/13/12	2	0
Vislab Tour	UT: EOE MITE	Vislab	6/14/12	50	50
Vislab Tour	UT DDCE	Vislab	6/26/12	9	9
Vislab	TEDxGlobal Simulcast	Vislab	6/27/12	22	Not tracked.
Vislab	Skillpoint Alliance Gameon!	Vislab	6/28/12	40	40
Machine Room Tour	Shadow a Scientist- Ritu Arora	PRC	6/29/12	40	0

19.5.2 Education

Scientific Computing Curriculum and Courses

TACC's Spring 2012 courses continued in Q2 with students expressing interest in additional training courses offered at TACC. The addition of the Scientific and Technical Computing to the Spring course offerings, making it available Fall and Spring will be revaluated due to the demands on TACC Research Associates and Scientists who teach the courses. Spring enrollment totals were 36 graduate students and 45 undergraduates. TACC staff taught: Introduction to Scientific Programming (4 graduate students and 27 undergraduates), Parallel Computing for Science & Engineering (16 graduates and 11 undergraduates) and Scientific and Technical Computing (16 graduates and 7 undergraduates). Students were assigned accounts on *Ranger*. In addition, the GPU Programming class was offered in the long summer session, with an enrollment of 16 students. The class is being revamped for the future to account for the advancements in GPUs and GPU programming. This summer included content on GPGPUs.

The courses are offered in the Flawn Academic Center in a customized classroom housing both lecture space and a computer instruction laboratory. The classroom customization was made possible through a partnership with Chevron to increase instruction in scientific computing. Fall 2012 is targeted for a pilot webcasting of instruction from the lab to a higher education institution in the state of Texas without scientific computation course offerings.

The partnership with the UT Austin Division of Statistics and Scientific Computation (DSSC)

entered its third year and continued to be marketed in the Scientific Computation courses. DSSC offers an Undergraduate Certificate and a Graduate Portfolio Program in Scientific Computation. Four of the scientific computing classes taught by TACC fulfill requirements for the certificate and portfolio. Documentation of students completing the certificate program appears on their transcript as a notation. Five students have completed the program to date. As of Spring 2012, 36 undergraduates and 10 graduate students are enrolled in the program. DSSC's scientific computing course descriptions are online at:

http://www.tacc.utexas.edu/education/academic-courses

19.5.3 Training

TACC conducted 9 training workshops and staff members at the Cornell Center for Advanced Computing conducted one training workshop related to the Ranger project during the reporting period. A total of 640 students attended the workshops either in person or via webcast. The following table lists the date, title, location, and attendance for each workshop.

Date	Class	Attendance	Underrepresented Community	Location
4/9/2012	Visualization Workshop	20	9	Clark Atlantic University
4/19/2012	XSEDE New User Training	57	19	TACC, webcast
4/23/2012	HPC Python Tutorial	116	20	TACC, webcast
5/15/2012	Visualization Workshop	21	12	North Carolina University (SURA)
5/16-17/2012	Parallel Computing Workshop on Ranger and Lonestar	40	11	Cornell
5/17/2012	HPCToolkit and PerfExpert Tutorials	25	8	TACC, webcast
June 2012	XSEDE Scholars – C Programming	20	17	TACC, webcast
6/14/2012	Bring your own code workshop	22	2	M.D. Anderson Cancer Center
6/26/2012	Data Storage – Architectures and Networking	28	9	TACC, webcast

The Virtual Workshop provides users access to twenty-two training modules with new modules under development and existing modules being reviewed for updates. Users who are logged in to the XSEDE portal can pass-through to the Virtual Workshop, or they can use guest registration.

Available Modules
An Introduction to Linux An Introduction to C Programming An Introduction to Fortran Programming Python on Ranger and Lonestar Balancing Scripts and Compiled Code in Scientific Applications MATLAB Programming Parallel Programming Concepts and High-Performance Computing Ranger Environment Message Passing Interface (MPI) MPI Point-to-Point Communications MPI Collective Communications MPI One-Sided Communication MPI Advanced Topics OpenMP Hybrid Programming with OpenMP and MPI Profiling and Debugging Optimization and Scalability Series – Part 1: Planning for Parallel Computational Steering Large Data Visualization ParaView VisIt Using Databases
Modules Currently in Work

Allocations   Data Transfers
Multi-node Map Reduce   Parallel I/O   PerfExpert   R   Transition to Stampede

Table: Virtual Workshop Usage

	Page Loads	Unique Visitors	First Time Visitors	Returning Vi
Q1 '11	4,456	920	730	
Q2 '11	16,281	2,988	2,509	
July – Sept 2011	9,208	2,905	2,457	
Oct – Dec 2011	10,068	3,615	3,019	
Jan – Mar 2012	16,800	5,318	4,249	
Apr – Jun 2012	17,875	5,860	4,795	

Note: the Q2 '11 numbers were a result of high activity after an online news release on the Virtual Workshop was sent out.

19.6 SP Collaborations

Provided in annual report.

19.7 SP-Specific Activities

Provided in annual report.

19.8 Publications

Potluri S., Wang H., Bureddy D., Singh A., Rosales C., Panda D "Optimizing MPI Communications on Multi-GPU Systems using CUDA Inter-process Communication" *The*

Second International Workshop on Accelerators and Hybrid Exascale Systems (AsHES), IPDPS

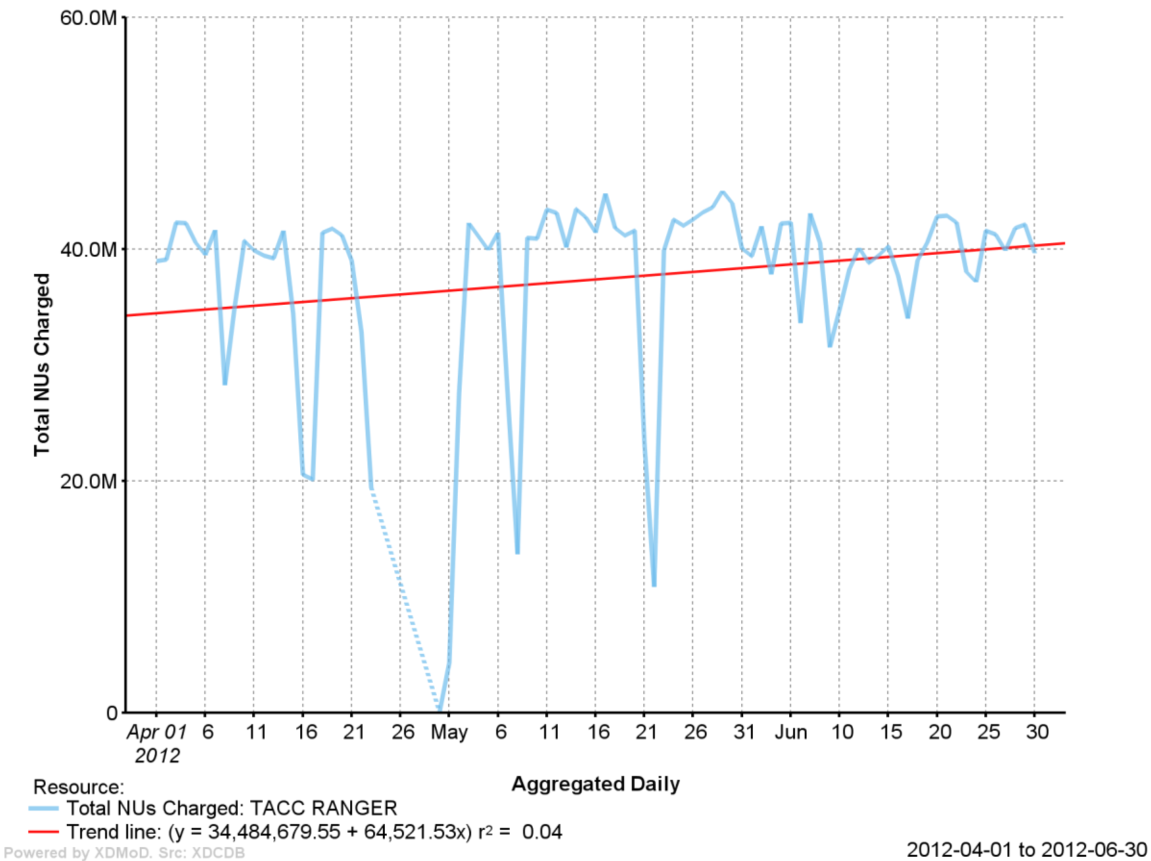
19.9 Metrics

19.9.1 Standard Systems Metrics

Total NUs Charged by Resource

Resource = TACC-RANGER

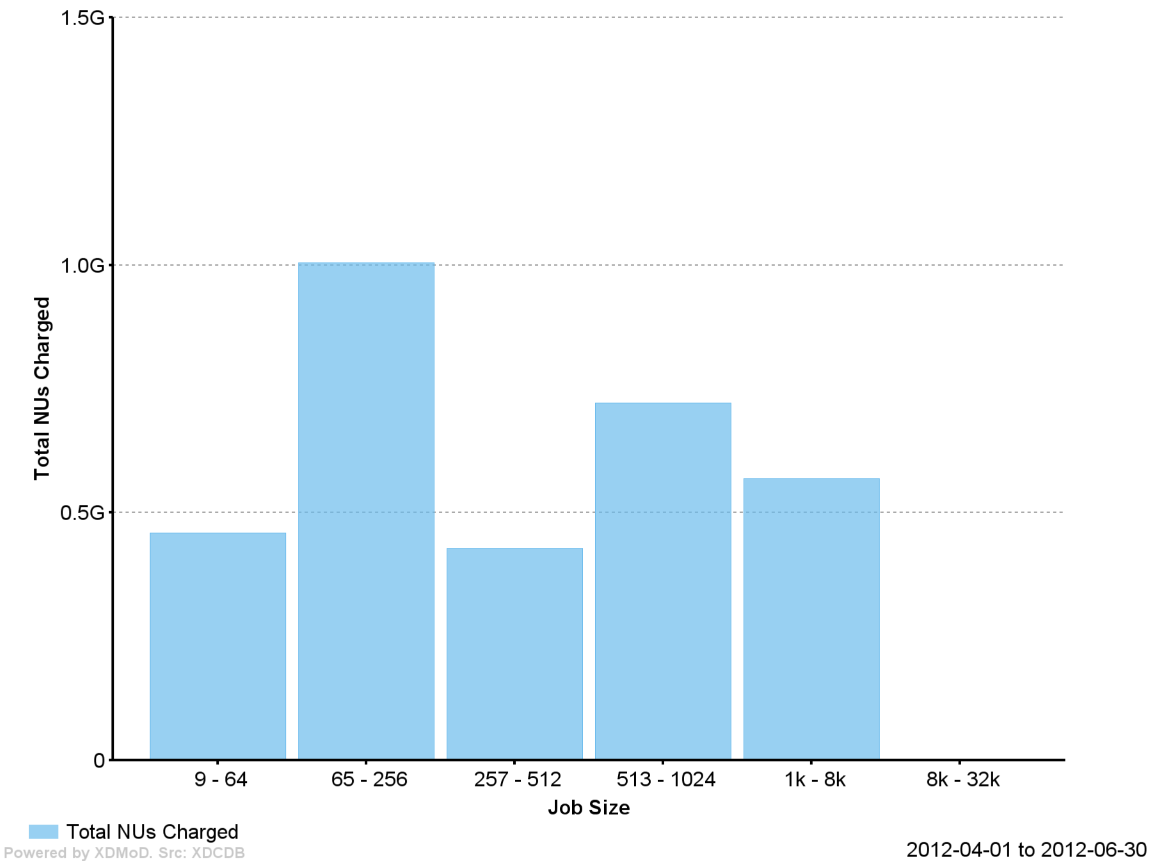
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = TACC-RANGER

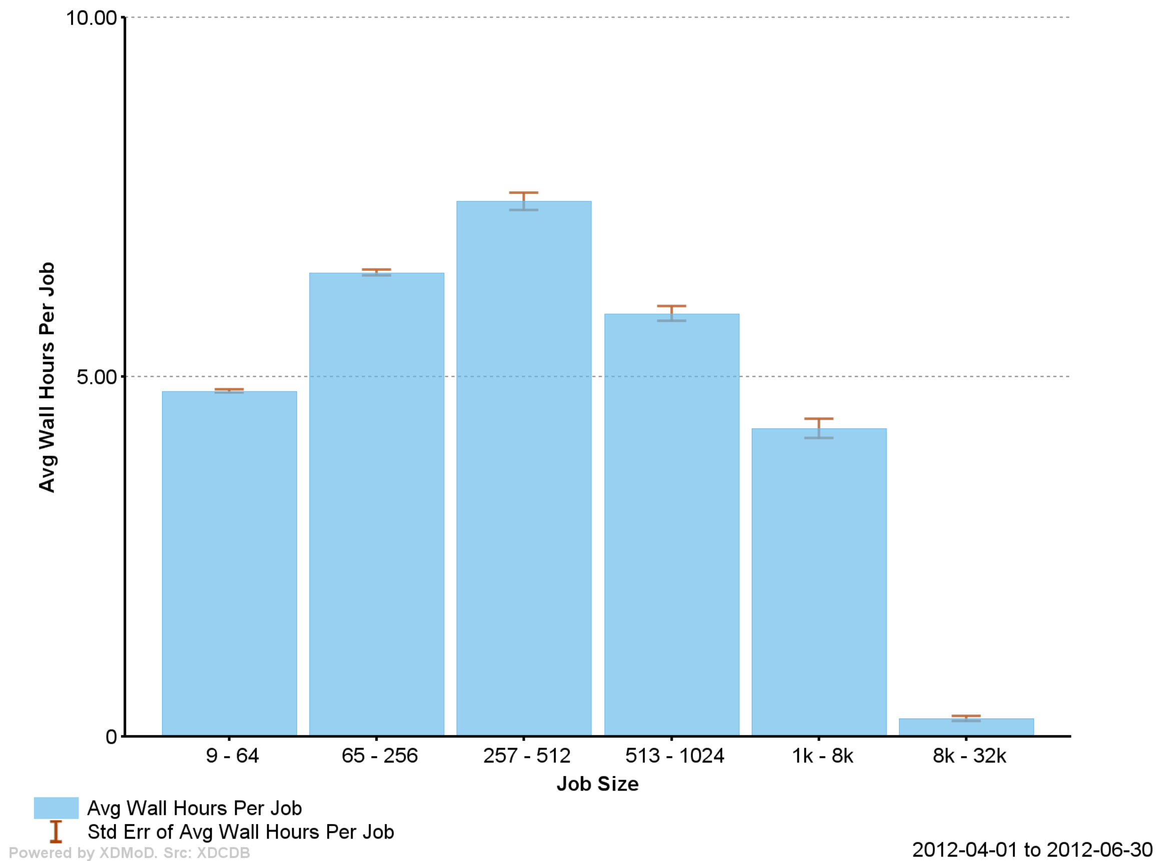
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = TACC-RANGER

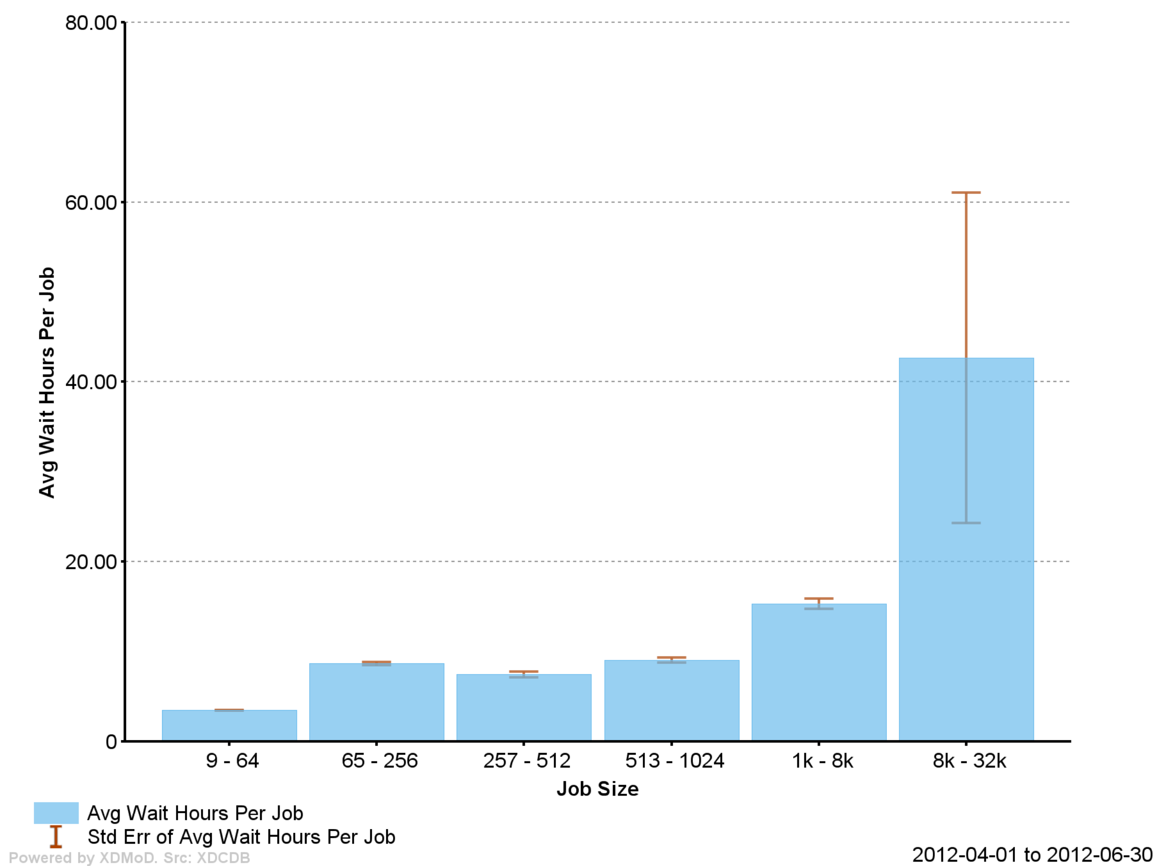
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = TACC-RANGER

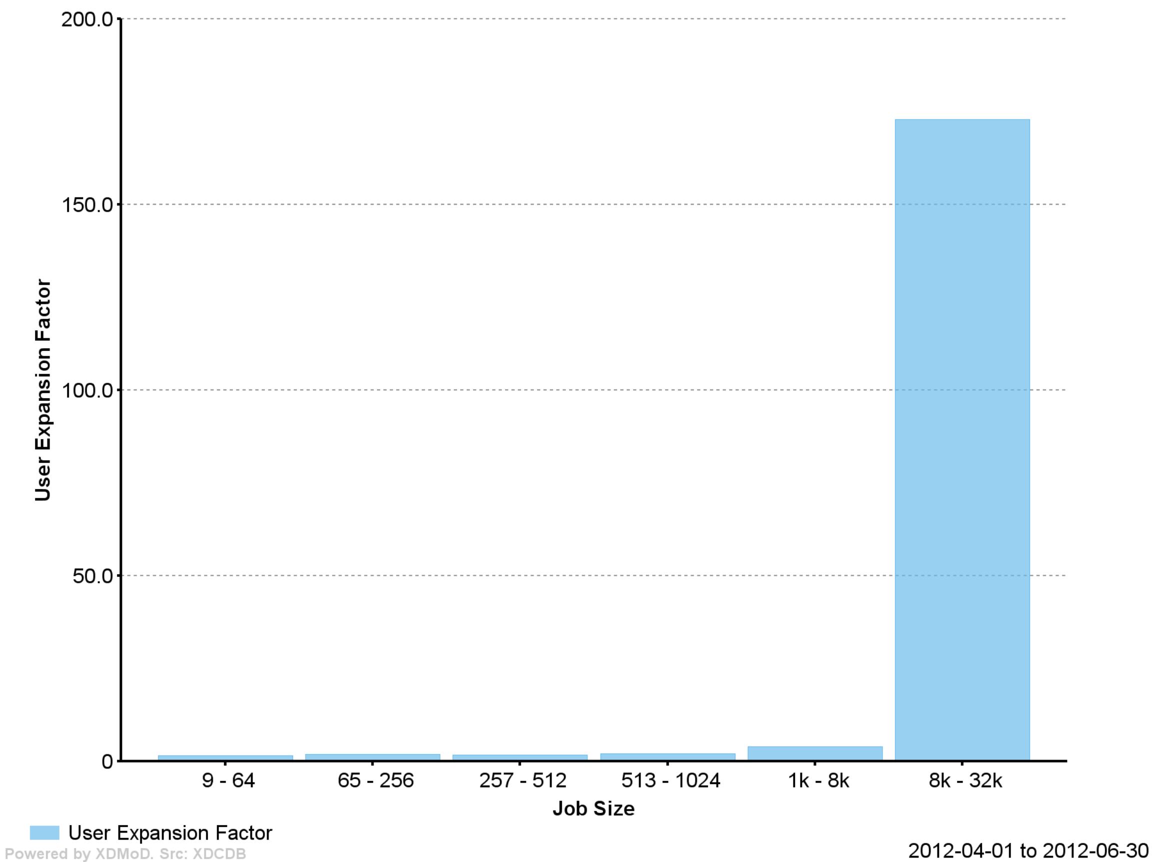
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = TACC-RANGER

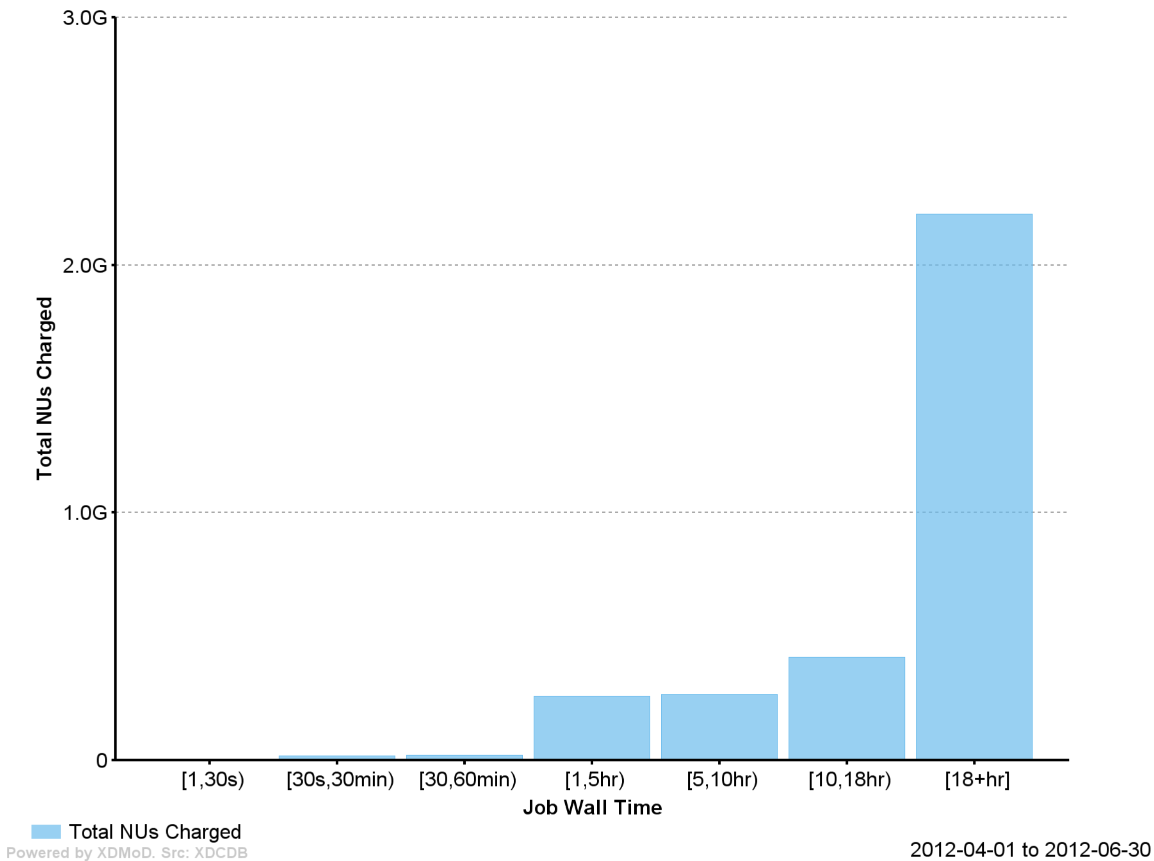
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = TACC-RANGER

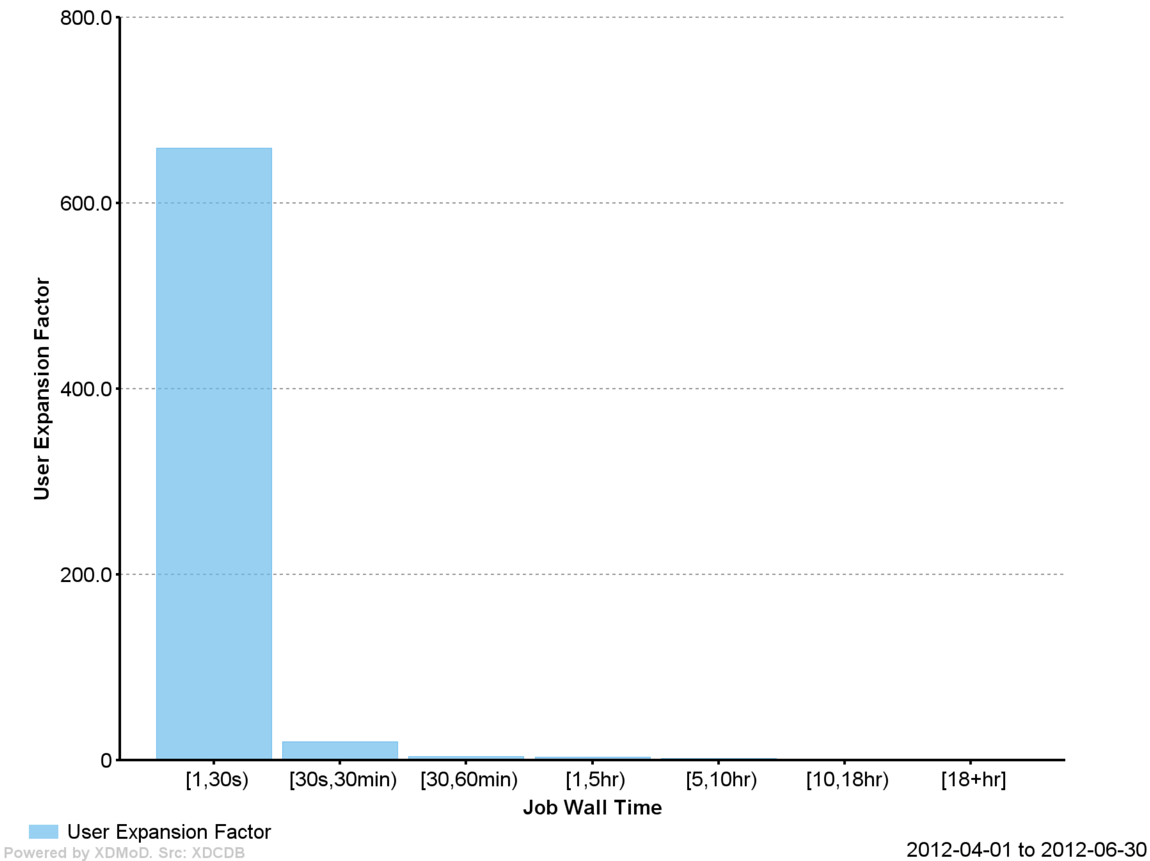
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = TACC-RANGER

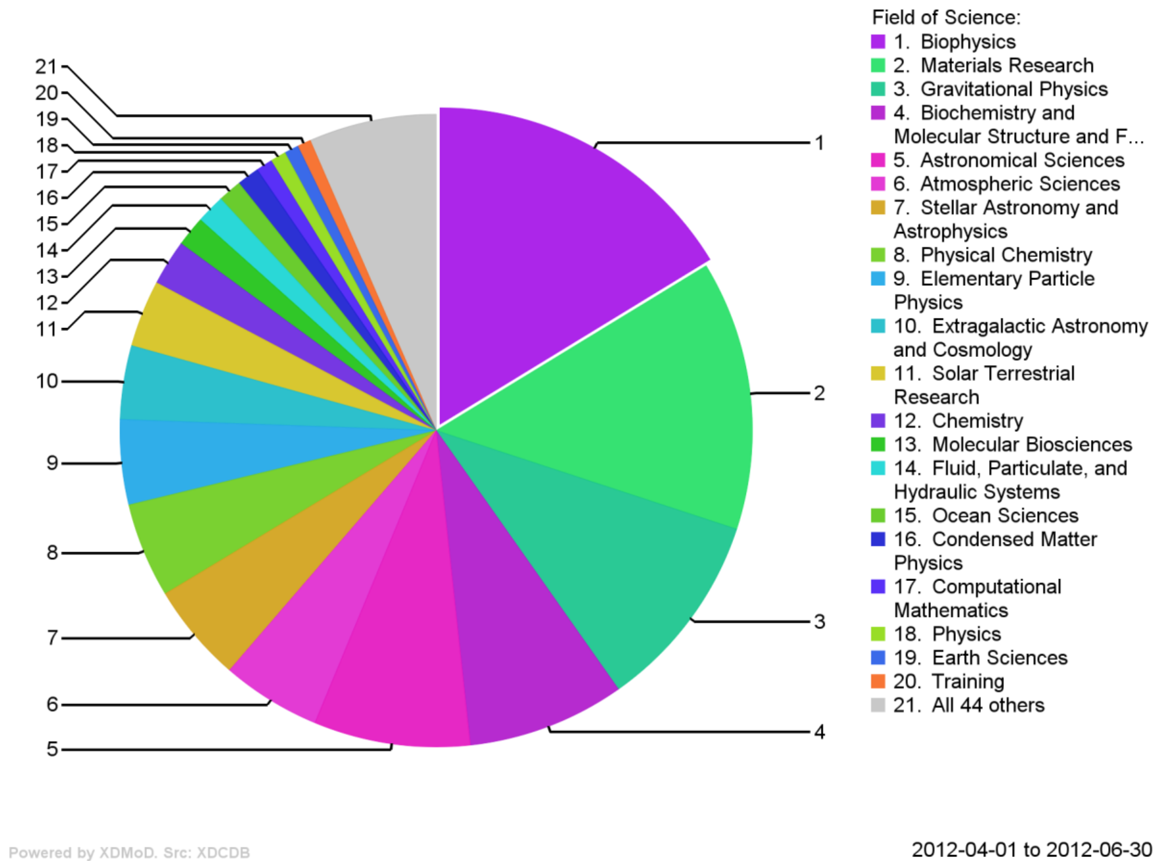
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = TACC-RANGER

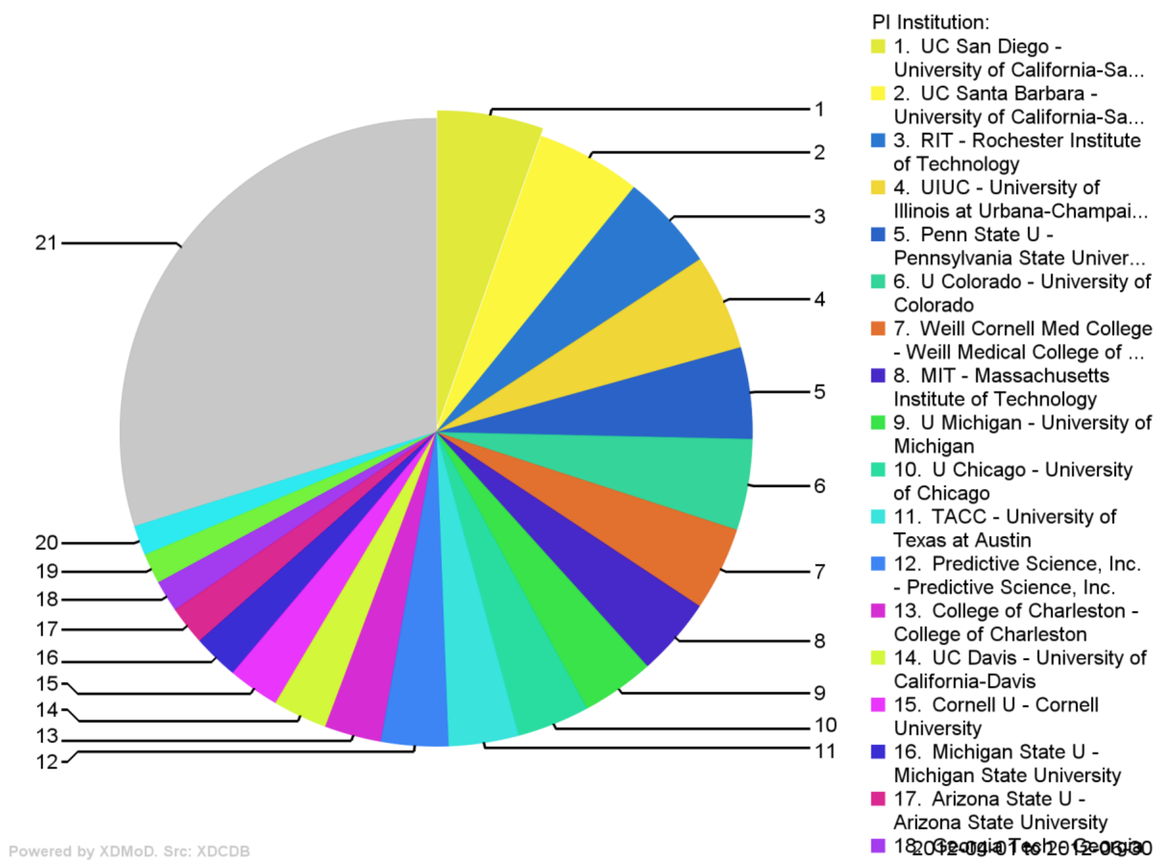
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = TACC-RANGER

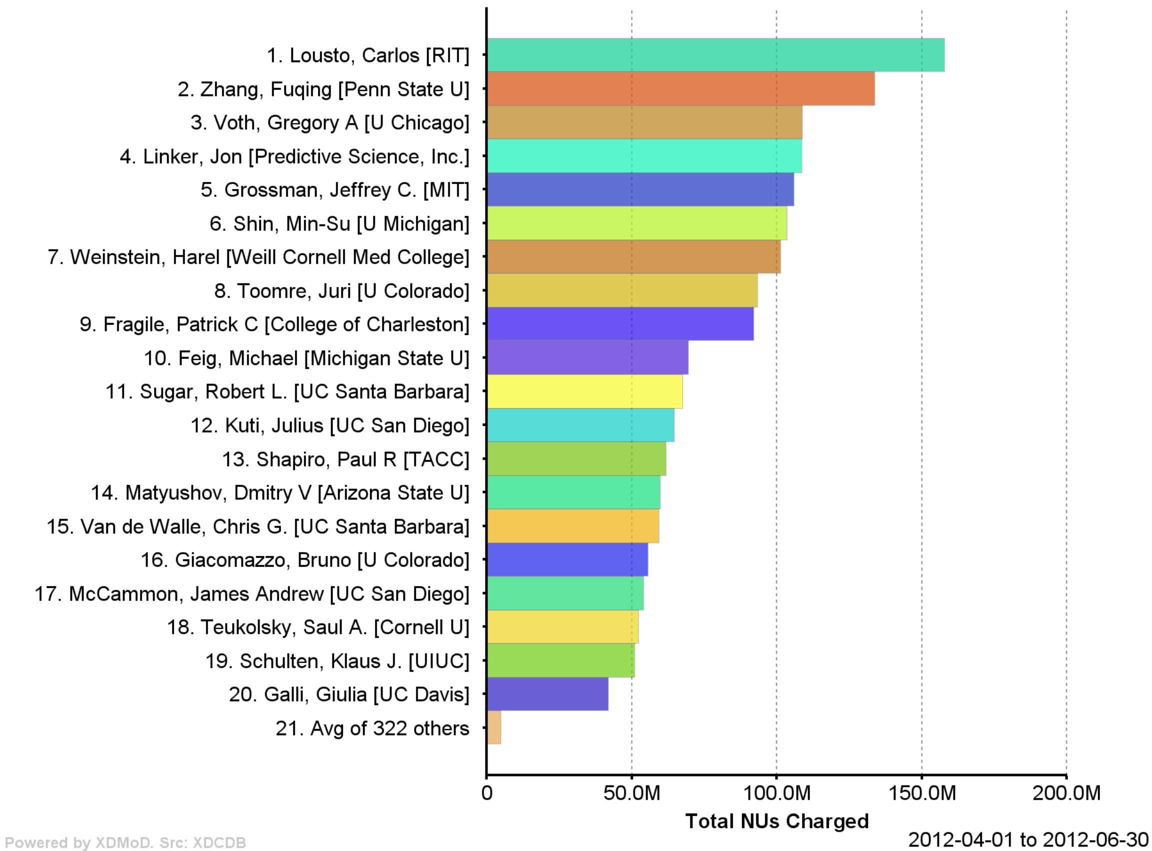
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = TACC-RANGER

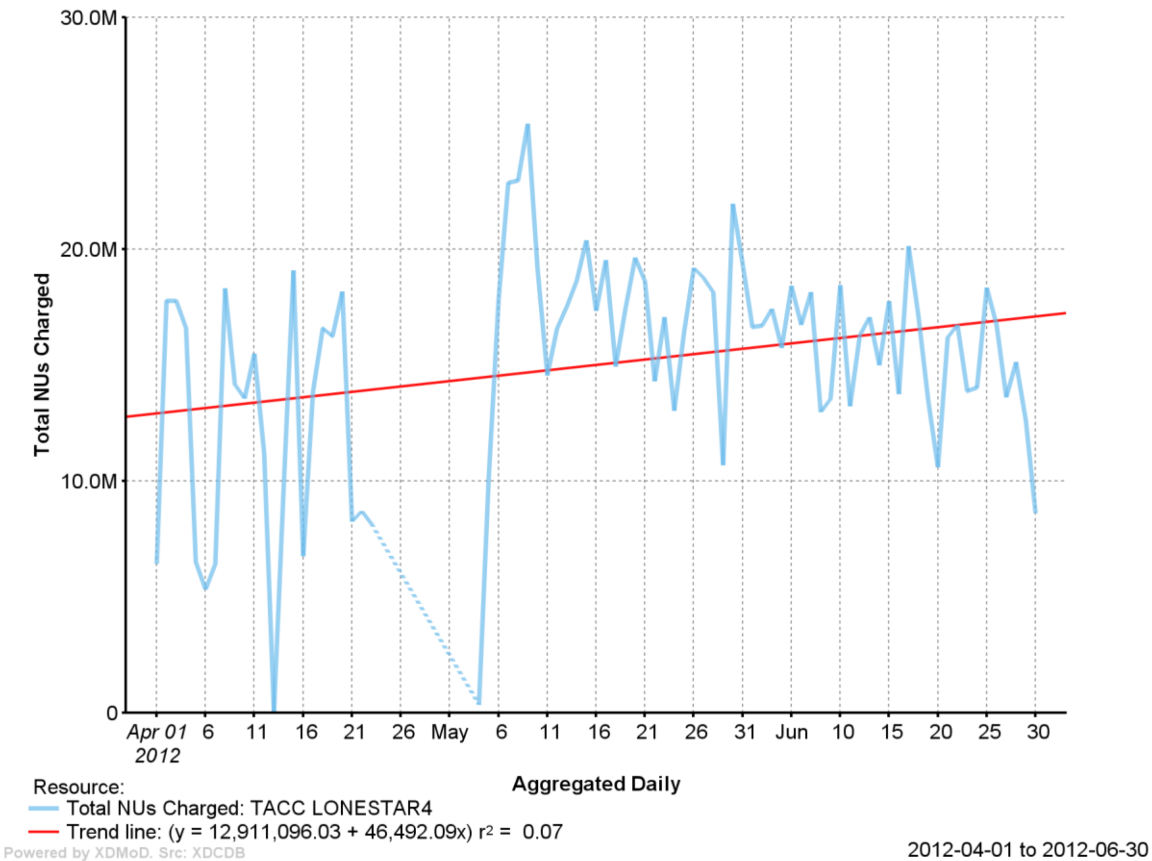
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = TACC-LONESTAR4

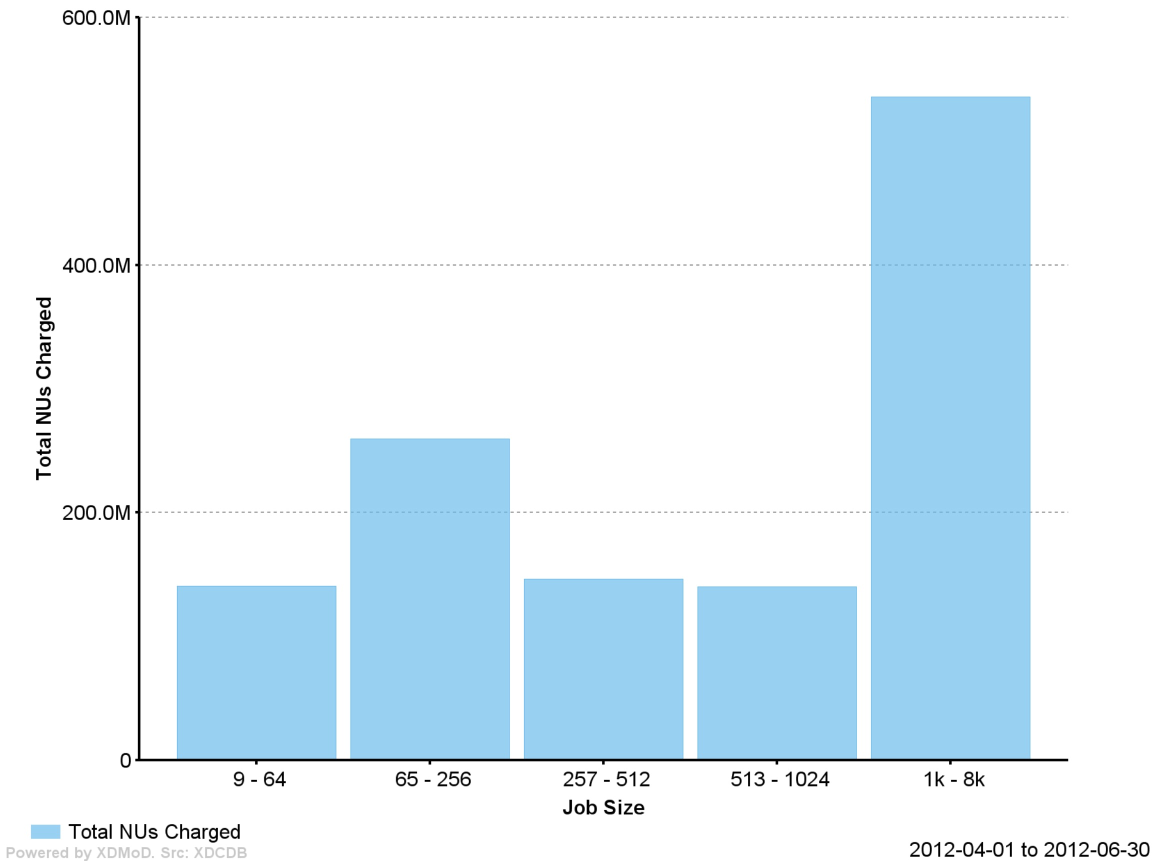
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = TACC-LONESTAR4

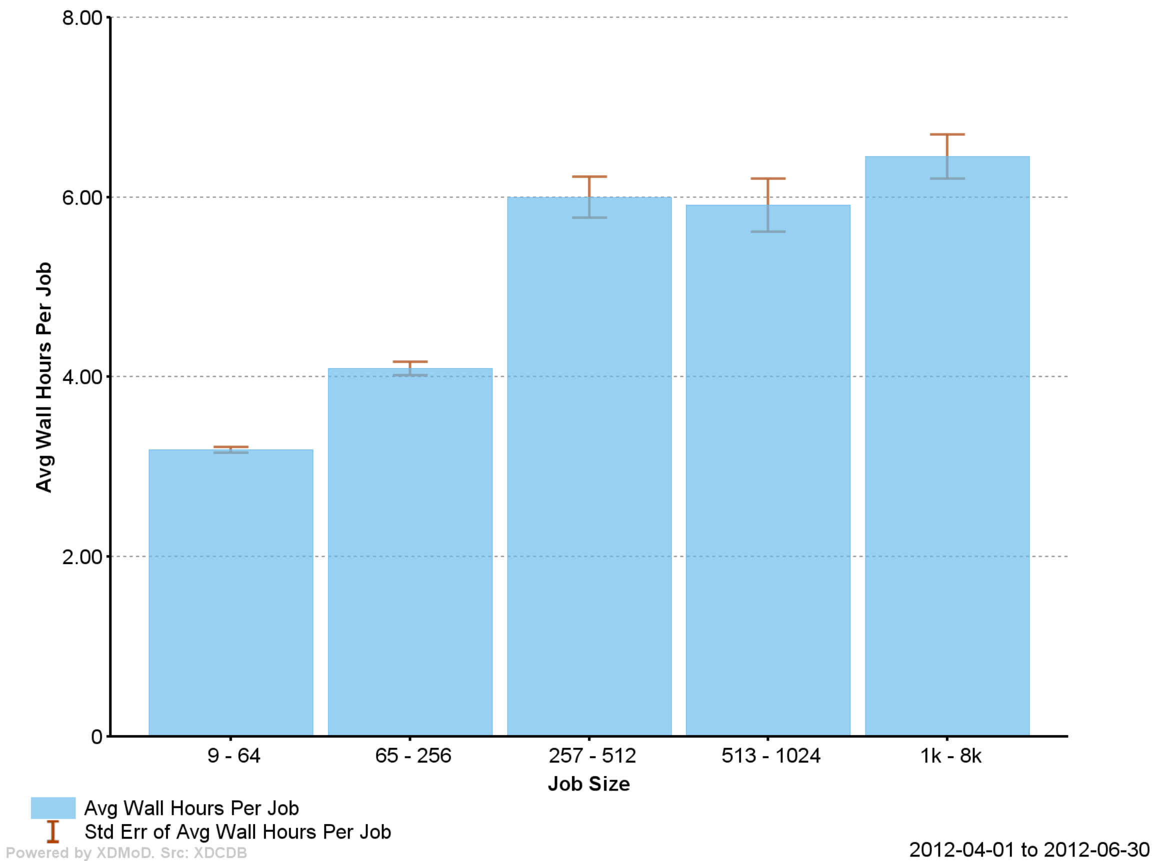
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = TACC-LONESTAR4

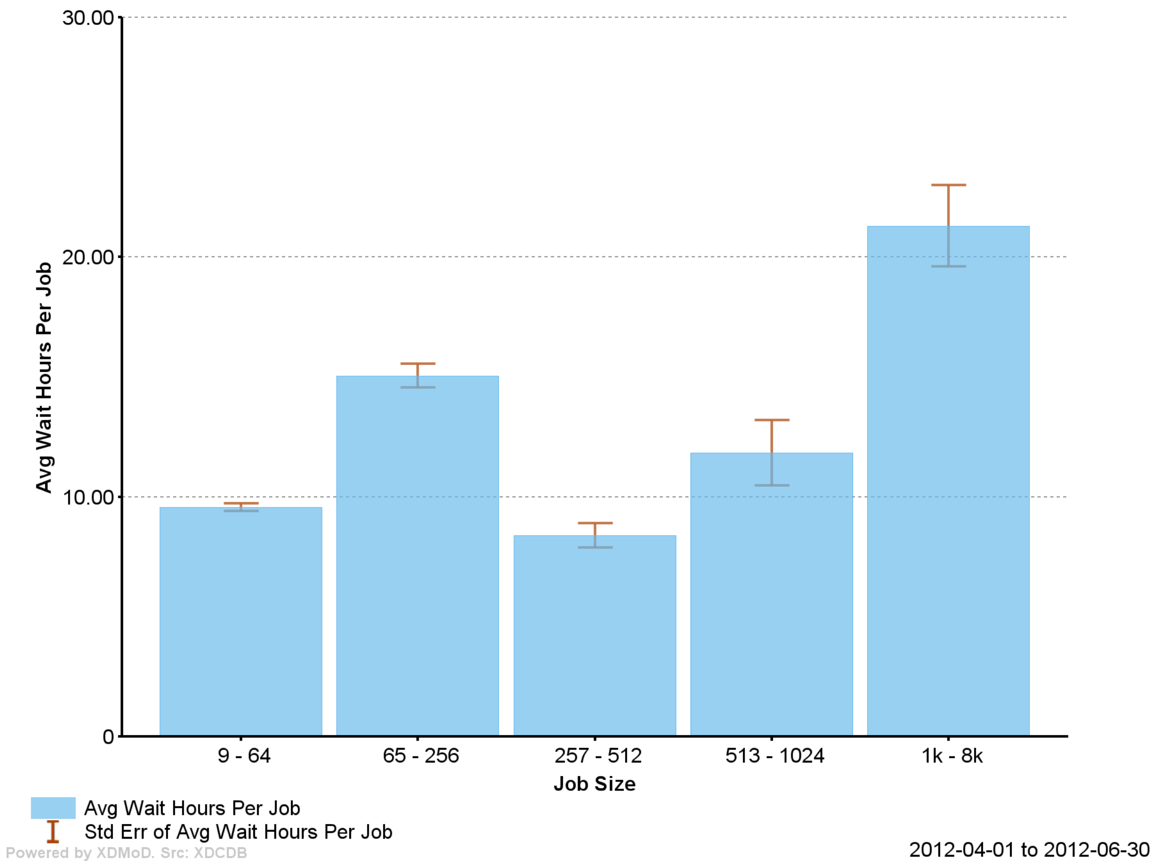
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = TACC-LONESTAR4

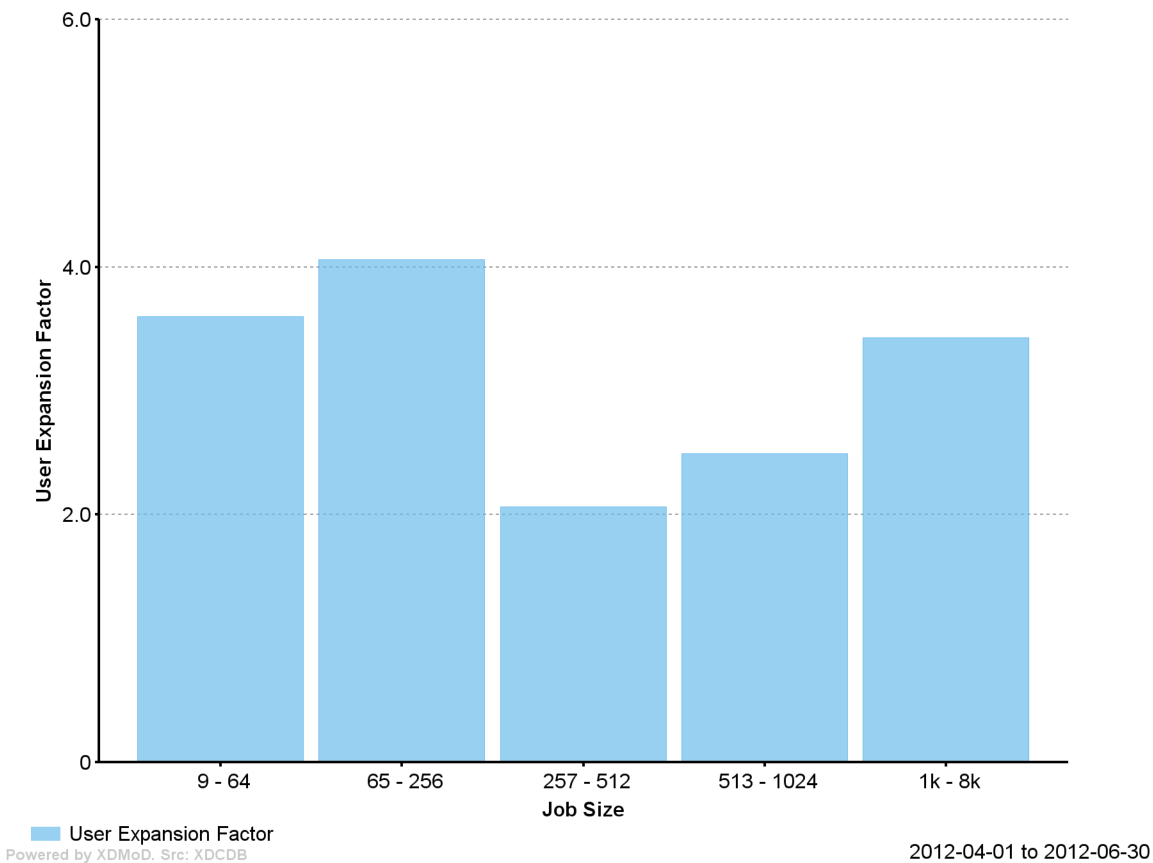
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = TACC-LONESTAR4

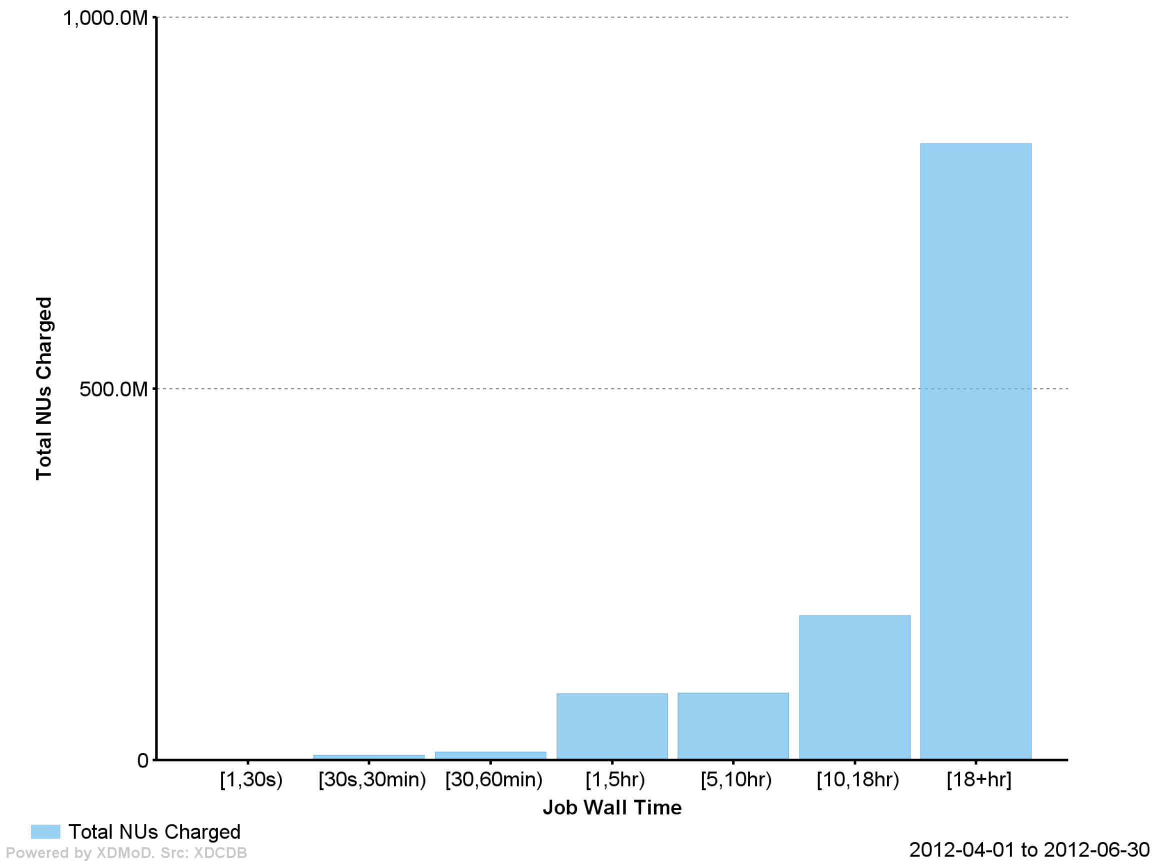
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = TACC-LONESTAR4

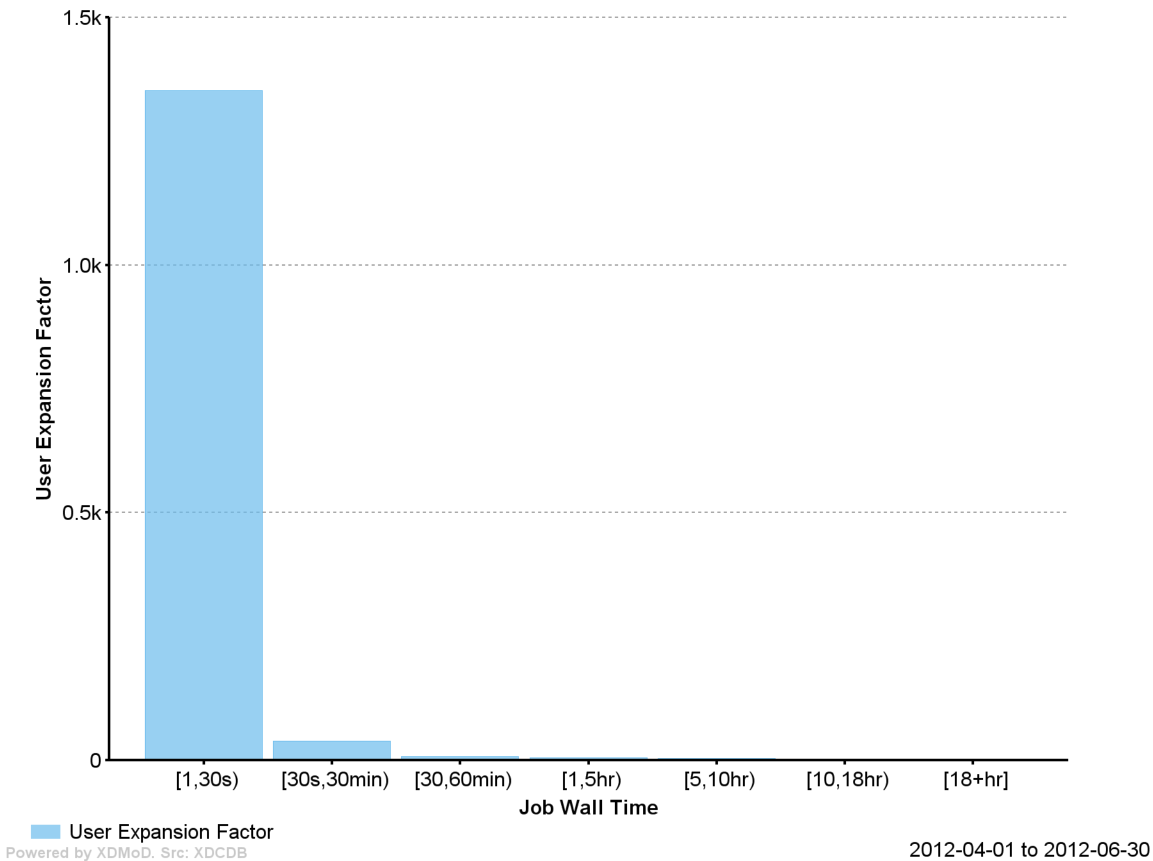
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = TACC-LONESTAR4

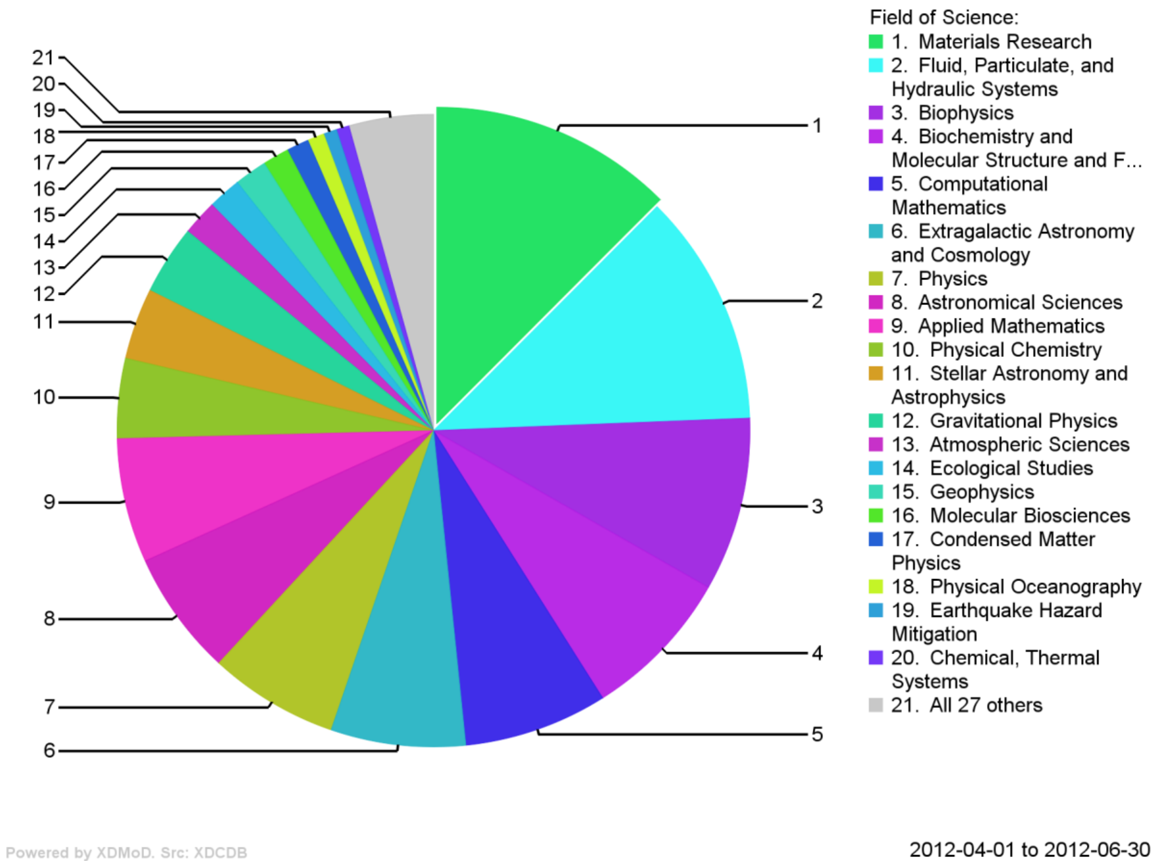
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = TACC-LONESTAR4

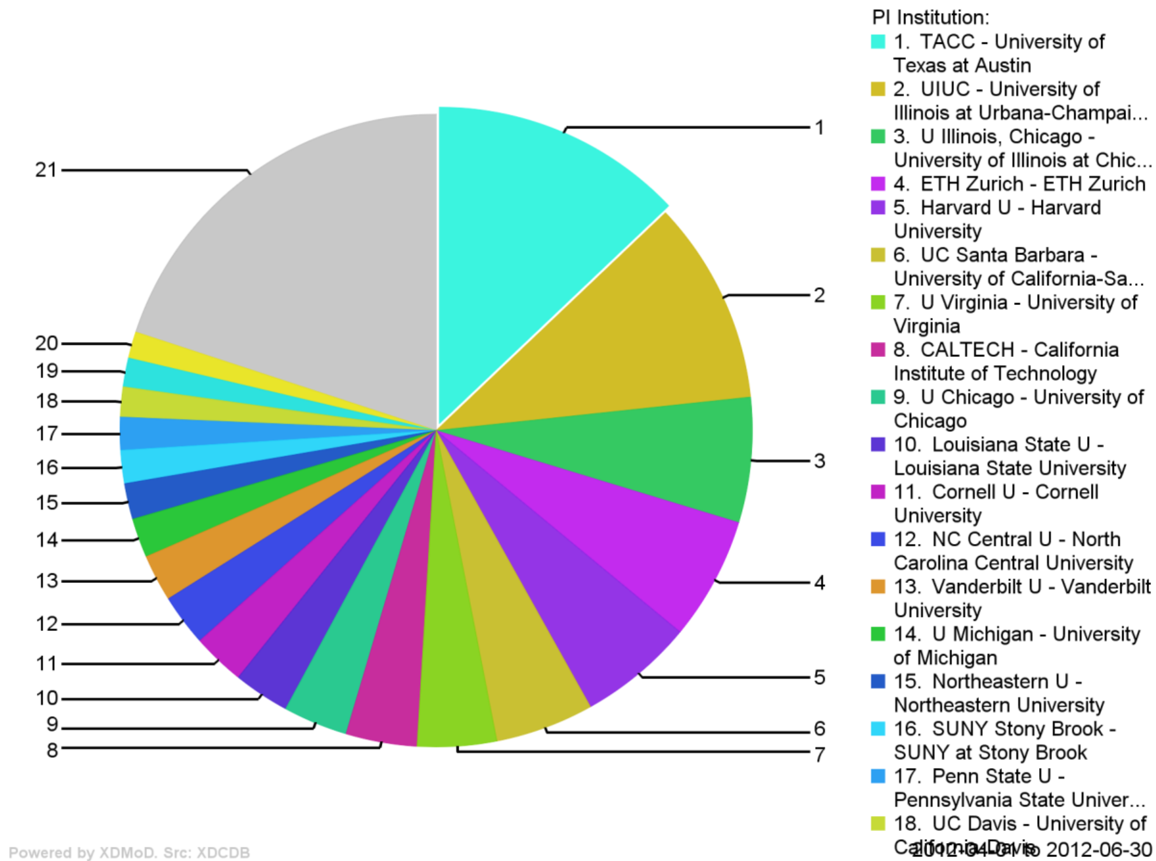
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = TACC-LONESTAR4

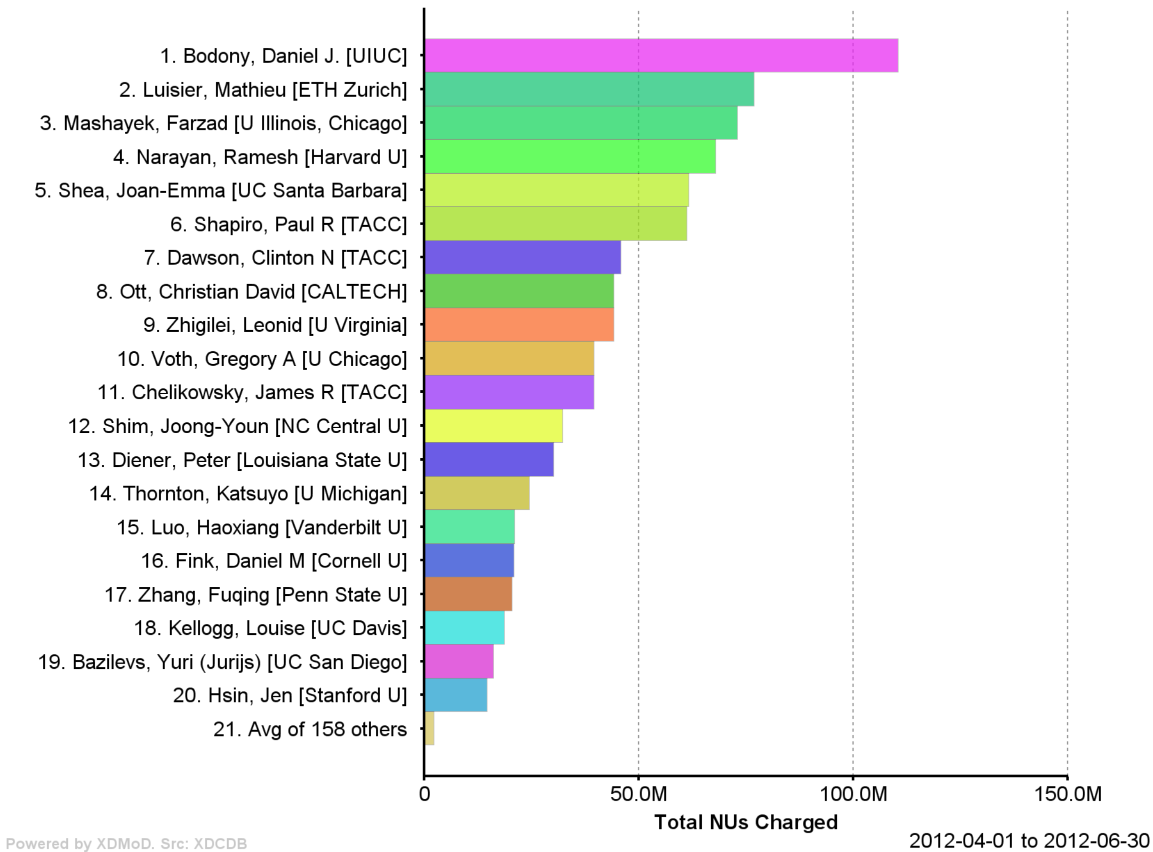
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = TACC-LONESTAR4

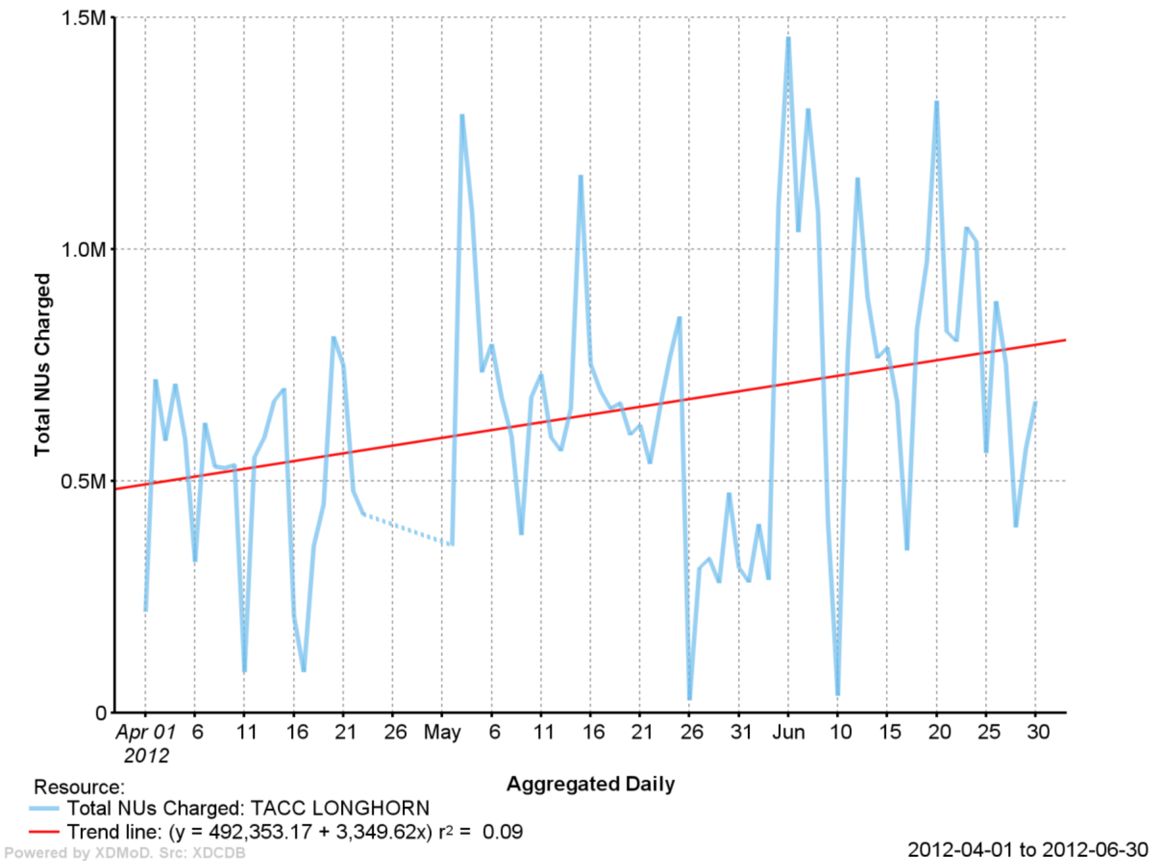
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = TACC-LONGHORN

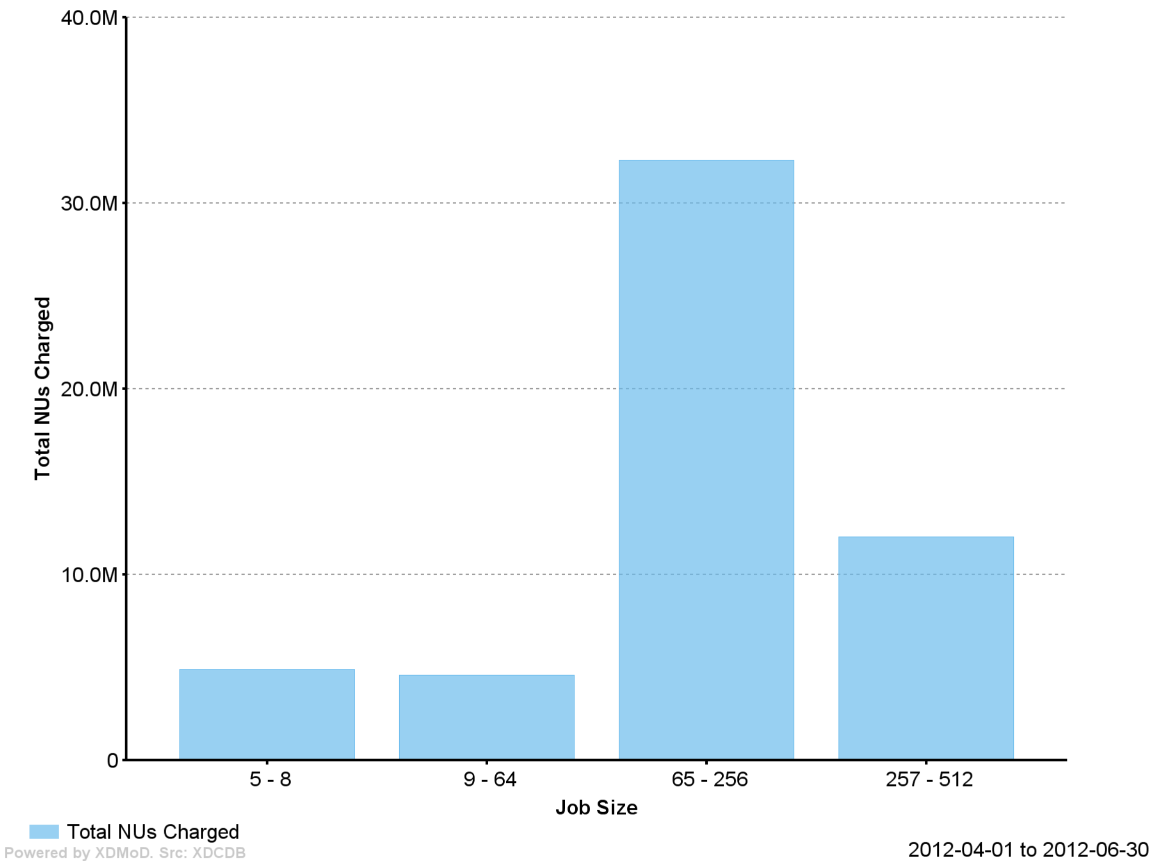
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = TACC-LONGHORN

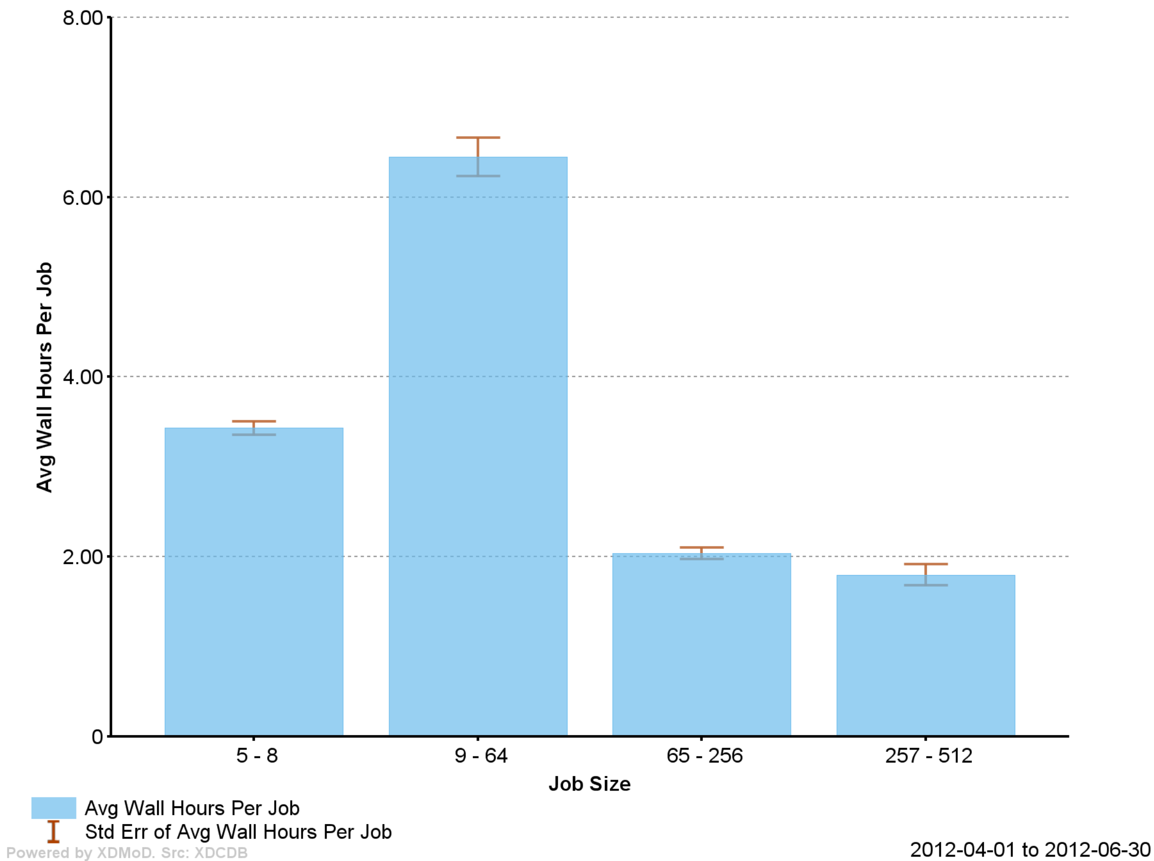
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = TACC-LONGHORN

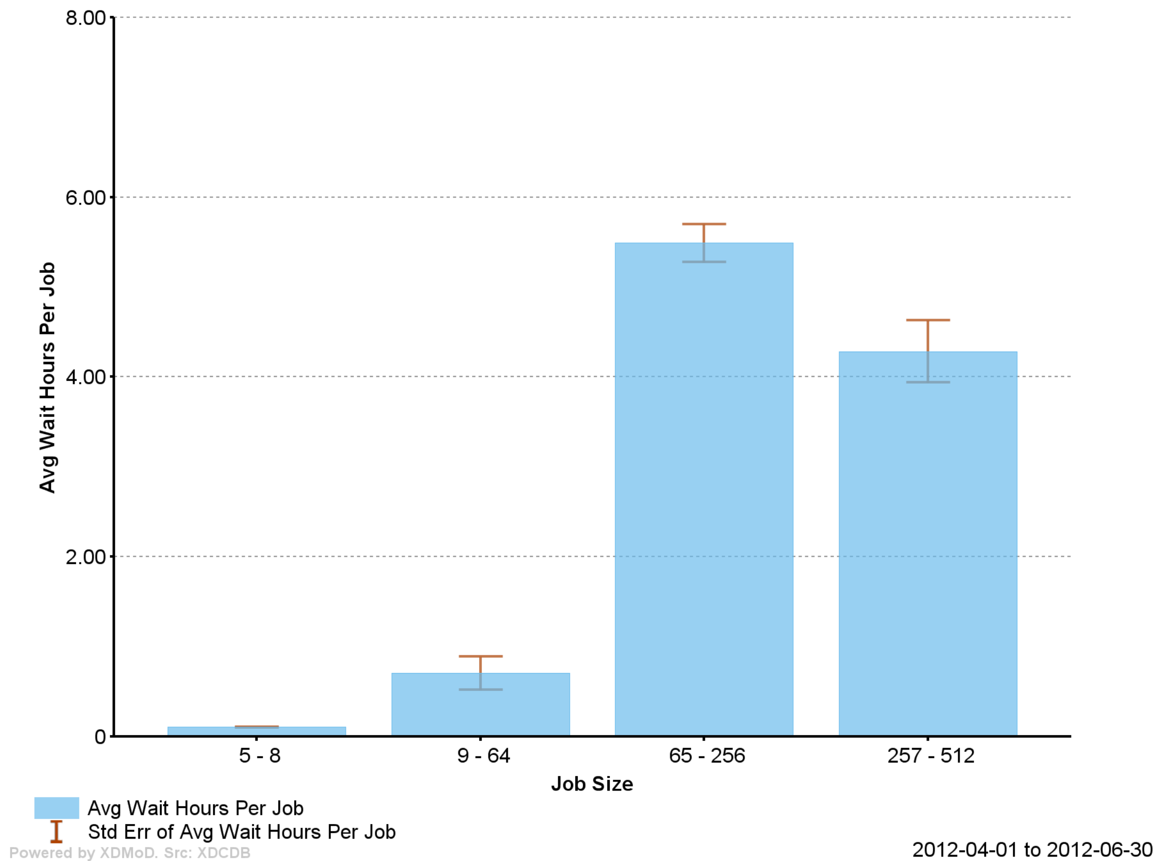
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = TACC-LONGHORN

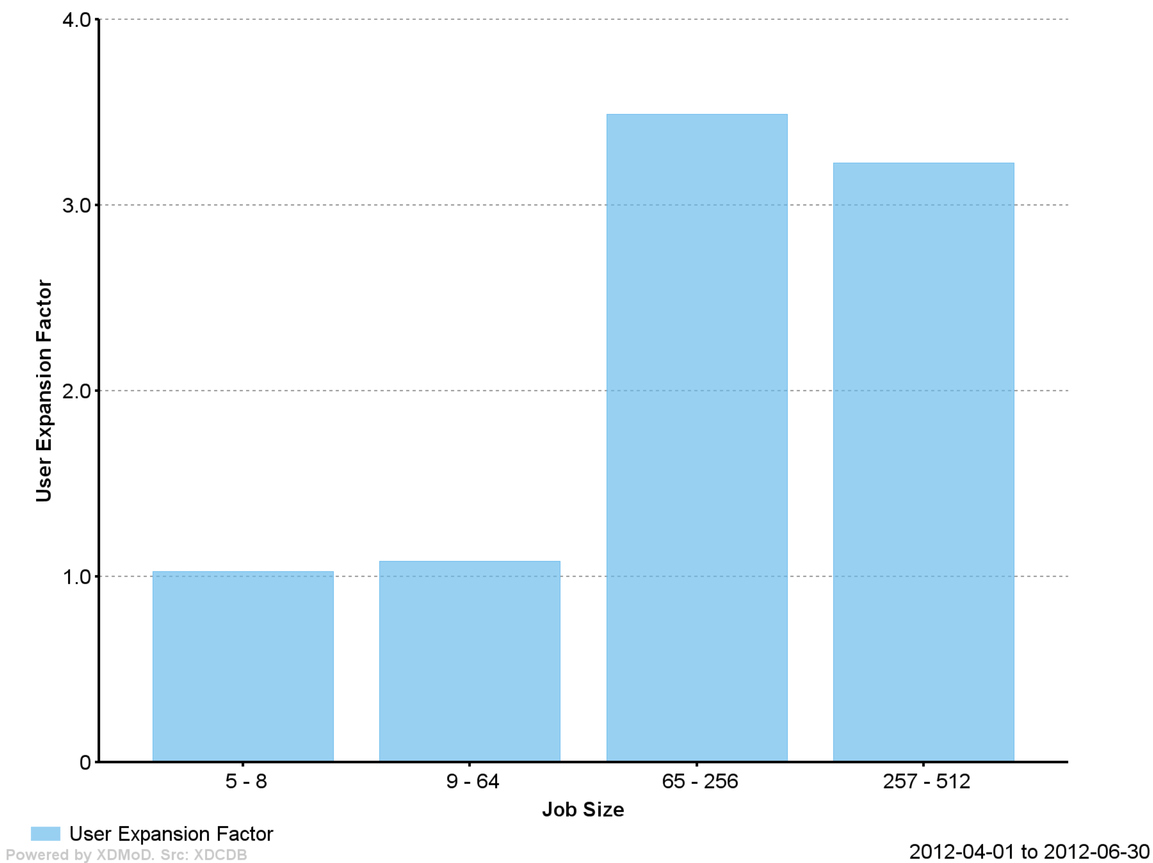
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = TACC-LONGHORN

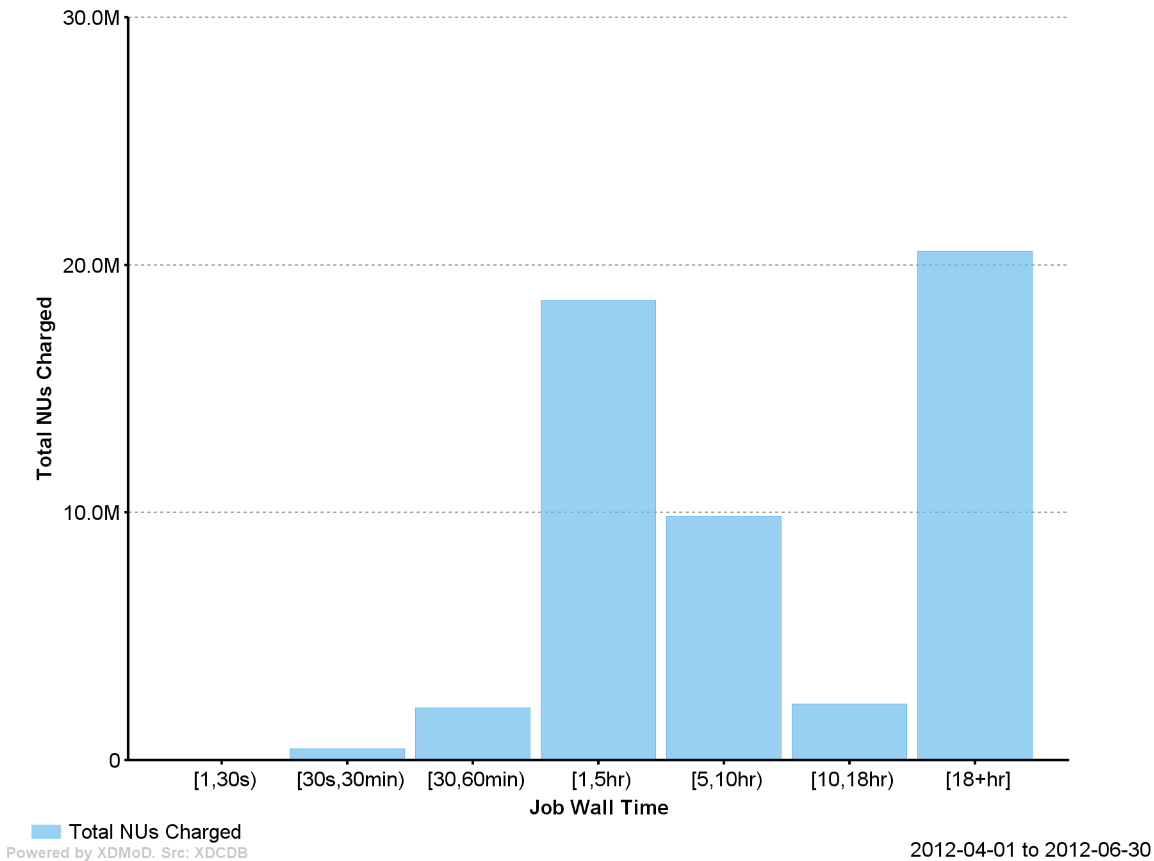
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = TACC-LONGHORN

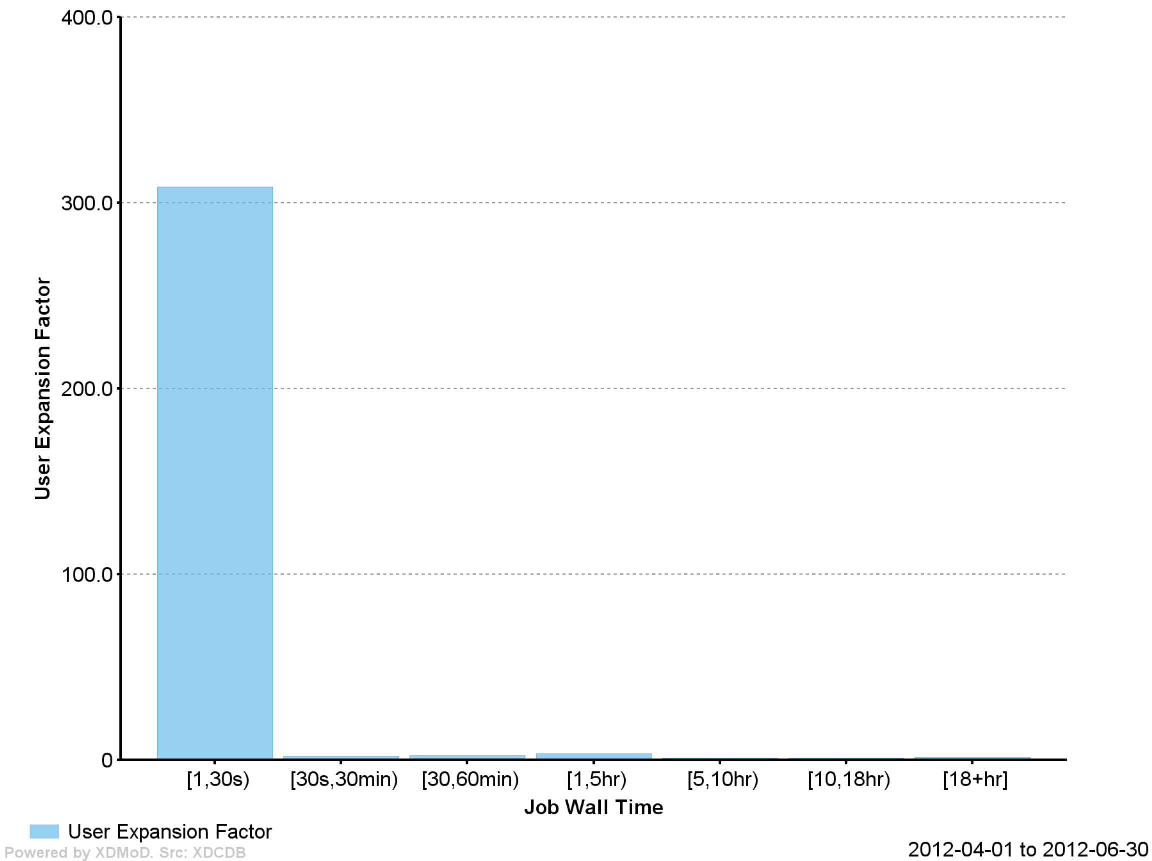
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = TACC-LONGHORN

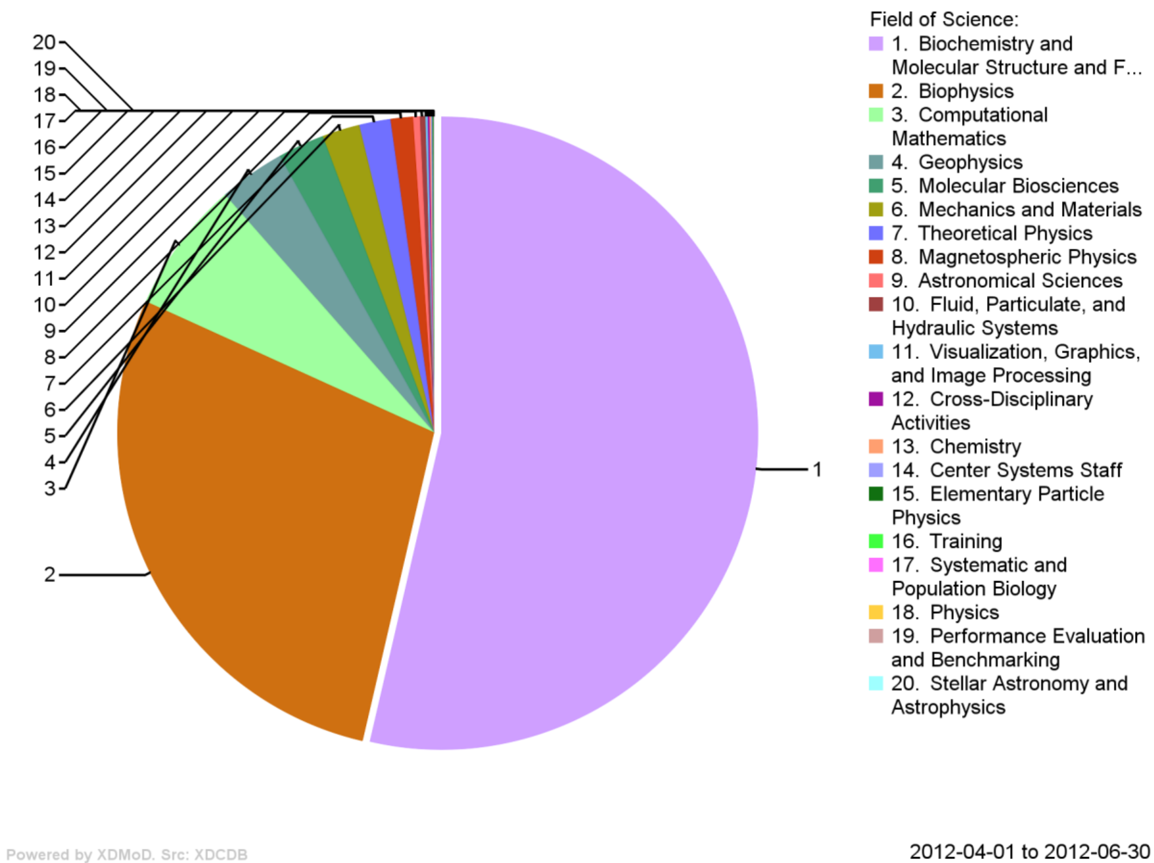
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = TACC-LONGHORN

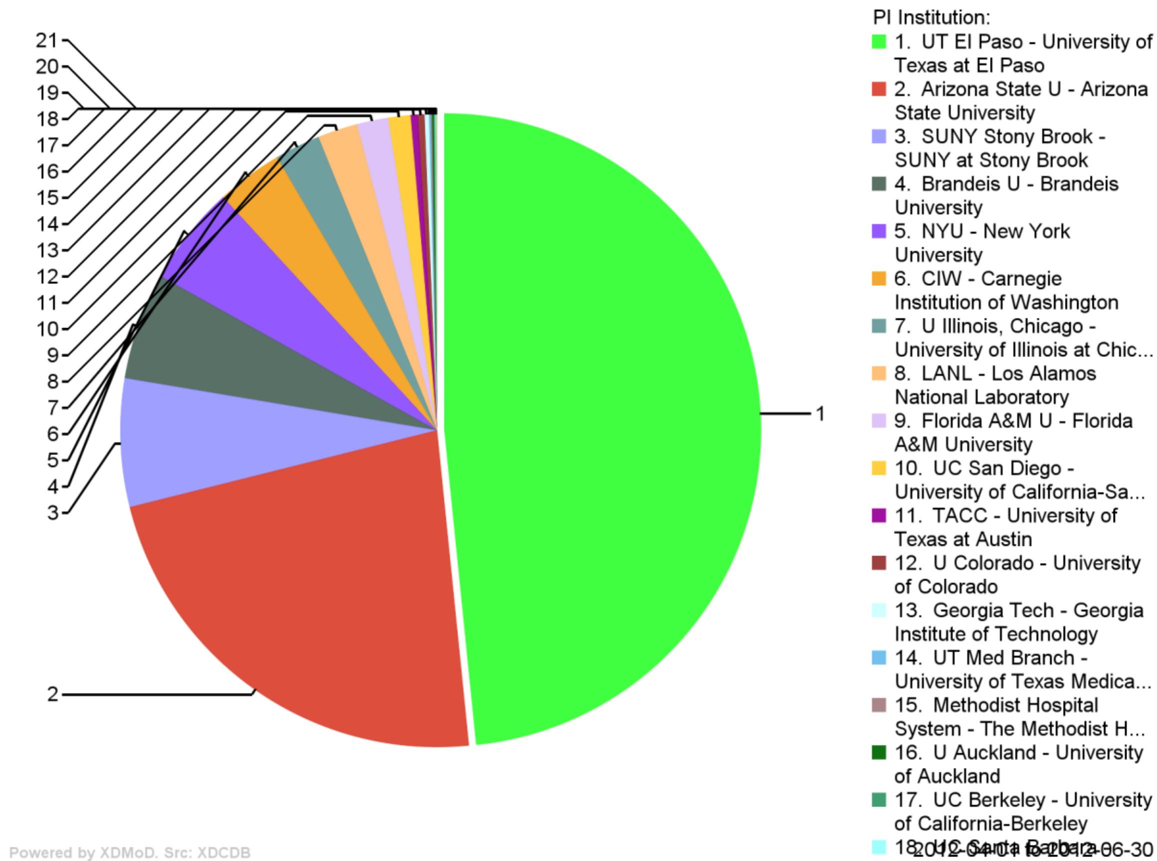
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = TACC-LONGHORN

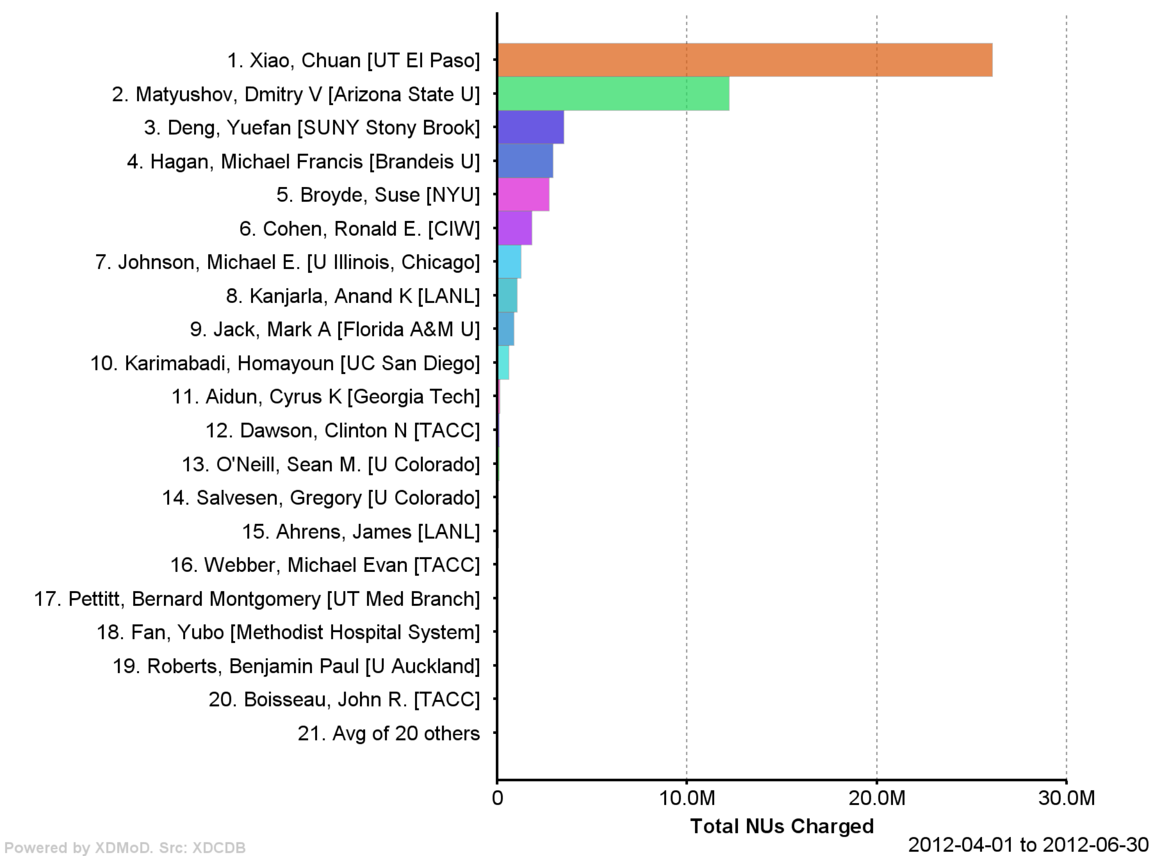
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = TACC-LONGHORN

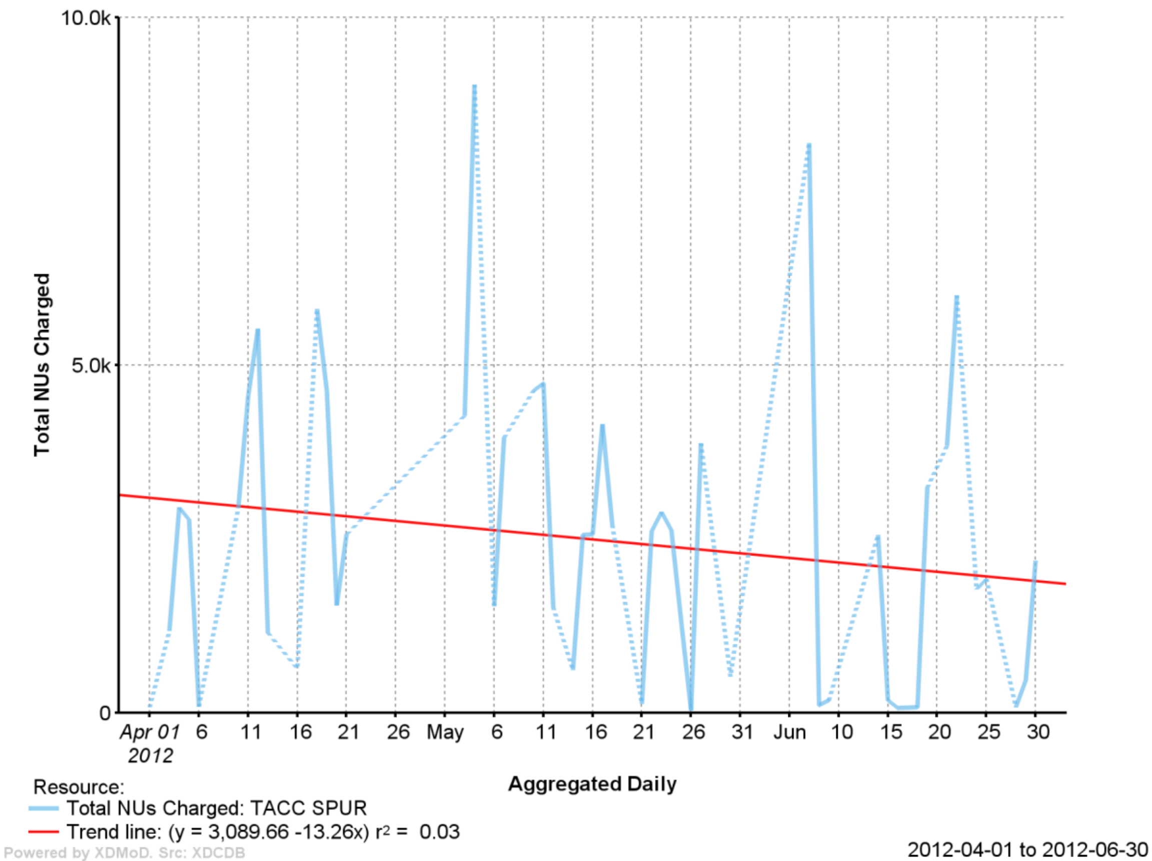
2012-04-01 to 2012-06-30



Total NUs Charged by Resource

Resource = TACC-SPUR

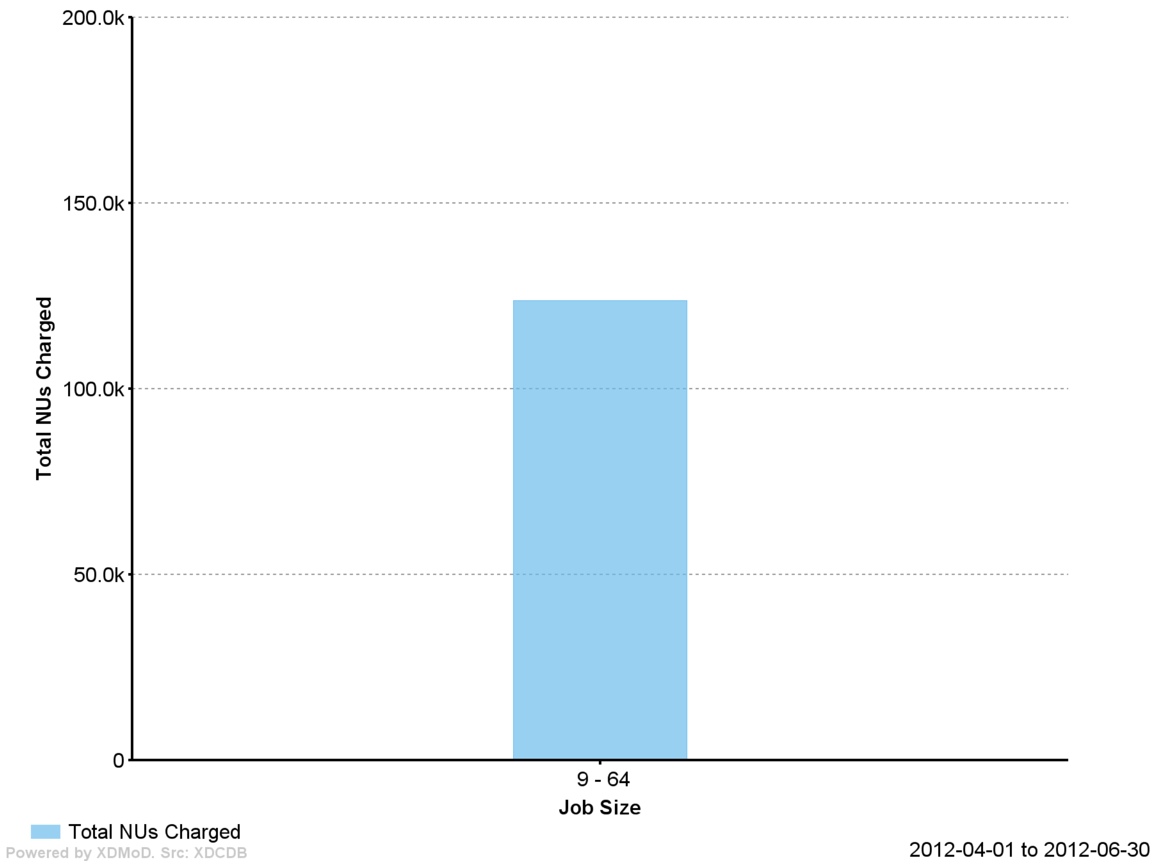
2012-04-01 to 2012-06-30



Total NUs Charged by Job Size

Resource = TACC-SPUR

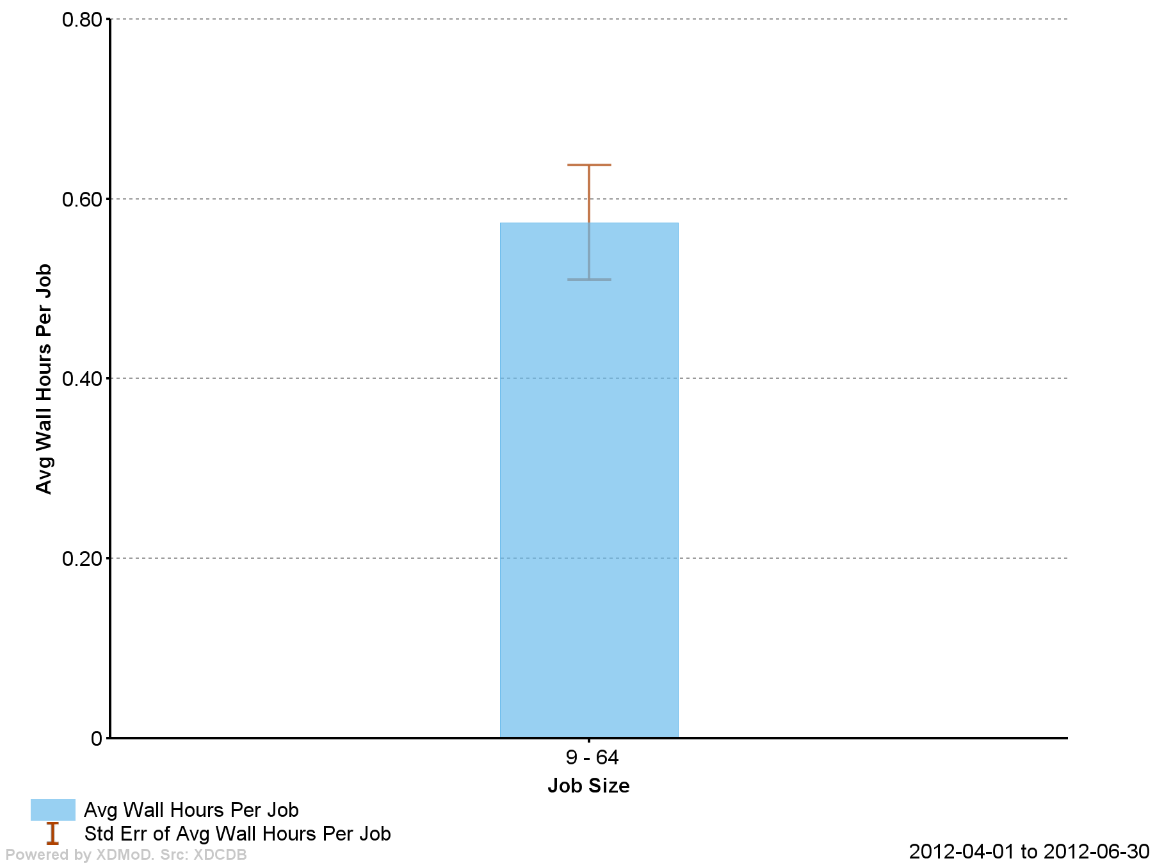
2012-04-01 to 2012-06-30



Avg Wall Hours Per Job by Job Size

Resource = TACC-SPUR

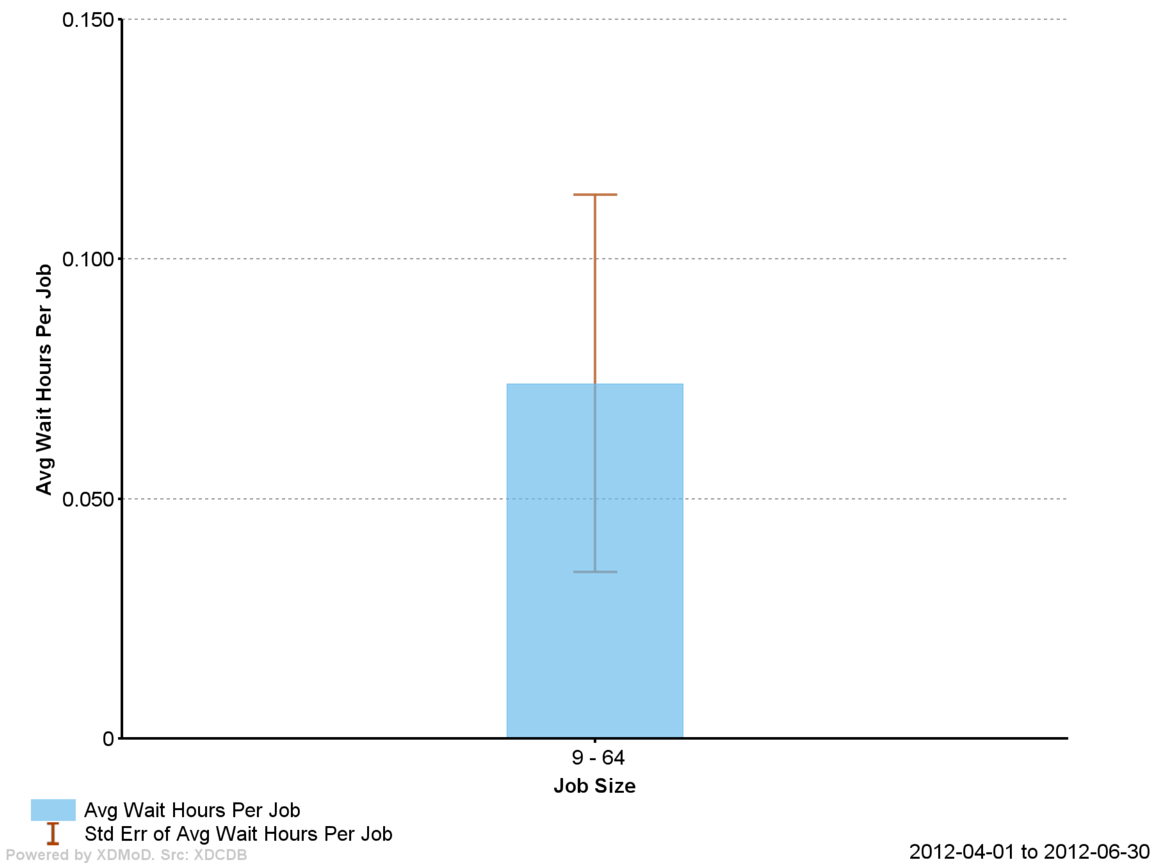
2012-04-01 to 2012-06-30



Avg Wait Hours Per Job by Job Size

Resource = TACC-SPUR

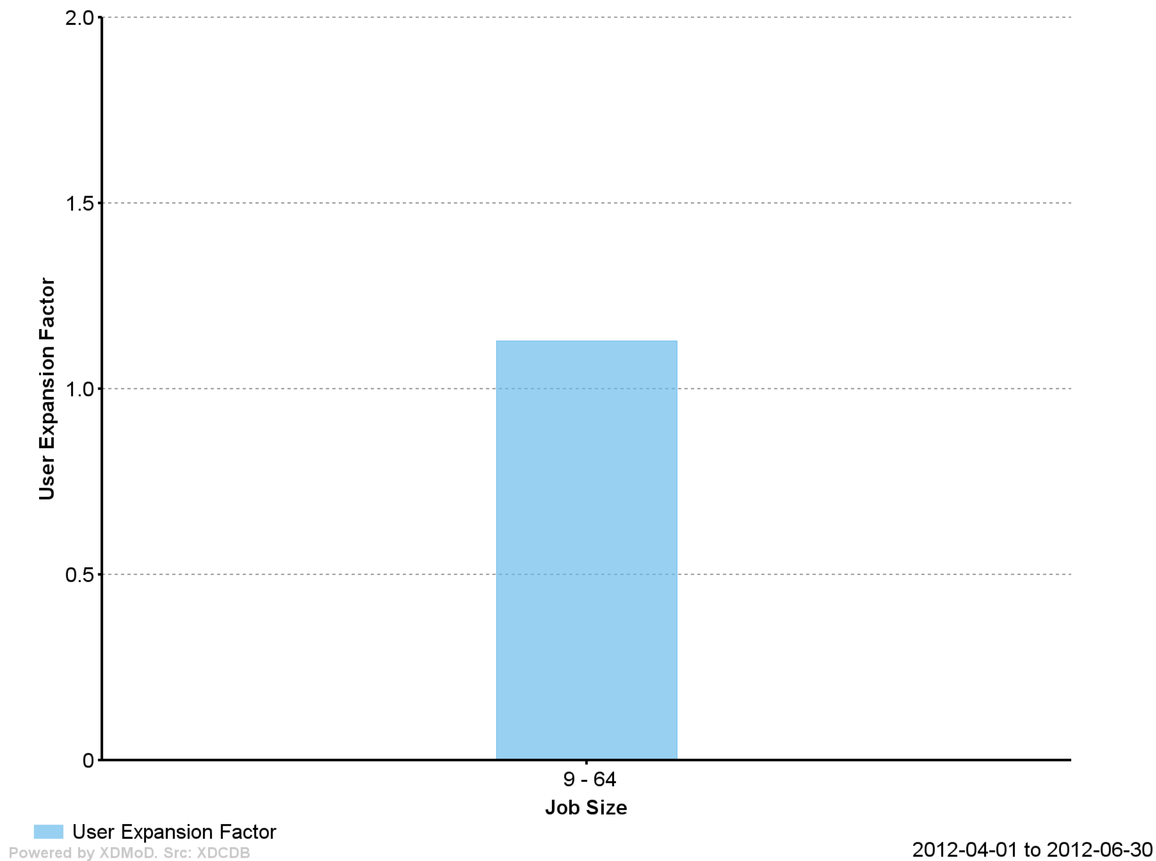
2012-04-01 to 2012-06-30



User Expansion Factor by Job Size

Resource = TACC-SPUR

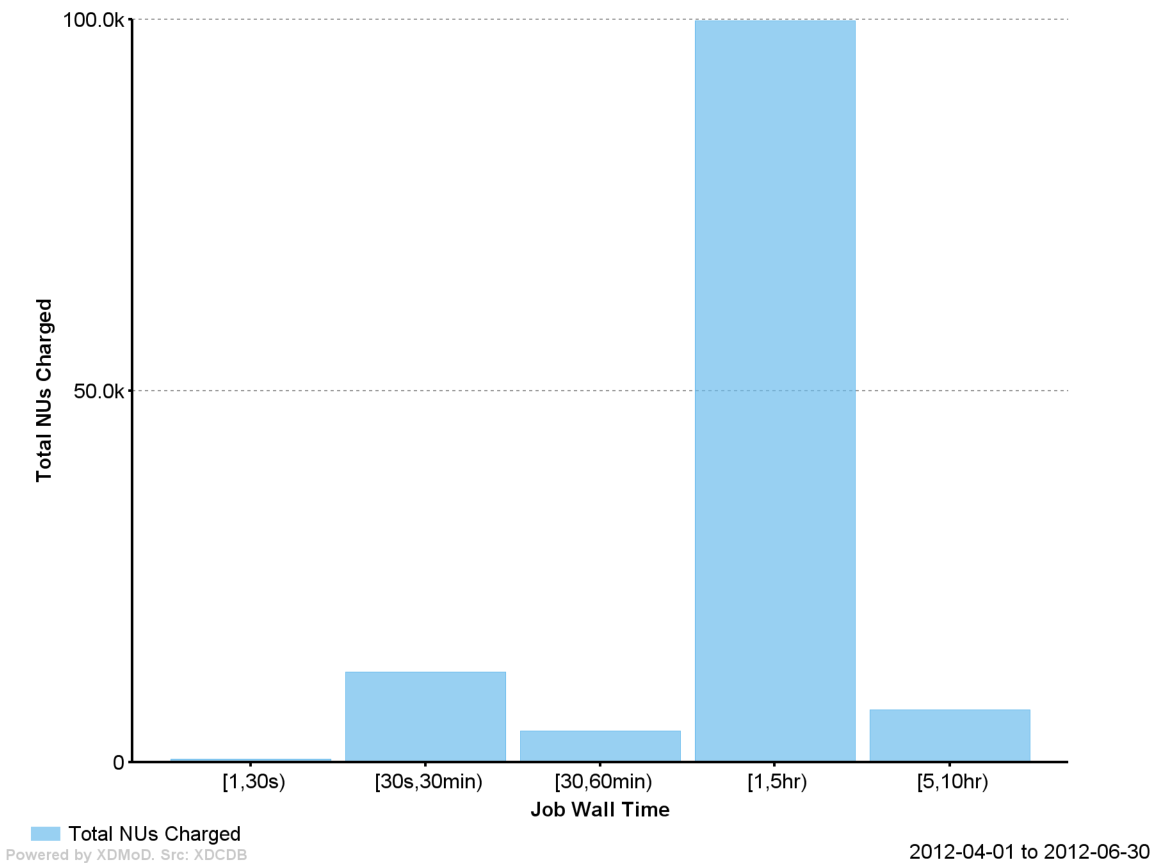
2012-04-01 to 2012-06-30



Total NUs Charged by Job Wall Time

Resource = TACC-SPUR

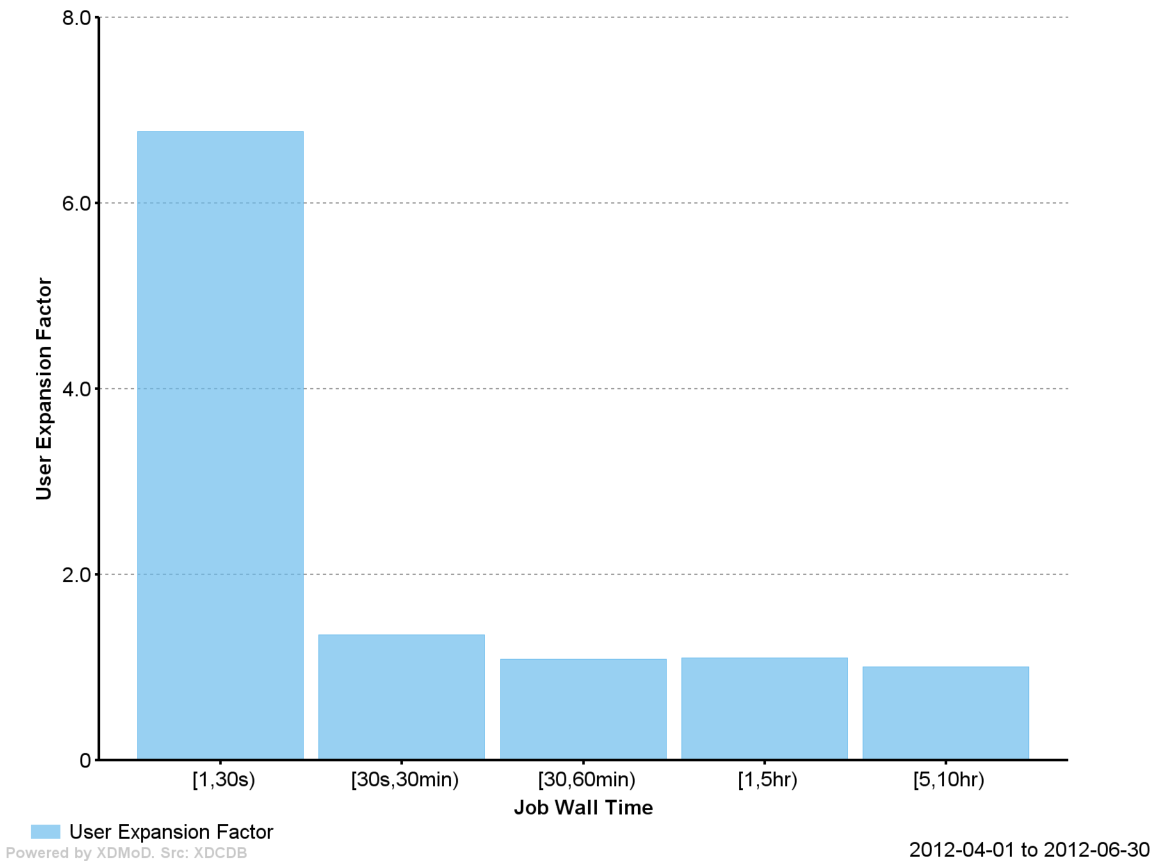
2012-04-01 to 2012-06-30



User Expansion Factor by Job Wall Time

Resource = TACC-SPUR

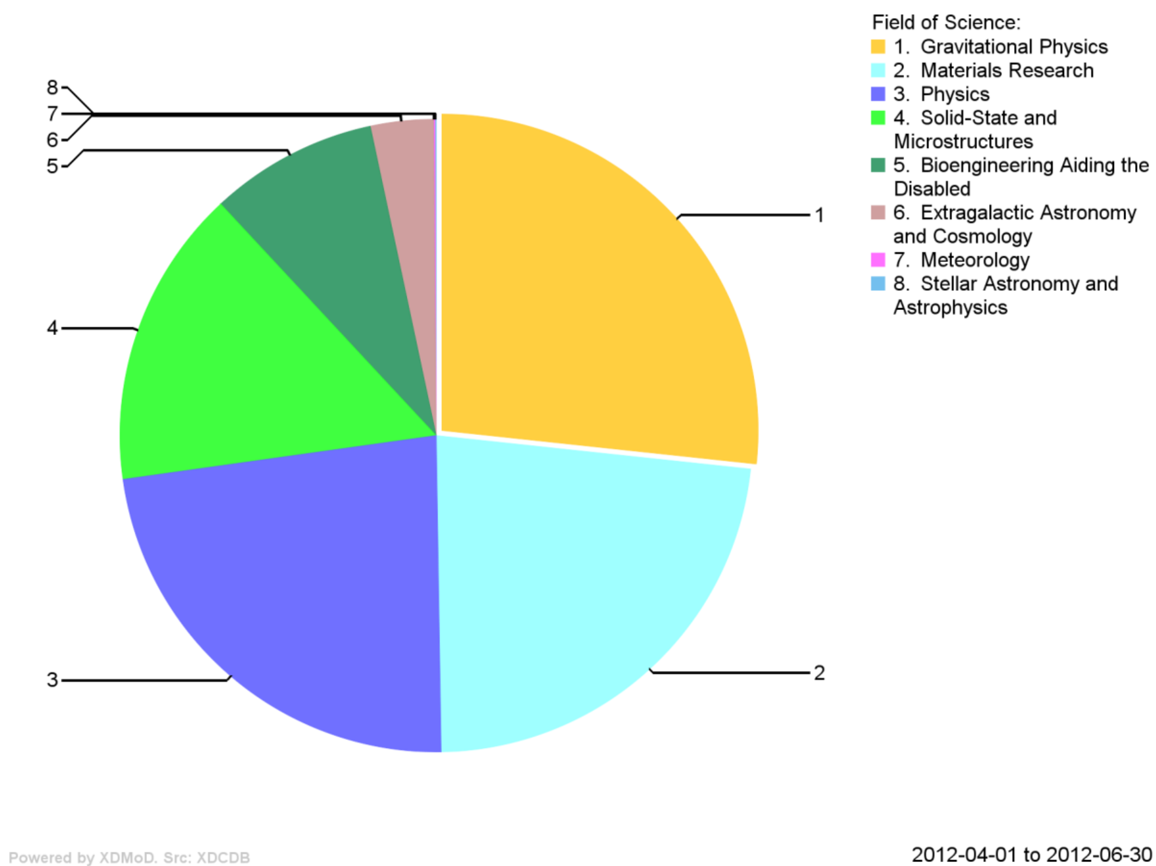
2012-04-01 to 2012-06-30



Total NUs Charged by Field of Science

Resource = TACC-SPUR

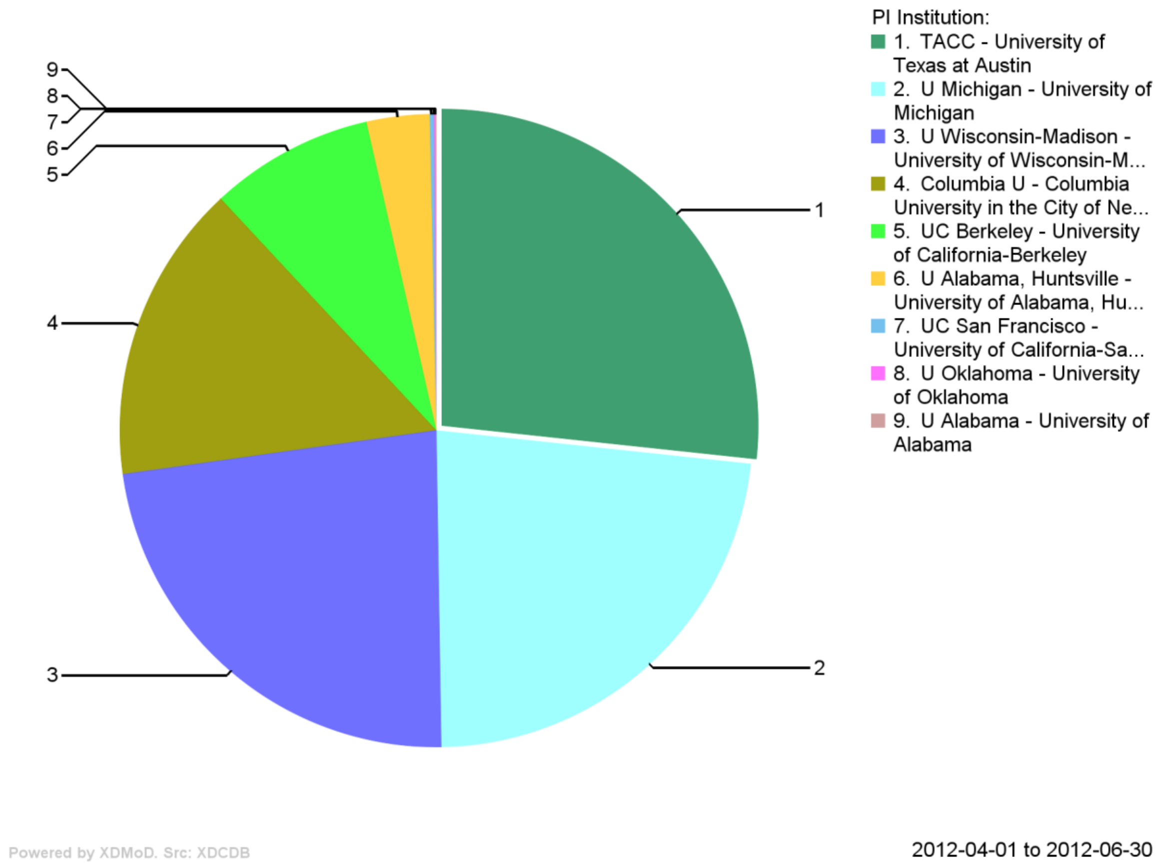
2012-04-01 to 2012-06-30



Total NUs Charged by PI Institution

Resource = TACC-SPUR

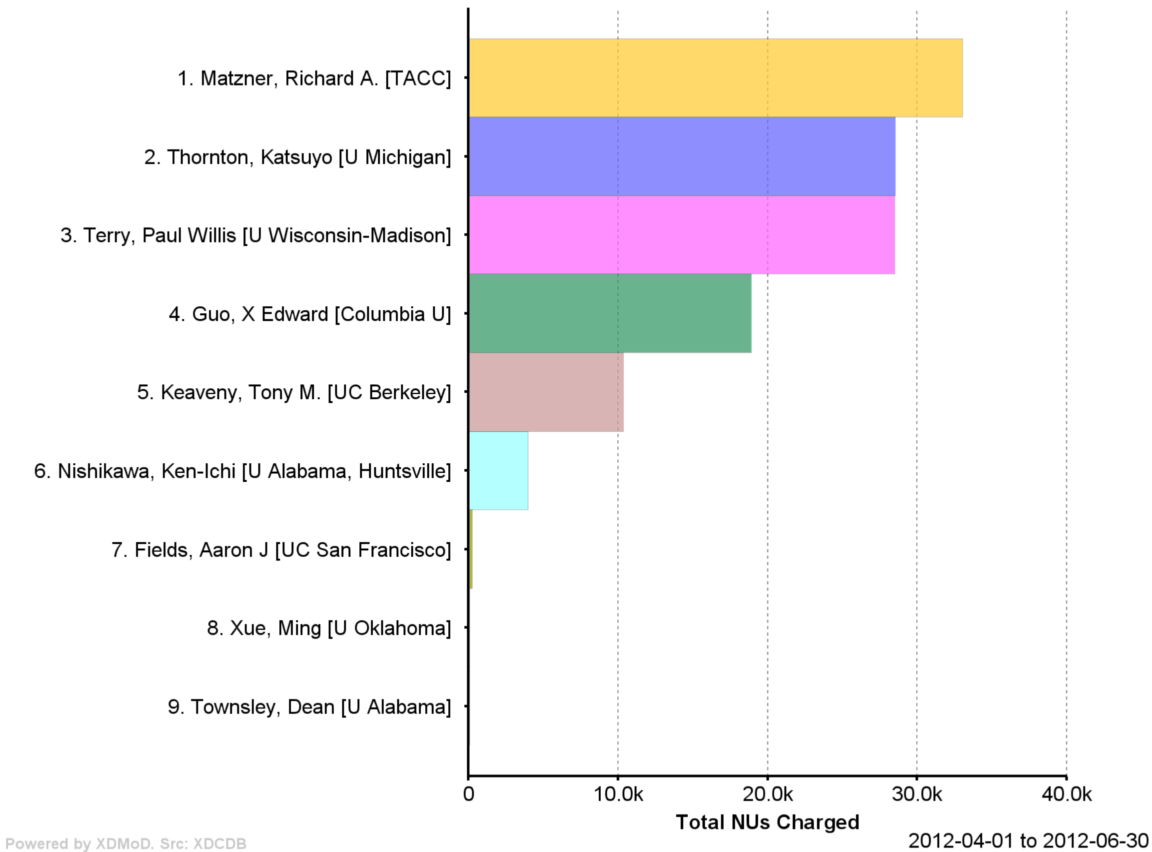
2012-04-01 to 2012-06-30



Total NUs Charged by PI

Resource = TACC-SPUR

2012-04-01 to 2012-06-30



19.9.3 *Standard User Assistance Metrics*

TACC staff members continue to provide trouble ticket support via the XSEDE ticket system and the TACC Consulting System. 312 tickets, submitted via the XSEDE ticket system, were handled by TACC staff during the report period with 306 being closed. Both trouble ticket systems are monitored 7x24x365 and approximately 25 TACC staff members are engaged in this front-line support activity. The following table indicates the number of tickets opened, closed, and a breakdown of the ticket category.

Issue Category	Number of tickets opened	Number of tickets closed
Jobs/Batch Queues	100	98
Software/Applications	57	55
Login/Access Issues	88	88
System Issues	11	10
Account Issues	17	17
Filesystem Issues	13	12
Other	26	26

TACC ticket resolution times by category from the XSEDE ticket system.

Time to Resolution	account issues	file systems	grid software	jobs/batch queues	login/access issues	mss/data issues	network issues	software/apps	system issues	other
0-1 hr				4	5			2	2	1
1-24 hr	5	1		9	30	2		5	5	4
1-7 d	6	2		29	21	4		11		3
1-2 wk		2		16	7			14	1	1
> 2 wk	2	4	1	31	3	4		24		3
Still Open		2		4	1		1	3		1

XSEDE users also may submit requests for assistance via the TACC User Portal. During the reporting period 246 tickets were submitted via through the TUP; 223 have been resolved, 10 are pending user response, and 13 are in progress.

19.9.4 *SP-specific Metrics*

The following table contains uptime statistics for the reporting period for TACC compute, visualization, and storage resources.

TACC Resource Uptime Statistics

	Lonestar					Spur				
	Uptime					Uptime				
	PM		Outage			PM		Outage		
Month	#	Hrs	#	Hrs	%Up	#	Hrs	#	Hrs	%Up
2012-4	0	0.00	0	0.00	100	0	0.00	0	0.00	100
2012-5	1	9.50	0	0.00	98.72	2	28.00	0	0.00	96.24
2012-6	0	0.00	0	0.00	100	0	0.00	0	0.00	100
	Ranger					Ranch				
	Uptime					Uptime				
	PM		Outage			PM		Outage		
Month	#	Hrs	#	Hrs	%Up	#	Hrs	#	Hrs	%Up
2012-4	0	0.00	0	0.00	100	0	0.00	1	0.50	99.93
2012-5	2	28.00	0	0.00	96.24	0	0.00	0	0.00	100
2012-6	0	0.00	0	0.00	100	0	0.00	1	39.50	94.51
	Longhorn									
	Uptime									
	PM		Outage							
Month	#	Hrs	#	Hrs	%Up					
2012-4	0	0.00	0	0.00	100					
2012-5	0	0.00	0	0.00	100					
2012-6	0	0.00	0	0.00	100					

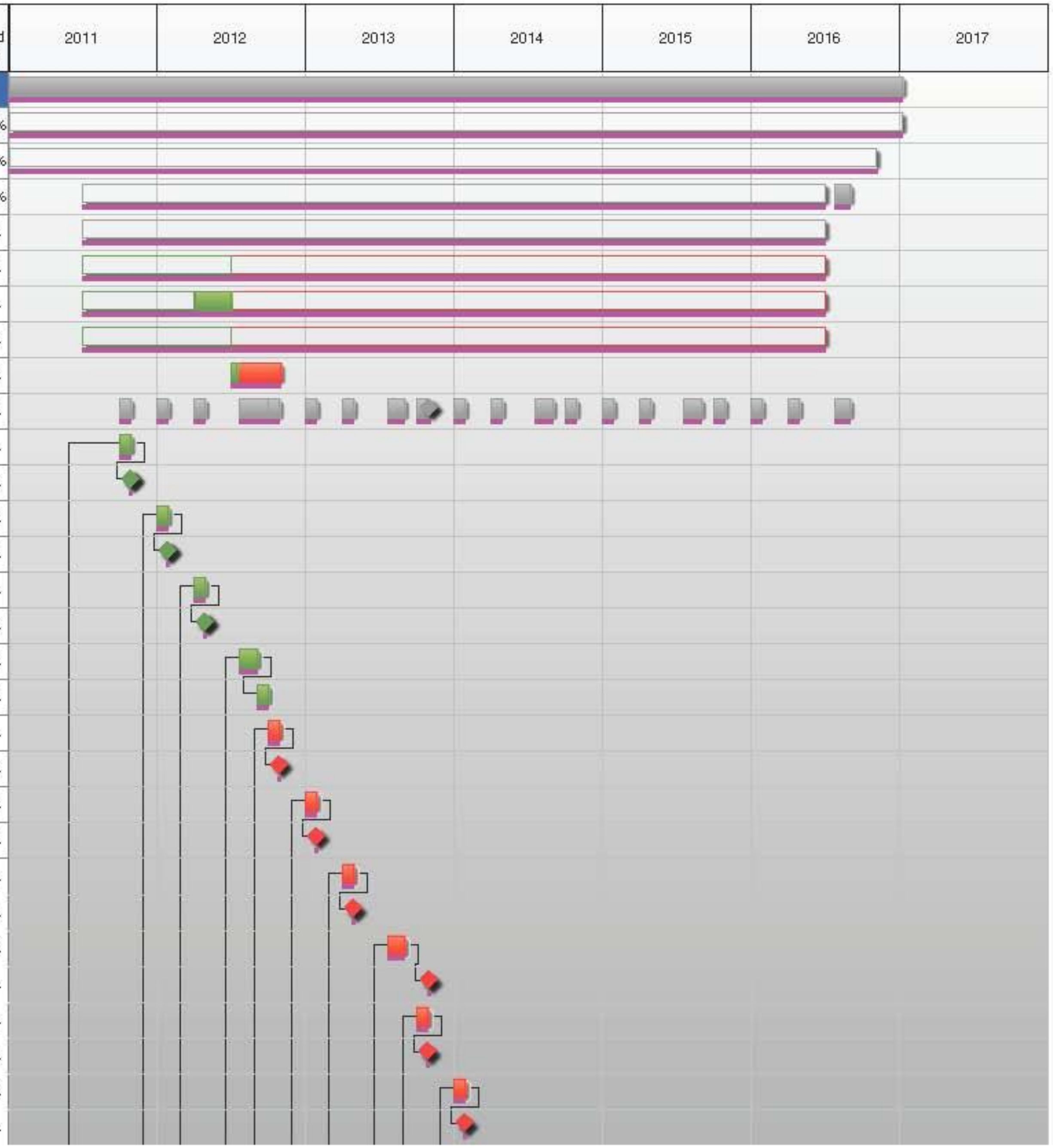
A XSEDE Project Milestones Update

Content for this appendix is pending finalizing the XSEDE Architecture that is being reworked due to the merging of the XSEDE and XROADS proposals.

B XSEDE Schedule with Progress Update

XSEDE

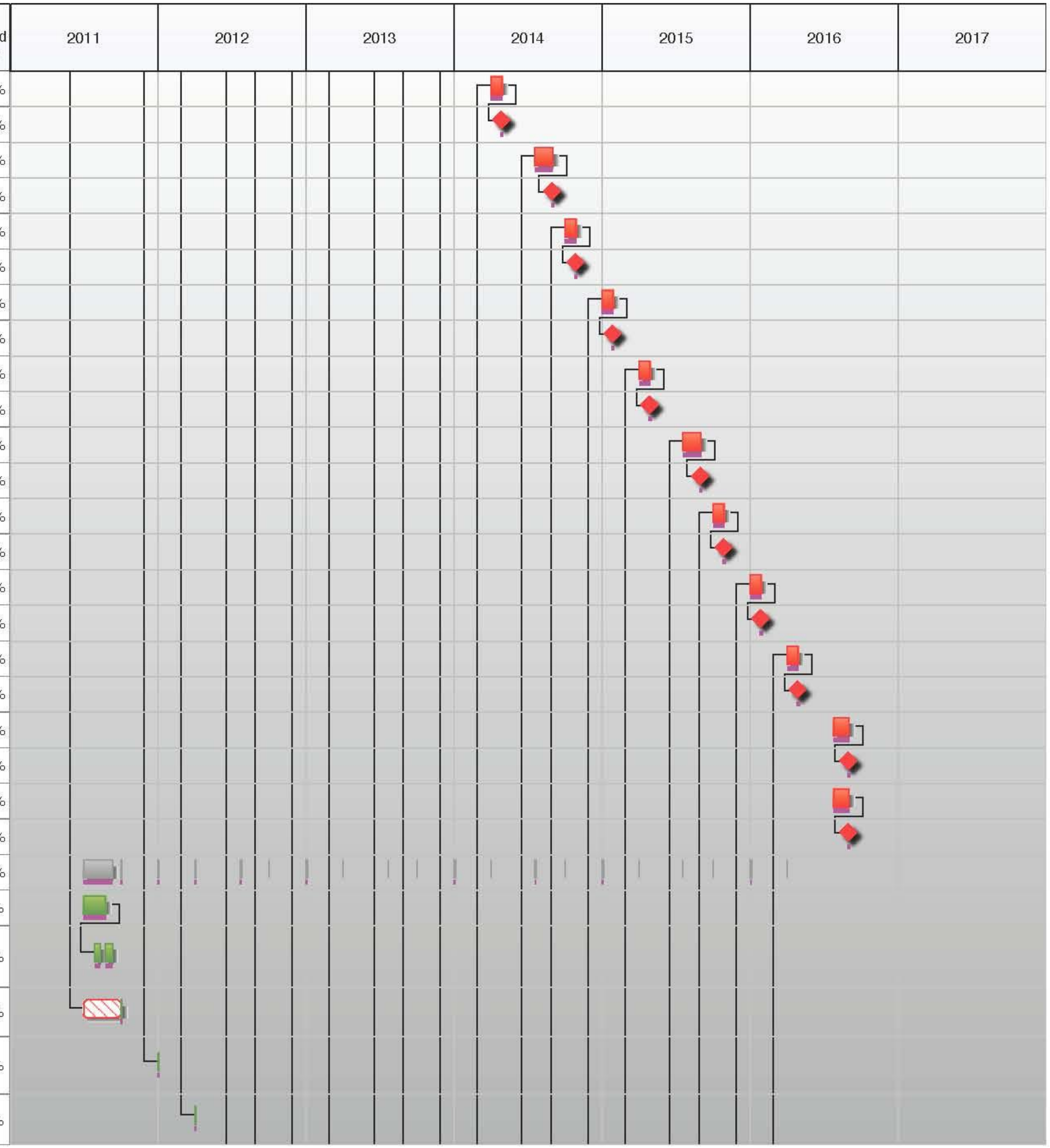
Name	Start	Finish	% Completed - XSEDE
Prodution: XSEDE Schedule	1/3/11	1/6/17	
XSEDE	1/3/11	1/6/17	0.00%
Project Office	1/3/11	11/3/16	0.00%
Project Management and Reporting	7/1/11	8/31/16	0.00%
Management	7/1/11	7/1/16	0.00%
Ongoing: Senior Management Team - Oversight and management of XSEDE	7/1/11	6/30/16	20.00%
Ongoing: Business and finance office support - subaward management	7/1/11	7/1/16	20.00%
Ongoing: Annual Planning, Budgeting, Change Control resulting from SEMP	7/1/11	6/30/16	20.00%
Update Project Execution Plan	7/2/12	10/31/12	10.00%
Project Reporting	10/3/11	8/31/16	0.00%
Prepare 2011 Q3 Quarterly Report	10/3/11	10/28/11	100.00%
Milestone: 2011 Q3 Quarterly Report completed	10/28/11	10/28/11	100.00%
Prepare 2011 Q4 Quarterly Report	1/2/12	1/27/12	100.00%
Milestone: 2011 Q4 Quarterly Report completed	1/27/12	1/27/12	100.00%
Prepare 2012 Q1 Quarterly Report	4/2/12	4/27/12	100.00%
Milestone: 2012 Q1 Quarterly Report completed	4/27/12	4/27/12	100.00%
Prepare Y1 Annual Report	7/23/12	8/31/12	100.00%
Milestone: Y1 Annual Report completed	9/3/12	9/28/12	100.00%
Prepare 2012 Q3 Quarterly Report	10/1/12	10/26/12	0.00%
Milestone: 2012 Q3 Quarterly Report completed	10/26/12	10/26/12	0.00%
Prepare 2012 Q4 Quarterly Report	1/1/13	1/28/13	0.00%
Milestone: 2012 Q4 Quarterly Report completed	1/28/13	1/28/13	0.00%
Prepare 2013 Q1 Quarterly Report	4/1/13	4/26/13	0.00%
Milestone: 2013 Q1 Quarterly Report completed	4/26/13	4/26/13	0.00%
Prepare Y2 Annual Report	7/22/13	8/30/13	0.00%
Milestone: Y2 Annual Report completed	11/1/13	11/1/13	0.00%
Prepare 2013 Q3 Quarterly Report	10/1/13	10/29/13	0.00%
Milestone: 2013 Q3 Quarterly Report completed	10/29/13	10/29/13	0.00%
Prepare 2013 Q4 Quarterly Report	1/1/14	1/28/14	0.00%
Milestone: 2013 Q4 Quarterly Report completed	1/28/14	1/28/14	0.00%



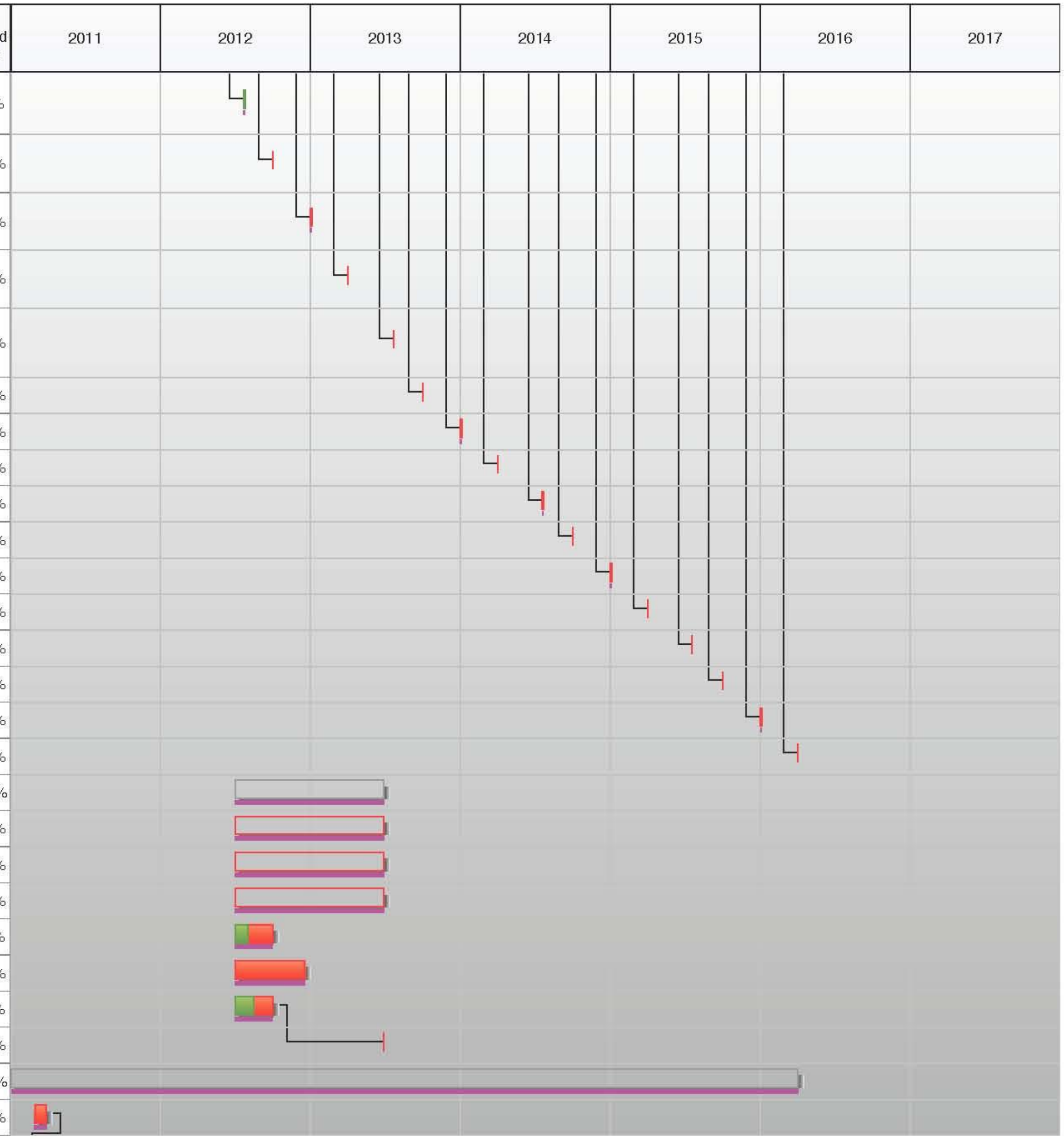
Sciforma 4.0

Printed on **Thursday, August 23, 2012**

Name	Start	Finish	% Completed - XSEDE
Prepare 2014 Q1 Quarterly Report	4/1/14	4/28/14	0.00%
Milestone: 2014 Q1 Quarterly Report completed	4/28/14	4/28/14	0.00%
Prepare Y3 Annual Report	7/21/14	8/29/14	0.00%
Milestone: Y3 Annual Report completed	8/29/14	8/29/14	0.00%
Prepare 2014 Q3 Quarterly Report	10/1/14	10/28/14	0.00%
Milestone: 2014 Q3 Quarterly Report completed	10/28/14	10/28/14	0.00%
Prepare 2014 Q4 Quarterly Report	1/1/15	1/28/15	0.00%
Milestone: 2014 Q4 Quarterly Report completed	1/28/15	1/28/15	0.00%
Prepare 2015 Q1 Quarterly Report	4/1/15	4/28/15	0.00%
Milestone: 2015 Q1 Quarterly Report completed	4/28/15	4/28/15	0.00%
Prepare Y4 Annual Report	7/20/15	8/31/15	0.00%
Milestone: Y4 Annual Report completed	8/31/15	8/31/15	0.00%
Prepare 2015 Q3 Quarterly Report	10/1/15	10/28/15	0.00%
Milestone: 2015 Q3 Quarterly Report completed	10/28/15	10/28/15	0.00%
Prepare 2015 Q4 Quarterly Report	1/1/16	1/28/16	0.00%
Milestone: 2015 Q4 Quarterly Report completed	1/28/16	1/28/16	0.00%
Prepare 2016 Q1 Quarterly Report	4/1/16	4/28/16	0.00%
Milestone: 2016 Q1 Quarterly Report completed	4/28/16	4/28/16	0.00%
Prepare Y5 Annual Report	7/25/16	8/31/16	0.00%
Milestone: Y5 Annual Report completed	8/31/16	8/31/16	0.00%
Prepare Final Report	7/25/16	8/31/16	0.00%
Milestone: Final Report completed	8/31/16	8/31/16	0.00%
Project Management - Risk Management	7/1/11	4/1/16	0.00%
Identify risks with Level 3 WBS Managers	7/1/11	8/25/11	100.00%
Review risk register with Level 2 WBS Managers	7/29/11	9/8/11	100.00%
Quarterly review of risk register - 2011 Q3	10/3/11	10/4/11	100.00%
Quarterly review of risk register - 2011 Q4	1/2/12	1/3/12	100.00%
Quarterly review of risk register - 2012 Q1	4/2/12	4/3/12	100.00%



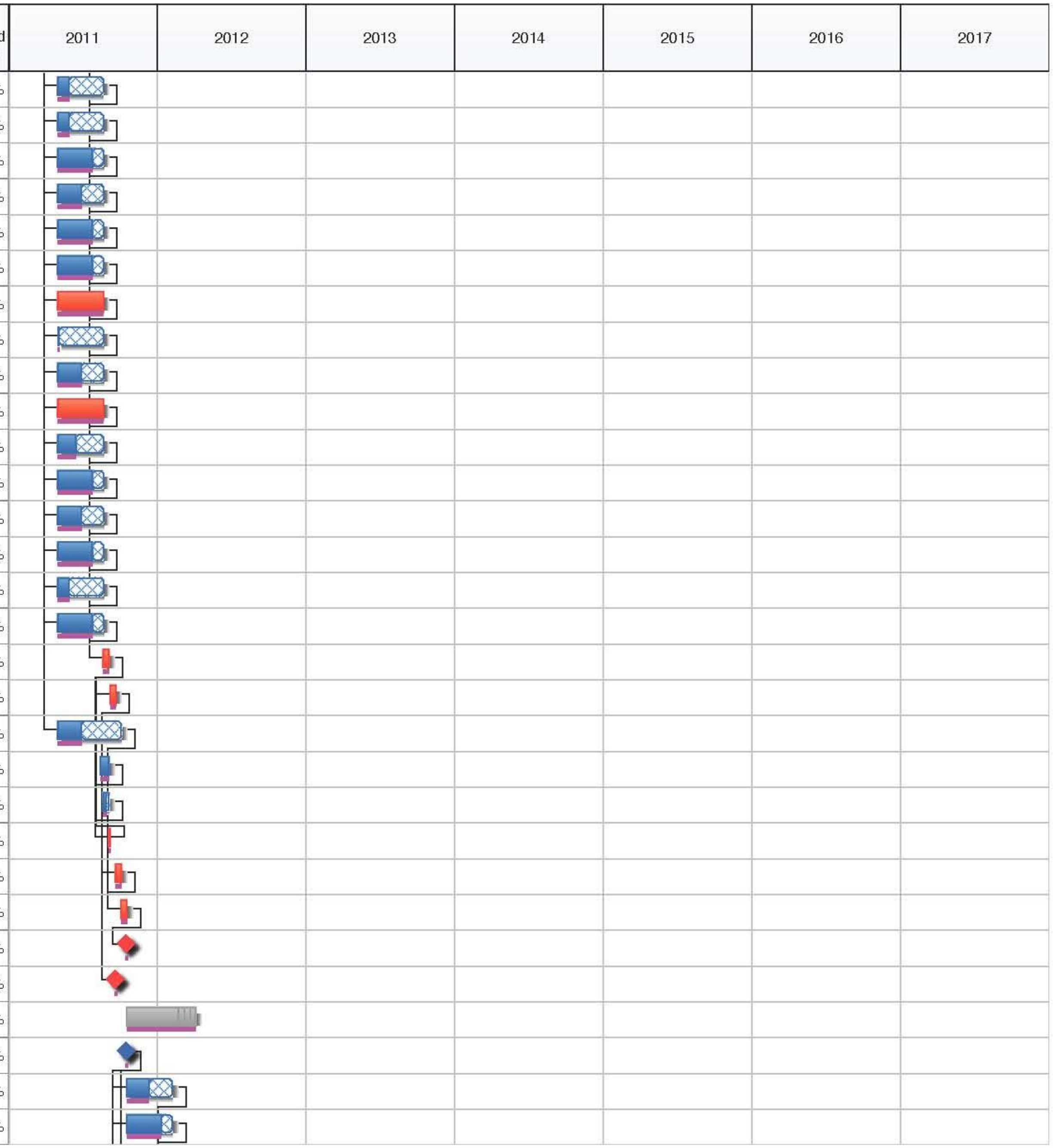
Name	Start	Finish	% Completed - XSEDE
Quarterly review of risk register - 2012 Q2	7/23/12	7/23/12	100.00%
Quarterly review of risk register - 2012 Q3	10/1/12	10/1/12	0.00%
Quarterly review of risk register - 2012 Q4	1/1/13	1/1/13	0.00%
Quarterly review of risk register - 2013 Q1	4/1/13	4/1/13	0.00%
Quarterly review of risk register - 2013 Q2	7/22/13	7/22/13	0.00%
Quarterly review of risk register - 2013 Q3	10/1/13	10/1/13	0.00%
Quarterly review of risk register - 2013 Q4	1/1/14	1/1/14	0.00%
Quarterly review of risk register - 2014 Q1	4/1/14	4/1/14	0.00%
Quarterly review of risk register - 2014 Q2	7/21/14	7/21/14	0.00%
Quarterly review of risk register - 2014 Q3	10/1/14	10/1/14	0.00%
Quarterly review of risk register - 2014 Q4	1/1/15	1/1/15	0.00%
Quarterly review of risk register - 2015 Q1	4/1/15	4/1/15	0.00%
Quarterly review of risk register - 2015 Q2	7/20/15	7/20/15	0.00%
Quarterly review of risk register - 2015 Q3	10/1/15	10/1/15	0.00%
Quarterly review of risk register - 2015 Q4	1/1/16	1/1/16	0.00%
Quarterly review of risk register - 2016 Q1	4/1/16	4/1/16	0.00%
⊟ Systems Engineering	7/2/12	6/28/13	0.00%
User needs collection	7/2/12	6/28/13	0.00%
Requirements analysis	7/2/12	6/28/13	0.00%
Managing the UREP	7/2/12	6/28/13	0.00%
Updating the Requirements Management Plan	7/2/12	9/28/12	35.00%
Updating the SEMP	7/2/12	12/14/12	0.00%
XDOR/IDEALS: establishing procedures/developing documentation	7/2/12	9/28/12	50.00%
Management of XDOR process	6/28/13	6/28/13	0.00%
⊟ Systems Architecture	1/3/11	3/30/16	0.00%
Deploy Grid Middleware Infrastructure	3/1/11	3/28/11	0.00%



Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Milestone: Grid Middleware Infrastructure deployed	3/28/11	3/28/11	0.00%							
Deploy Data Management software	4/1/11	3/29/12	0.00%							
Milestone: Data Management software deployed	3/29/12	3/29/12	0.00%							
Deploy Account Management software	3/1/11	3/28/11	0.00%							
Milestone: Account Management software deployed	3/28/11	3/28/11	0.00%							
Deploy Information Services Infrastructure	3/1/11	3/28/11	0.00%							
Milestone: Information Services Infrastructure deployed	3/28/11	3/28/11	0.00%							
Deploy Common User Environment	3/1/11	3/28/11	0.00%							
Milestone: Common User Environment deployed	3/28/11	3/28/11	0.00%							
Deploy System of Systems Test Environment	3/1/11	3/28/11	0.00%							
Milestone: System of Systems Test Environment deployed	3/28/11	3/28/11	0.00%							
Spiral 1.0	1/3/11	12/21/11	0.00%							
Inc 1.0 Start	1/3/11	1/3/11	0.00%							
Refine Engineering Plan	1/3/11	1/21/11	0.00%							
Prod. Baseline 1.0 Approved	1/3/11	1/3/11	0.00%							
UNICORE Stack (1.0)	1/3/11	3/4/11	0.00%							
Genesis II Stack (1.0)	1/3/11	2/25/11	0.00%							
Execution Mgmt. (1.0)	1/3/11	2/25/11	0.00%							
Execution Mon. Sys. (1.0)	1/3/11	1/28/11	0.00%							
Information Serv. (1.0)	1/3/11	1/28/11	0.00%							
GFFS Initial	1/3/11	1/11/11	0.00%							
XWFS Initial	1/3/11	1/21/11	0.00%							
Data Management (1.0)	1/3/11	3/25/11	0.00%							
G&VO Management (1.0)	1/3/11	3/25/11	0.00%							
API&CLT (1.0)	1/3/11	3/25/11	0.00%							
GUI Tools (1.0)	1/3/11	1/28/11	0.00%							
Replica Management Sys (1.0)	1/3/11	3/25/11	0.00%							
AAM System (1.0)	1/3/11	1/28/11	0.00%							
Ticket System (1.0)	1/3/11	1/28/11	0.00%							
Ops Ctr Spt System (1.0)	1/3/11	1/28/11	0.00%							

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
XSEDE Gateway Framework (1.0)	1/3/11	3/25/11	0.00%							
Campus Interoperability Framework (1.0)	1/3/11	3/25/11	0.00%							
Investigate Identity Mgmt and Auth w InCommon	1/3/11	4/22/11	0.00%							
Security Mgmt and Services(1.0)	1/3/11	6/17/11	0.00%							
Create Security PSG	9/1/11	10/12/11	0.00%							
Approval of Security PSG	10/12/11	10/12/11	0.00%							
XSEDE Certificate Authority	9/1/11	12/21/11	0.00%							
Unified System Log (1.0)	1/3/11	1/28/11	0.00%							
CM & Deployment Mgmt System (1.0)	1/3/11	1/21/11	0.00%							
OSG Bridge (1.0)	1/3/11	2/25/11	0.00%							
XSEDE Portal	1/3/11	3/25/11	0.00%							
System Test Planning (1.0)	5/23/11	6/10/11	0.00%							
Usability Panel Planning (1.0)	5/23/11	6/3/11	0.00%							
TRR for Inc 1.0	7/4/11	7/5/11	0.00%							
Inc 1.0 Integration	6/20/11	7/1/11	0.00%							
Inc 1.0 System Test	7/6/11	8/2/11	0.00%							
Inc 1.0 Spiral I&T Complete	8/2/11	8/2/11	0.00%							
Inc 1.0 Deployment Planning	7/4/11	7/15/11	0.00%							
Training Material Dev. (1.0)	3/28/11	5/20/11	0.00%							
Inc 1.0 Deployment	8/3/11	8/16/11	0.00%							
Inc 1.0 Deployment Complete	8/16/11	8/16/11	0.00%							
XSEDE IOC	8/16/11	8/16/11	0.00%							
Spiral 2.0	5/3/11	10/18/11	0.00%							
Prod. Baseline 2.0 Approved	5/3/11	5/3/11	0.00%							
Syslog Analysis Sys (2.0)	5/3/11	5/3/11	0.00%							
AAM System (2.0)	5/3/11	6/27/11	0.00%							
Execution Monitoring Sys (2.0)	5/3/11	6/27/11	0.00%							
RQOSA Mgmt Sys (2.0)	5/3/11	7/25/11	0.00%							
Security Mgmt and Services (2.0)	5/3/11	8/22/11	0.00%							
API&CLT (2.0)	5/3/11	6/27/11	0.00%							

Name	Start	Finish	% Completed - XSEDE
GUI Tools (2.0)	5/3/11	5/30/11	0.00%
CM & Deploy Mgmt (2.0)	5/3/11	5/30/11	0.00%
Campus Interop Fwrk (2.0)	5/3/11	7/25/11	0.00%
Ticket System (2.0)	5/3/11	6/27/11	0.00%
XSEDE Gateway Framework (2.0)	5/3/11	7/25/11	0.00%
Replica Mgmt (2.0)	5/3/11	7/25/11	0.00%
File System (2.0)	5/3/11	8/22/11	0.00%
Ops Ctr Spt Sys (2.0)	5/3/11	5/3/11	0.00%
Execution Mgmt (2.0)	5/3/11	6/27/11	0.00%
Info Services (2.0)	5/3/11	8/22/11	0.00%
Foreign Grid Policy Adapters (2.0)	5/3/11	6/13/11	0.00%
Bandwidth Provisioning (2.0)	5/3/11	7/25/11	0.00%
XSEDE Portal (2.0)	5/3/11	6/27/11	0.00%
G&VO Mgmt (2.0)	5/3/11	7/25/11	0.00%
Collab. Tools (2.0)	5/3/11	5/30/11	0.00%
Data Management (2.0)	5/3/11	7/25/11	0.00%
Inc 2.0 Integration	8/23/11	9/5/11	0.00%
Inc 2.0 System Test	9/7/11	9/20/11	0.00%
Training Material Dev. (2.0)	5/3/11	6/27/11	0.00%
System Test Planning (2.0)	8/15/11	9/2/11	0.00%
Usability Panel Planning (2.0)	8/22/11	8/26/11	0.00%
TRR for Inc 2.0	9/6/11	9/6/11	0.00%
Inc 2.0 Deployment Planning	9/21/11	10/4/11	0.00%
Inc 2.0 Deployment	10/5/11	10/18/11	0.00%
Inc 2.0 Deployment Complete	10/18/11	10/18/11	0.00%
Inc 2.0 Spiral I&T Complete	9/20/11	9/20/11	0.00%
Spiral 3.0	10/18/11	4/4/12	0.00%
Prod. Baseline 3.0 Approved	10/18/11	10/18/11	0.00%
Execution Mgmt (3.0)	10/18/11	12/12/11	0.00%
G&VO Mgmt (3.0)	10/18/11	1/9/12	0.00%



Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Data Managementt (3.0)	10/18/11	12/12/11	0.00%							
Info Services (3.0)	10/18/11	2/6/12	0.00%							
Replica Mgmt Sys (3.0)	10/18/11	10/18/11	0.00%							
Bandwidth Provisioning (3.0)	10/18/11	10/18/11	0.00%							
RQOSA Mgmt (3.0)	10/18/11	10/18/11	0.00%							
AAM Sys (3.0)	10/18/11	12/12/11	0.00%							
Execution Mon Sys (3.0)	10/18/11	11/14/11	0.00%							
Collab Tools (3.0)	10/18/11	11/14/11	0.00%							
API&CLT (3.0)	10/18/11	11/14/11	0.00%							
GUI Tools (3.0)	10/18/11	10/31/11	0.00%							
File System (3.0)	10/18/11	1/9/12	0.00%							
Ops Ctr Support Sys (3.0)	10/18/11	10/18/11	0.00%							
Syslog Analysis Sys (3.0)	10/18/11	10/31/11	0.00%							
XSEDE Portal (3.0)	10/18/11	11/14/11	0.00%							
System Test Planning (3.0)	2/1/12	2/21/12	0.00%							
Usability Panel Planning (3.0)	2/7/12	2/13/12	0.00%							
TRR for Inc 3.0	2/22/12	2/22/12	0.00%							
Security Mgmt and Services (3.0)	10/18/11	1/9/12	0.00%							
Inc 3.0 Integration	2/7/12	2/20/12	0.00%							
Inc 3.0 System Test	2/23/12	3/7/12	0.00%							
Inc 3.0 Spiral I&T Complete	3/7/12	3/7/12	0.00%							
Training Material Dev. (3.0)	11/7/11	12/30/11	0.00%							
Inc 3.0 Deployment Planning	3/8/12	3/21/12	0.00%							
Inc 3.0 Deployment	3/22/12	4/4/12	0.00%							
Inc 3.0 Deployment Complete	4/4/12	4/4/12	0.00%							
XSEDE FOC	4/4/12	4/4/12	0.00%							
Ongoing: Incremental improvements continue via SEMP Spiral Design Process	4/5/12	3/30/16	0.00%							
⊟ Public Facing XSEDE Architecture Document	12/5/11	3/2/12	0.00%							
Agreement of contents and level of detail	12/5/11	12/5/11	100.00%							
Establish time frame to produce public facing architecture document	12/9/11	12/9/11	100.00%							

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Outline for initial level 1 & level 2-decomposition documentation	1/6/12	1/6/12	100.00%							
First draft of public facing document	1/13/12	1/13/12	100.00%							
Revise, comment add content to document as necessary	2/2/12	2/2/12	100.00%							
Architects review first draft with Bachman	2/10/12	2/10/12	100.00%							
Identify remaining steps to complete first draft	2/13/12	2/13/12	100.00%							
A&D team including liaisons from SD&I, Security, Campus Bridging and	2/17/12	2/17/12	100.00%							
Architects address comments/revisions and request endorsement from A&D	2/28/12	2/28/12	100.00%							
First version of the XSEDE Architecture Document (Level 1 & 2 Decomposition)	3/2/12	3/2/12	100.00%							
⊟ Campus Bridging - Architectural Response to Stakeholder Requirements	1/19/12	5/3/12	0.00%							
Preliminary background work	1/19/12	2/16/12	100.00%							
Documentation and review of use cases and requirements matrix completed	2/17/12	2/23/12	100.00%							
Architectural response at a Level 3 Decomposition prepared by the	2/24/12	3/22/12	95.00%							
Stakeholder review of Architectural response	3/23/12	4/19/12	33.00%							
Stakeholder approved Architectural response for Campus Bridging	4/20/12	5/3/12	80.00%							
⊟ Science Gateways - Architectural Response to Stakeholder Requirements	5/4/12	8/9/12	0.00%							
Documentation and review of use cases and requirements matrix completed	5/4/12	5/31/12	25.00%							
Architectural response at a Level 3 Decomposition prepared by the	6/1/12	6/28/12	0.00%							
Stakeholder review of Architectural response	6/29/12	6/29/12	0.00%							
Stakeholder review of Architectural response (cont.)	7/2/12	7/26/12	0.00%							
Stakeholder approved Architectural response for Gateways incorporated into	7/27/12	8/9/12	0.00%							
⊟ Computing - Architectural Response to Stakeholder Requirements	8/10/12	11/15/12	0.00%							
Documentation and review of use cases and requirements matrix completed	8/10/12	9/6/12	10.00%							
Architectural response at a Level 3 Decomposition prepared by the	9/7/12	10/4/12	0.00%							
Stakeholder review of Architectural response	10/5/12	11/1/12	0.00%							
Stakeholder approved Architectural response for Security incorporated into	11/2/12	11/15/12	0.00%							
⊟ BIG Data - Architectural Response to Stakeholder Requirements	11/16/12	3/7/13	0.00%							
Documentation and review of use cases and requirements matrix completed	11/16/12	12/13/12	5.00%							
Architectural response at a Level 3 Decomposition prepared by the	12/14/12	1/25/13	0.00%							
Stakeholder review of Architectural response	1/28/13	2/21/13	0.00%							
Stakeholder approved Architectural response for Integration with Other	2/22/13	3/7/13	0.00%							

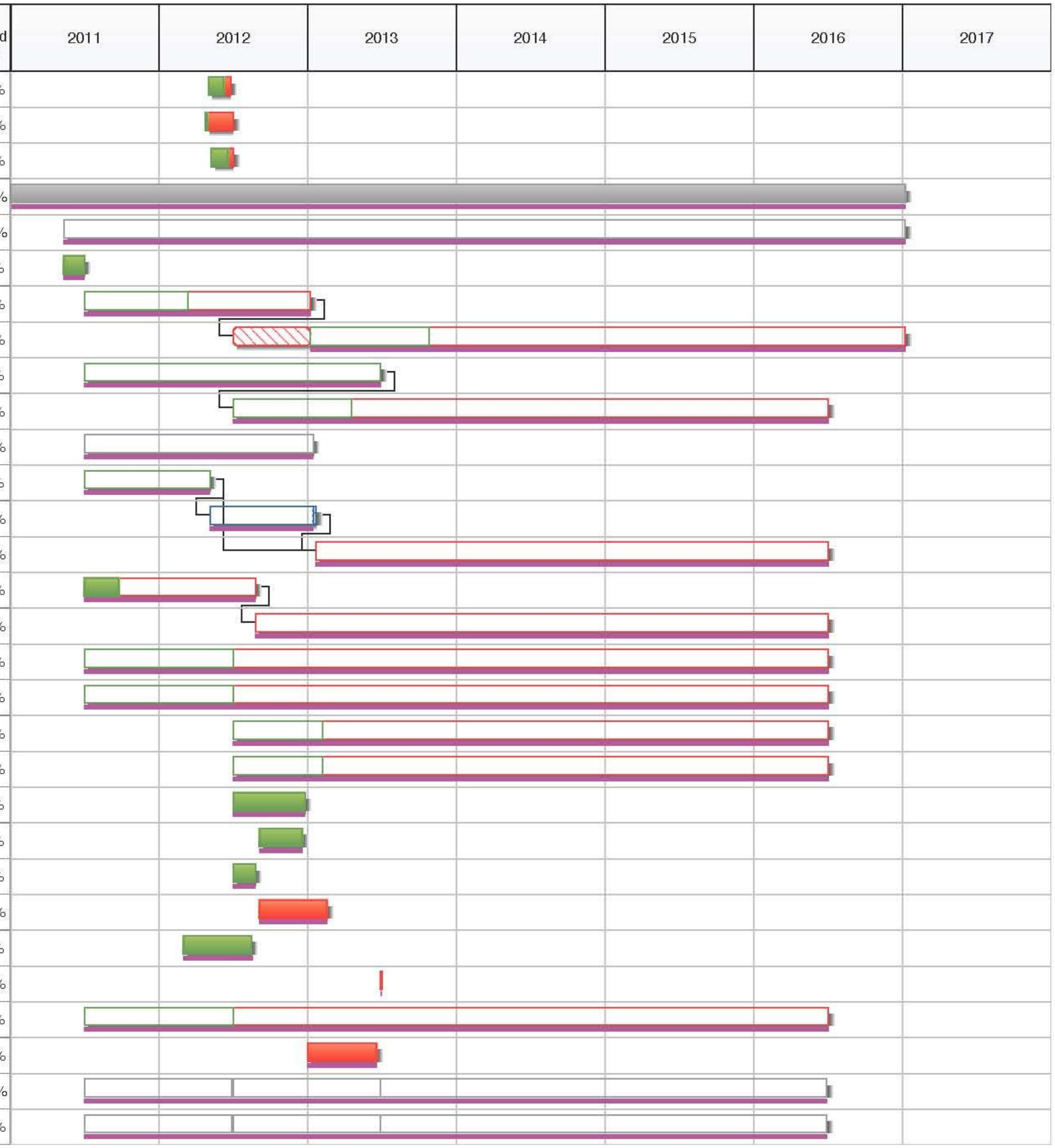
Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Connecting Instrumentation - Architectural Response to Stakeholder Requirements	3/8/13	6/13/13	0.00%							
Documentation and review of use cases and requirements matrix completed	3/8/13	4/4/13	5.00%							
Architectural response at a Level 3 Decomposition prepared by the	4/5/13	5/2/13	0.00%							
Stakeholder review of Architectural response	5/3/13	5/30/13	0.00%							
Stakeholder approved Architectural response for Replication and Fault	5/31/13	6/13/13	0.00%							
Collaboration - Architectural Response to Stakeholder Requirements	6/14/13	9/19/13	0.00%							
Documentation and review of use cases and requirements matrix completed	6/14/13	6/28/13	9.00%							
Documentation and review of use cases and requirements matrix completed	7/1/13	7/11/13	0.00%							
Architectural response at a Level 3 Decomposition prepared by the	7/12/13	8/8/13	0.00%							
Stakeholder review of Architectural response	8/9/13	9/5/13	0.00%							
Stakeholder approved Architectural response for Interprocess	9/6/13	9/19/13	0.00%							
External Relations	**1/3/11**	**11/3/16**	**0.00%**							
Generate Publications: Highlights (Science, EOT, Digital Resources)	4/2/12	11/3/16	0.00%							
SciHi subcommittee from XSEDE ER established	4/9/12	4/9/12	100.00%							
Collect story ideas	5/25/12	5/25/12	100.00%							
Story choices approved by XSEDE leadership	6/29/12	6/29/12	14.00%							
Graphic designer selected	7/2/12	7/2/12	100.00%							
About 15 science highlights stories selected, edited (incl tech review) and	7/27/12	7/27/12	10.00%							
Cover-to-cover edit complete	8/10/12	8/10/12	0.00%							
Overall design and test story mockup complete and reviewed	9/5/12	9/5/12	0.00%							
Final design complete	9/20/12	9/20/12	0.00%							
Completed book delivered to printer	9/21/12	9/21/12	0.00%							
Milestone: Science Highlights published	11/21/12	11/21/12	0.00%							
Ongoing: Repeat previous 10 tasks annually	4/2/12	11/3/16	0.00%							
Create XSEDE website and translate relevant TG website content	1/3/11	4/18/11	0.00%							
XSEDE WebsiXSEDE website committee established	1/3/11	1/3/11	100.00%							
Website requirements document complete	1/17/11	1/17/11	100.00%							
Content requirements document complete	1/17/11	1/17/11	100.00%							
First rev of design reviewed by website committee	2/15/11	2/15/11	100.00%							
Rev of website reviewed by XSEDE leadership	3/1/11	3/1/11	100.00%							

Name	Start	Finish	% Completed - XSEDE
Short form usability test completed	3/15/11	3/15/11	100.00%
Final version of website reviewed by website committee	3/22/11	3/22/11	100.00%
Content approved	4/1/11	4/1/11	100.00%
Final build complete	4/7/11	4/7/11	100.00%
Content ported and built	4/15/11	4/15/11	100.00%
Initial version of XSEDE website launched	4/18/11	4/18/11	100.00%
Milestone: XSEDE website completed	4/18/11	4/18/11	100.00%
Generate Annual Conference Proceedings	3/1/12	7/14/16	0.00%
Proceedings chair selected and incorporated into event planning	3/1/12	3/1/12	100.00%
Submission site publicized	3/1/12	3/1/12	100.00%
Approval from ACM or whatever publisher received	6/1/12	6/1/12	100.00%
Templates and copyright permission forms out to authors	6/5/12	6/5/12	100.00%
All papers and copyright permission forms received from authors and OKed	6/22/12	6/22/12	100.00%
All papers sent to production house for reproduction to USB drive	6/25/12	6/25/12	100.00%
Milestone: Annual Conference Proceedings Complete	7/12/12	7/12/12	100.00%
Ongoing: Repeat previous 7 tasks annually	7/13/12	7/14/16	20.00%
Milestone: Ongoing - Generate press releases & website content	4/4/11	3/24/16	20.00%
Milestone: Ongoing - Generate publicity via social media	1/3/12	3/24/16	20.00%
Monthly: Gather, edit and format content for internal (plain text) e-newsletter; distribute to	1/10/12	6/9/16	0.00%
Milestone: Ongoing - Generate monthly internal e-newsletter	1/10/12	6/9/16	20.00%
Monthly: Gather, edit and format content for external (HTML/designed) e-newsletter;	1/25/12	6/23/16	0.00%
Milestone: Ongoing - Generate monthly external e-newsletter	1/25/12	6/23/16	20.00%
Industry Relations	7/1/11	6/30/16	0.00%
Workforce Development	7/1/11	6/30/16	0.00%
Increase industry partners' awareness of all XSEDE SP's training opportunites	7/1/11	6/30/16	0.00%
Elicit industry partners' input to enhance training programs for workforce	7/1/11	6/30/16	0.00%
Software Development	7/12/12	7/10/14	0.00%
Hold conference call with XAB to flesh out software development activity	7/12/12	7/12/12	1.00%
Select and execute the software development project	7/13/12	7/10/14	0.00%
SD&I	9/1/11	6/28/13	0.00%

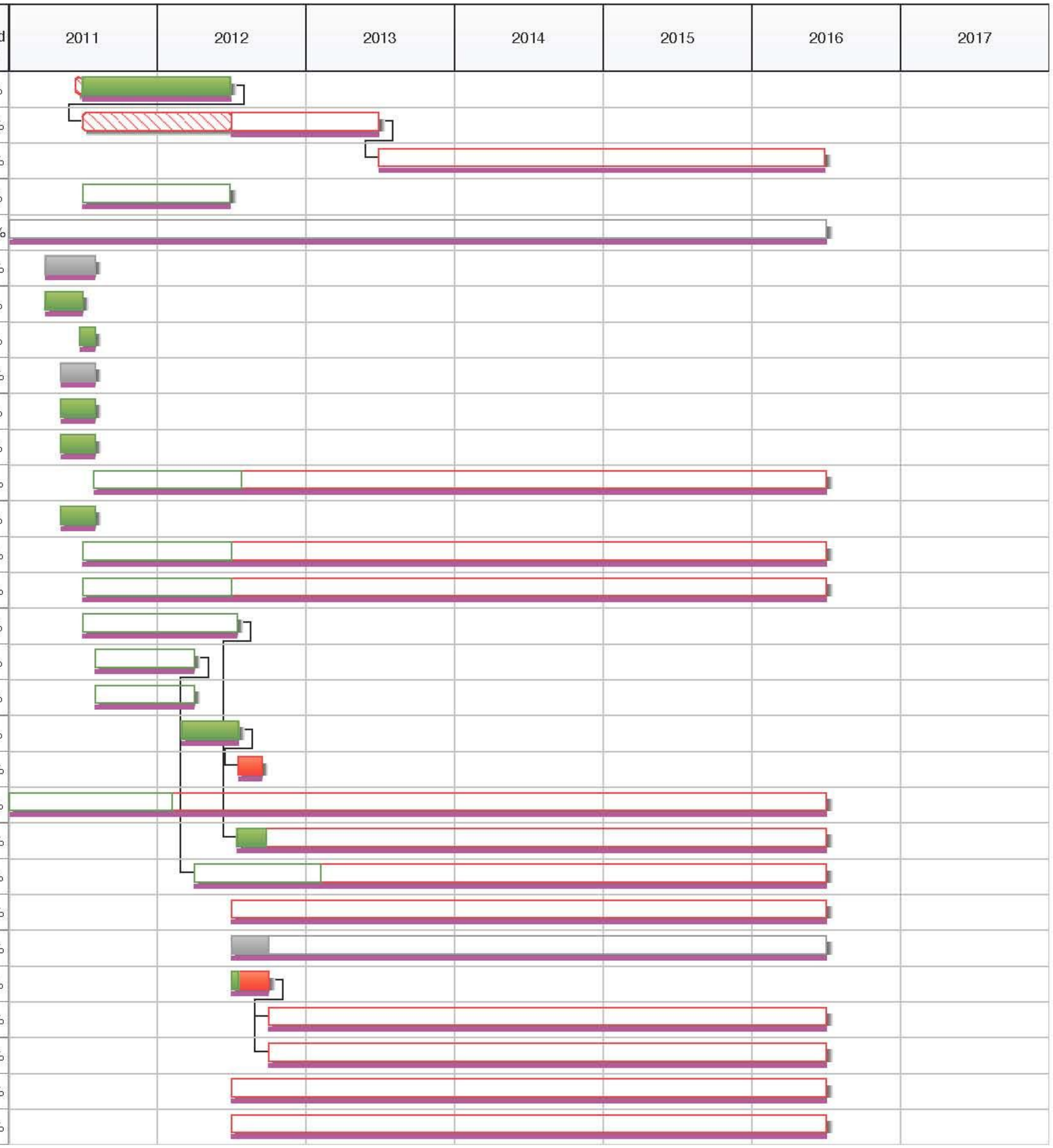
Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
PDR	9/1/11	9/1/11	100.00%							
Increment Planning	9/1/11	9/1/11	0.00%							
develop incrment plan	9/1/11	9/1/11	100.00%							
IRR	9/7/11	9/7/11	0.00%							
conduct IRR	9/7/11	9/7/11	100.00%							
IRR complete and passed	9/7/11	9/7/11	100.00%							
CI Detailed Design	9/8/11	9/8/11	0.00%							
develop detail design	9/8/11	9/8/11	0.00%							
GFFS	9/8/11	9/8/11	100.00%							
Execution Management	9/8/11	9/8/11	100.00%							
XUAS Data	9/8/11	9/8/11	100.00%							
CDR	9/9/11	9/15/11	0.00%							
conduct CDR	9/9/11	9/9/11	100.00%							
CDR complete and passed	9/15/11	9/15/11	100.00%							
CI Development	9/16/11	9/16/11	0.00%							
develop CI	9/16/11	9/16/11	0.00%							
GFFS	9/16/11	9/16/11	100.00%							
Execution Management	9/16/11	9/16/11	100.00%							
XUAS Data	9/16/11	9/16/11	100.00%							
CI TRR	9/19/11	10/7/11	0.00%							
conduct CI TRR	9/19/11	9/19/11	0.00%							
GFFS	9/19/11	9/19/11	100.00%							
Execution Management	9/19/11	9/19/11	100.00%							
XUAS Data	9/19/11	9/19/11	100.00%							
CI TRR complete and passed	10/7/11	10/7/11	100.00%							
CI Tests	10/10/11	12/15/11	0.00%							
conduct Ci tests	10/10/11	10/10/11	0.00%							
GFFS	10/10/11	10/10/11	100.00%							
Execution Management	10/10/11	10/10/11	100.00%							
XUAS Data	10/10/11	10/10/11	100.00%							

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
CI Tests complete and passed	12/15/11	12/15/11	100.00%							
⊟ STRR	12/16/11	12/16/11	0.00%							
conduct STRR	12/16/11	12/16/11	100.00%							
STRR complete and passed	12/16/11	12/16/11	100.00%							
⊟ System Integration Test	12/19/11	12/28/11	0.00%							
conduct system test	12/19/11	12/19/11	100.00%							
System test complete and passed	12/28/11	12/28/11	100.00%							
⊟ ORR	12/29/11	12/29/11	0.00%							
conduct ORR	12/29/11	12/29/11	100.00%							
ORR complete and passed	12/29/11	12/29/11	100.00%							
⊟ Increment De-brief	12/30/11	1/6/12	0.00%							
increment reflection workshop	12/30/11	12/30/11	100.00%							
reflection and practice report	1/2/12	1/2/12	100.00%							
de-brief complete	1/6/12	1/6/12	100.00%							
Implement Open, Continuous Planning	7/2/12	6/28/13	100.00%							
Implement Continuous Development and Integration	7/2/12	6/28/13	100.00%							
Implement Engineering Improvements	7/2/12	6/28/13	100.00%							
SDIACT-10 - Deliver Operational Tests with Cis	4/19/12	6/29/12	28.00%							
SDIACT-15 - Genesis II/UNICORE 6 GAML SAML	4/3/12	7/2/12	33.00%							
SDIACT-18 - Replicated/Synchronized stateful resource	4/18/12	6/18/12	100.00%							
SDIACT-28 - GO Transfer REST API as XSEDE Production service	4/25/12	7/2/12	17.00%							
SDIACT-31 - Improve GridFTP for SPs	4/25/12	7/2/12	60.00%							
SDIACT-43 - Genesis II Documentation	4/30/12	7/2/12	100.00%							
Activity 44. Campus bridging beta support	4/30/12	7/2/12	100.00%							
SDIACT-49 - Link Globus Online into XSEDE User Portal	4/25/12	7/2/12	90.00%							
SDIACT-50 - MyProxy OAuth Limited Proxy Support	4/20/12	7/2/12	33.00%							
SDIACT-54 - Register new increment 1 components	4/19/12	6/5/12	5.00%							
SDIACT-73 - System information publishing pilot	4/24/12	6/29/12	25.00%							
SDIACT-75 - GridFTP in UNICORE 6	5/1/12	7/2/12	33.00%							
SDIACT-96 - Identify XSEDE not TeraGrid in resource names for new OSG resource	3/22/12	5/9/12	100.00%							

Name	Start	Finish	% Completed - XSEDE
SDIACT-97 - Basic Execution Service	5/2/12	6/25/12	70.00%
SDIACT-100 - Globus Online Increment 2 addressing security concerns	4/25/12	7/2/12	3.00%
SDIACT-101 - EMS and GFFS Increment 2 updates	5/8/12	7/2/12	71.00%
Operations	1/3/11	1/6/17	0.00%
Cybersecurity	5/12/11	1/6/17	0.00%
Setup coordination of XSEDE incident response	5/12/11	7/1/11	100.00%
Setup and deploy XSEDE Certificate Authority	7/1/11	1/4/13	46.00%
Ongoing: Maintain Certificate Authority	1/4/13	1/6/17	20.00%
Develop and deploy Security Awareness program	7/1/11	6/27/13	100.00%
Ongoing: Maintain Security Awareness program	7/2/12	7/1/16	20.00%
Develop and deploy two factor authentication service	7/1/11	1/11/13	0.00%
Evaluate Implementation Options	7/1/11	5/4/12	100.00%
Deploy two factor authentication	5/4/12	1/11/13	0.00%
Ongoing: Maintain two factor authentication service	1/21/13	7/1/16	0.00%
Integrate and deploy InCommon Federated Authentication service	7/1/11	8/23/12	20.00%
Ongoing: Maintain InCommon Federated Authentication service	8/24/12	6/30/16	0.00%
Ongoing: Develop and implement improvements based on SEMP Spiral Design	7/1/11	6/30/16	20.00%
Ongoing: XSEDE Incident response	7/1/11	6/30/16	20.00%
Ongoing: Conduct on-demand security reviews for SD&I	7/2/12	6/30/16	15.00%
Ongoing: Conduct security reviews for ST&D	7/2/12	7/1/16	15.00%
Conduct XSEDE security risk assessment	7/2/12	12/21/12	100.00%
Setup XSEDE Nessus vulnerability assessment capability	9/3/12	12/14/12	100.00%
Setup secure wiki	7/2/12	8/24/12	100.00%
Install and update intrusion detection capability and security monitoring	9/3/12	2/15/13	0.00%
Obtain InCommon membership for XSEDE	3/1/12	8/14/12	100.00%
Conduct annual XSEDE security meeting	6/28/13	7/2/13	0.00%
Ongoing: Participate in XSEDE communications plan	7/1/11	6/30/16	20.00%
Prototype NSFv4 wide area file system and investigate security implications	1/1/13	6/17/13	0.00%
Data Services	7/1/11	6/29/16	0.00%
XSEDE Wide File System	7/4/11	6/29/16	0.00%



Name	Start	Finish	% Completed - XSEDE
Evaluate/Select global parallel file system(s)	7/4/11	6/28/12	100.00%
Deploy/extend global parallel file system(s)	6/29/12	6/27/13	0.00%
Ongoing: Maintain/extend global parallel file system(s)	6/28/13	6/29/16	0.00%
Evaluate and design archival replication framework	7/1/11	6/27/12	100.00%
XSEDENet Networking	**1/3/11**	**7/1/16**	**0.00%**
Transition current TG network to XSEDE network	4/1/11	7/29/11	0.00%
Execute NLR contracts	4/1/11	7/1/11	100.00%
Connect XSEDE sites to NLR	6/27/11	7/29/11	100.00%
xsede.org nameservice	5/9/11	7/29/11	0.00%
delegate subdomains to XSEDE partners: NICS, TACC, PSC	5/9/11	7/29/11	100.00%
establish DNS entries for non-subdomain names (network.xsede.org, etc.)	5/9/11	7/29/11	100.00%
Ongoing: Maintain and Monitor XSEDEnet	7/29/11	6/30/16	20.00%
Preserve the PSC speed page	5/9/11	7/29/11	100.00%
Ongoing: Tune end-to-end performance for user applications	7/1/11	6/30/16	20.00%
Ongoing: Develop and implement improvements based on SEMP Spiral Design	7/1/11	6/30/16	20.00%
Setup perfSONAR	7/1/11	7/13/12	100.00%
Provide Operations tools to monitor the network	8/1/11	3/30/12	100.00%
Replace network.teragrid.org functions	8/1/11	3/30/12	100.00%
Evaluate alternatives to NLR	3/1/12	7/18/12	100.00%
Execute network vendor contracts	7/19/12	9/12/12	0.00%
Ongoing: Maintain DNS entries for non-subdomain names (network.xsede.org,	1/3/11	6/30/16	20.00%
Ongoing: Maintain perfSONAR	7/13/12	7/1/16	5.00%
Ongoing: Maintain Operations tools to monitor the network	3/30/12	6/30/16	20.00%
Ongoing: Develop/Implement special ops team process for network related application	7/2/12	7/1/16	0.00%
XSEDEnet Metrics	7/2/12	7/1/16	0.00%
Determine metrics	7/2/12	9/28/12	20.00%
Ongoing: Collect data	10/1/12	7/1/16	0.00%
Ongoing: Report metrics (dashboard, quarterly report)	10/1/12	7/1/16	0.00%
SMNP Access, Layer 2 monitoring and measurement	7/2/12	7/1/16	0.00%
Develop Application Standards for Network Performance Metrics	7/2/12	7/1/16	0.00%



Name	Start	Finish	% Completed - XSEDE
Software Testing and Deployment	7/1/11	6/30/16	0.00%
Test XSEDE software	7/4/11	11/29/12	0.00%
Test Grid Middleware - EMS (Unicore and Genesis II client)	7/4/11	11/29/12	100.00%
Test Grid Middleware - GFSS (Genesis II)	7/4/11	11/29/12	100.00%
Test Data Movement - Globus Online	7/4/11	11/29/12	100.00%
Deploy XSEDE software	11/30/12	6/27/13	0.00%
Deploy Grid Middleware - Unicore	11/30/12	6/27/13	50.00%
Deploy Grid Middleware - Genesis II (campus bridging)	11/30/12	6/27/13	50.00%
Deploy Grid Middleware - Globus Online	11/30/12	6/27/13	50.00%
Ongoing: Coordinate operational reviews of next increment of XSEDE software on	7/1/11	6/30/16	20.00%
Ongoing: Test next increment of XSEDE Software on demand	7/4/11	6/30/16	20.00%
Ongoing: Deploy next increment of XSEDE software to SPs on demand	7/4/11	6/30/16	20.00%
Ongoing: Coordinate campus bridging deployments	7/4/11	6/30/16	20.00%
Ongoing: Develop deployment work plans and documentation	7/4/11	6/30/16	20.00%
Ongoing: Coordinate Globus Grid software support	7/4/11	6/30/16	20.00%
Accounting and Account Management	5/9/11	7/1/16	0.00%
Ongoing: Maintain existing accounting and account management databases	7/1/11	6/30/16	20.00%
Ongoing: Investigate and improve accounting/account management processes	7/1/11	6/30/16	20.00%
Ongoing: Participate in XSEDE communications plan	7/1/11	6/30/16	20.00%
Optimize and upgrade the XDCDB system	7/1/11	3/29/12	100.00%
Ongoing: Develop and implement improvements based on SEMP Spiral Design	7/1/11	6/30/16	20.00%
Make appropriate name changes to {item}.xsede.org	5/9/11	7/29/11	0.00%
domain name changes	5/9/11	7/29/11	100.00%
web site changes	5/9/11	7/29/11	100.00%
documentation changes	5/9/11	7/29/11	100.00%
Complete the 2nd phase of vetted/unvetted account creation effort	5/9/11	3/30/12	100.00%
Ongoing: Maintain and enhance XSEDE-wide reporting, decision-support and	7/1/11	7/1/16	20.00%
Streamline and modernize the allocation-request and account	7/2/12	7/1/13	0.00%
Re-work and simplify the proposal submission process	7/2/12	6/28/13	0.00%
Improve the proposal review and approval process	1/1/13	7/1/13	0.00%

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Provide enhanced A&AM administrative capabilties	1/1/13	7/1/13	0.00%							
Streamline account-request process	7/2/12	1/4/13	0.00%							
Enhance user-based information delivery (allocation/usage info, etc.)	7/2/12	10/5/12	0.00%							
Improve overall A&AM-related documentation and training	7/2/12	1/11/13	0.00%							
Establish processes to improve campus bridging and new SP integration	7/2/12	10/5/12	0.00%							
Ongoing: Support hardware/infrastructure integrity (server maintenace, UPS, etc.)	7/1/11	6/30/16	20.00%							
System Operational Support	5/9/11	7/1/16	0.00%							
Setup XSEDE Operations Center	6/1/11	6/28/11	100.00%							
Ongoing: Operate XSEDE Operations Center	7/1/11	6/30/16	20.00%							
Deploy centralized XSEDE cyberinfrasture servers/services	5/30/11	8/31/12	75.00%							
Ongoing: Maintain centralized XSEDE cyberinfrastructure servers/services	8/31/12	7/1/16	20.00%							
Upgrade XDCDB hardware and split database and AMIE parts	5/9/11	8/12/11	0.00%							
upgrade hardware at SDSC and operating system at PSC	5/9/11	8/12/11	100.00%							
migrate AMIE to stand alone VM server (PSC and SDSC)	5/9/11	6/17/11	100.00%							
Evaluate XDCDB hardware at PSC and determine if refresh needed	5/9/11	6/29/12	100.00%							
Evaluate current TG services and classify them into XSEDE HA tiers	5/9/11	6/28/12	100.00%							
Plan and schedule a semi-annual XDCDB failover test (SDSC to PSC)	5/9/11	6/28/12	100.00%							
TeraGrid ticket system transition	7/1/11	8/11/11	100.00%							
Evaluate/Deploy XSEDE Centralized monitoring software	11/1/11	6/21/13	100.00%							
Participate in Ticket System evaluation in conjunction with User Support	7/1/11	2/2/12	100.00%							
Transition Ticket System: Legacy to New	2/2/12	7/4/13	50.00%							
Backup XOC setup at IU	11/1/11	6/28/13	0.00%							
Documentation/Training for XOC Backup	11/1/11	6/28/13	0.00%							
XOC Failover Test	11/1/11	6/28/13	0.00%							
Operations Annual Report	6/1/12	6/28/12	0.00%							
Prepare operational metrics annual report/internal Operational Assessment	6/1/12	6/28/12	100.00%							
Milestone: Operational metrics annual report completed	6/28/12	6/28/12	100.00%							
Ongoing: Participate in XSEDE communications plan	7/1/11	6/30/16	20.00%							
Ongoing: Provide VM support for Centralized Services	7/1/11	6/30/16	20.00%							
Ongoing: Provide VM support for Science Gateway services	7/1/11	6/30/16	20.00%							

Name	Start	Finish	% Completed - XSEDE
Ongoing: Maintain VM server hardware	7/1/11	6/30/16	20.00%
Ongoing: Maintain other server hardware	7/1/11	6/30/16	20.00%
⊟ User Services	1/3/11	7/1/16	0.00%
⊟ Training	1/3/11	6/30/16	0.00%
Milestone: Develop guidelines for online and in-person training materials across XSEDE	3/1/11	3/28/11	0.00%
⊟ Milestone: Develop 10 Training Modules Annually	4/1/11	6/30/16	0.00%
⊟ Milestone: Develop 10 Training Modules Year 1	4/1/11	3/30/12	0.00%
Develop and Post 2 new online training tutorials	4/1/11	7/1/11	0.00%
Develop and Post 2 new online training tutorials	7/5/11	9/29/11	100.00%
Develop and Post 3 new online training tutorials	10/3/11	12/30/11	100.00%
Develop and Post 3 new online training tutorials	1/4/12	3/30/12	100.00%
⊟ Milestone: Develop 10 Training Modules Year 2	7/2/12	6/28/13	0.00%
Develop and Post 2 new online training tutorials	7/2/12	9/28/12	0.00%
Develop and Post 2 new online training tutorials	10/1/12	12/31/12	0.00%
Develop and Post 3 new online training tutorials	1/1/13	3/29/13	0.00%
Develop and Post 3 new online training tutorials	4/1/13	6/28/13	0.00%
⊟ Milestone: Develop 10 Training Modules Year 3	7/1/13	6/30/14	0.00%
Develop and Post 2 new online training tutorials	7/1/13	9/30/13	0.00%
Develop and Post 2 new online training tutorials	10/1/13	12/31/13	0.00%
Develop and Post 3 new online training tutorials	1/1/14	3/31/14	0.00%
Develop and Post 3 new online training tutorials	4/1/14	6/30/14	0.00%
⊟ Milestone: Develop 10 Training Modules Year 4	7/1/14	6/30/16	0.00%
Develop and Post 2 new online training tutorials	7/1/14	9/30/14	0.00%
Develop and Post 2 new online training tutorials	10/1/14	12/31/14	0.00%
Develop and Post 3 new online training tutorials	1/1/15	3/31/15	0.00%
Develop and Post 3 new online training tutorials	4/1/16	6/30/16	0.00%
⊟ Milestone: Develop 10 Training Modules Year 5	7/1/15	6/30/16	0.00%
Develop and Post 2 new online training tutorials	7/1/15	9/30/15	0.00%
Develop and Post 2 new online training tutorials	10/1/15	12/31/15	0.00%
Develop and Post 3 new online training tutorials	1/1/16	3/31/16	0.00%

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Develop and Post 3 new online training tutorials	4/1/16	6/30/16	0.00%							
Conduct 50 Training Sessions Annually	4/1/11	6/30/16	0.00%							
Milestone: Conduct 50 Training sessions Year 1	4/1/11	3/30/12	0.00%							
Conduct first 10 in-person or webcast training sessions	4/1/11	7/1/11	100.00%							
Conduct first 15 in-person or webcast training sessions	7/5/11	12/30/11	100.00%							
Conduct first 10 in-person or webcast training sessions	10/3/11	12/30/11	75.00%							
Conduct first 15 in-person or webcast training sessions	1/4/12	3/30/12	100.00%							
Milestone: Conduct 50 Training sessions Year 2	7/2/12	6/28/13	0.00%							
Conduct first 10 in-person or webcast training sessions	7/2/12	9/28/12	0.00%							
Conduct first 15 in-person or webcast training sessions	10/1/12	12/31/12	0.00%							
Conduct first 10 in-person or webcast training sessions	1/1/13	3/29/13	0.00%							
Conduct first 15 in-person or webcast training sessions	4/1/13	6/28/13	0.00%							
Milestone: Conduct 50 Training sessions Year 3	7/1/13	6/30/14	0.00%							
Conduct first 10 in-person or webcast training sessions	7/1/13	9/30/13	0.00%							
Conduct first 15 in-person or webcast training sessions	10/1/13	12/31/13	0.00%							
Conduct first 10 in-person or webcast training sessions	1/1/14	3/28/14	0.00%							
Conduct first 15 in-person or webcast training sessions	4/1/14	6/30/14	0.00%							
Milestone: Conduct 50 Training sessions Year 4	7/1/14	6/30/15	0.00%							
Conduct first 10 in-person or webcast training sessions	7/1/14	9/30/14	0.00%							
Conduct first 15 in-person or webcast training sessions	10/1/14	12/31/14	0.00%							
Conduct first 10 in-person or webcast training sessions	1/1/15	3/31/15	0.00%							
Conduct first 15 in-person or webcast training sessions	4/1/15	6/30/15	0.00%							
Milestone: Conduct 50 Training sessions Year 5	7/1/15	6/30/16	0.00%							
Conduct first 10 in-person or webcast training sessions	7/1/15	9/30/15	0.00%							
Conduct first 15 in-person or webcast training sessions	10/1/15	12/31/15	0.00%							
Conduct first 10 in-person or webcast training sessions	1/1/16	3/31/16	0.00%							
Conduct first 15 in-person or webcast training sessions	4/1/16	6/30/16	0.00%							
Milestone: Complete Federation of existing online training materials with XSEDE	1/3/11	12/30/11	0.00%							

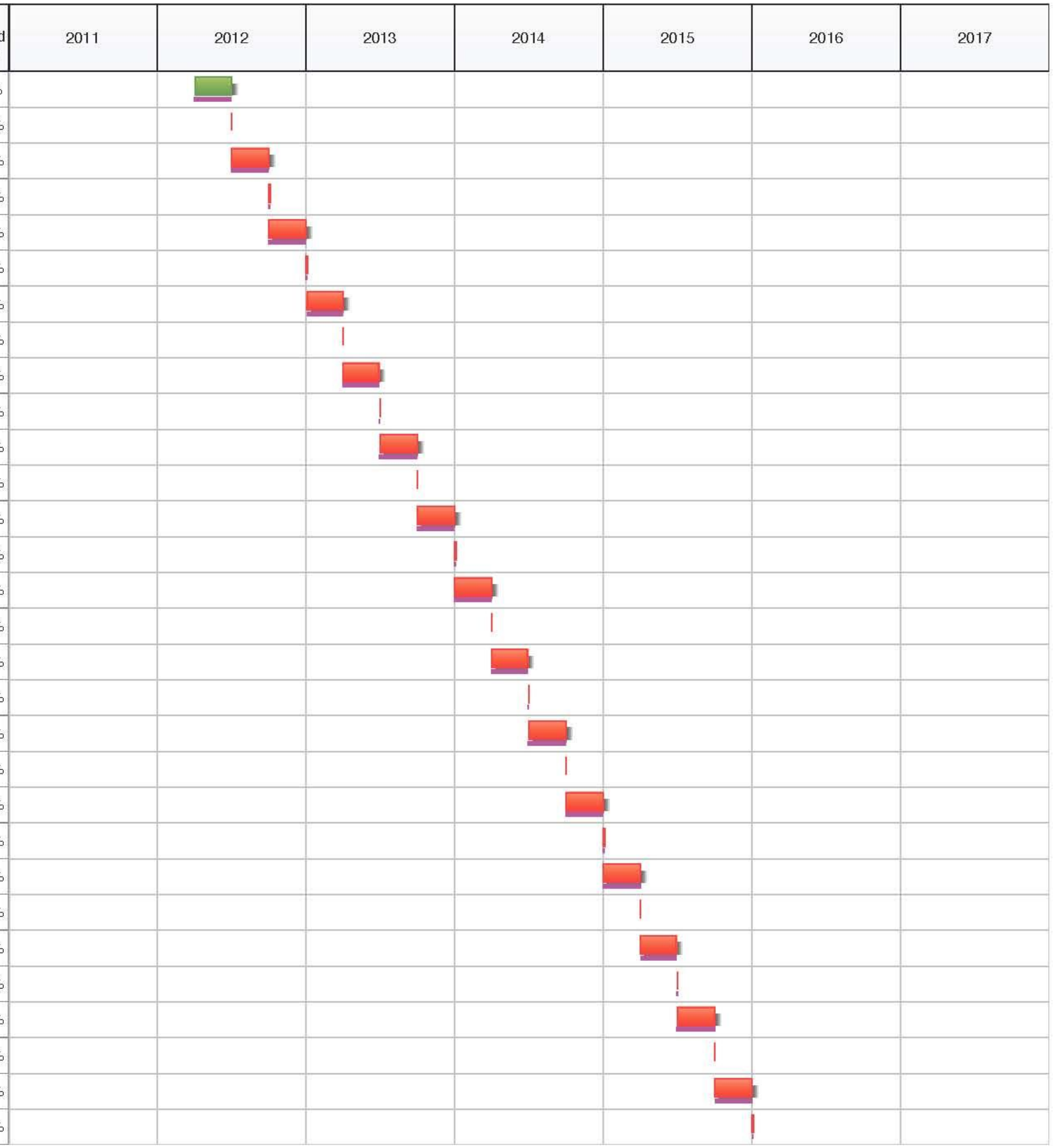
Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Milestone: Complete 2 targeted community workshops annually	7/1/11	6/30/16	0.00%							
Conduct first targeted community workshop Year 1	7/1/11	12/30/11	100.00%							
Conduct second targeted community workshop Year 1	1/2/12	6/29/12	100.00%							
Conduct first targeted community workshop Year 2	7/2/12	12/31/12	0.00%							
Conduct second targeted community workshop Year 2	1/1/13	6/28/13	0.00%							
Conduct first targeted community workshop Year 3	7/1/13	12/31/13	0.00%							
Conduct second targeted community workshop Year 3	1/1/14	6/2/14	0.00%							
Conduct first targeted community workshop Year 4	7/1/14	12/31/14	0.00%							
Conduct second targeted community workshop Year 4	1/1/15	6/30/15	0.00%							
Conduct first targeted community workshop Year 5	7/1/15	12/31/15	0.00%							
Conduct second targeted community workshop Year 5	1/1/16	6/30/16	0.00%							
Milestone: Conduct 4 technical training, content-based and mentoring webinars in	7/1/11	6/30/16	0.00%							
Conduct 4 webinars in support of XSEDE Scholars Program Year 1	7/1/11	6/29/12	100.00%							
Conduct 4 webinars in support of XSEDE Scholars Program Year 2	7/2/12	6/28/13	0.00%							
Conduct 4 webinars in support of XSEDE Scholars Program Year 3	7/2/13	6/30/14	0.00%							
Conduct 4 webinars in support of XSEDE Scholars Program Year 4	7/1/14	6/30/15	0.00%							
Conduct 4 webinars in support of XSEDE Scholars Program Year 5	7/1/15	6/30/16	0.00%							
Portfolio Review	1/3/11	1/3/11	100.00%							
Focus areas	1/3/11	12/30/11	0.00%							
Support of new systems (Stampede, Gordon, Keeneland, Blue Waters)	1/3/11	7/4/11	0.00%							
XSEDE architecture and tools	1/3/11	12/30/11	0.00%							
Security	1/3/11	7/4/11	0.00%							
Train the trainers	1/3/11	7/4/11	0.00%							
Trainiing for non-traditional areas	1/3/11	12/30/11	0.00%							
Portal to API for gateways	7/2/12	6/28/13	0.00%							
User Information Resources	**1/3/11**	**6/30/16**	**0.00%**							
Milestone: Release Production User Portal & Web Site	4/1/11	7/1/11	0.00%							
Transition existing production portal capabilities and web site	4/1/11	7/1/11	100.00%							
Define user information architecture	4/1/11	6/23/11	100.00%							
Milestone: Maintain Production User News System	4/1/11	7/1/11	0.00%							

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Transition existing user news system to production	4/1/11	7/1/11	100.00%							
Milestone: Release Allocation & User Guide for New and Transitioning Users	4/1/11	10/11/11	0.00%							
Document instructions for new users coming to XSEDE	4/1/11	7/1/11	100.00%							
Document instructions for existing users	7/4/11	9/5/11	100.00%							
Document allocation policies for resources	9/6/11	10/11/11	100.00%							
Milestone: Release new user guide with user comment capabilities	4/1/11	7/1/11	0.00%							
Create user guide template for HPC, Viz, Storage, etc.	4/1/11	6/23/11	100.00%							
Create user guide examples for each resource type	4/1/11	6/23/11	100.00%							
Ensure all user guides have been transitioned to template	4/1/11	7/1/11	100.00%							
Publish all user guides across web presence	4/1/11	7/1/11	100.00%							
Milestone: Production mobile user portal	4/1/11	6/30/11	0.00%							
Transition existing mobile framework	4/1/11	4/28/11	100.00%							
Evaluate requirements for mobile features	4/29/11	6/22/11	50.00%							
Create schedule for releasing future mobile features	4/29/11	6/30/11	50.00%							
Milestone: Release updated user news	7/1/11	9/22/11	0.00%							
Define requirements for user news system	7/1/11	7/28/11	100.00%							
Evaluate existing and alternative technologies	7/29/11	8/11/11	100.00%							
Release new user news system (if appropriate)	8/12/11	9/22/11	100.00%							
Milestone: Create new social media presence for XD	10/3/11	9/26/13	0.00%							
Define requirements for social media	10/3/11	9/3/12	50.00%							
Evaluate requirements that come out of User Engagement	10/31/11	9/26/13	0.00%							
Create twitter and facebook presence for users of XD and specifically XUP	7/2/12	12/31/12	0.00%							
Display twitter feeds on user portal	7/2/12	12/31/12	0.00%							
Milestone: Release collaborative capabilities to user portal	7/2/12	12/31/12	0.00%							
Define requirements based on User Engagement feedback	7/2/12	12/31/12	0.00%							
Enable users to be able to share calendars, chat, files, etc. with	7/2/12	12/31/12	0.00%							
Milestone: Release integrated training system	1/2/12	3/23/12	0.00%							
Define requirements for integrated training system based on TEOS and	1/2/12	1/27/12	100.00%							
Enable sites to post training courses on user portal	1/30/12	3/23/12	100.00%							
Enable sites to add online training resources to user portal	1/30/12	3/23/12	100.00%							

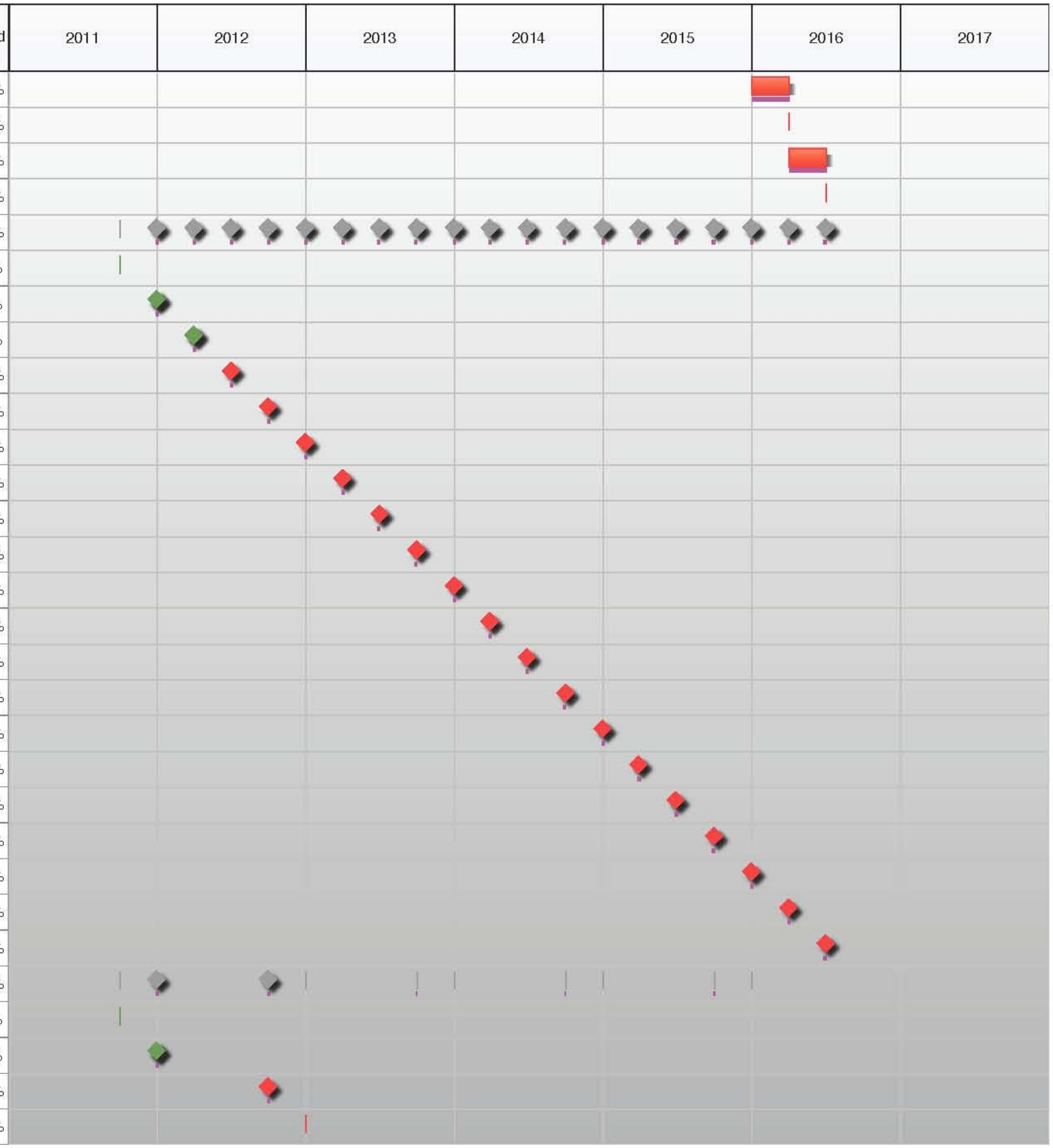
Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Enable users to register for training online via user portal	1/30/12	3/23/12	100.00%							
Create one stop shop for user training on user portal with calendar, SMS	1/30/12	3/23/12	100.00%							
Resource Selector	1/3/11	9/2/11	11.00%							
Ongoing: Develop and implement improvements based on SEMP Spiral Design	1/3/11	6/30/16	0.00%							
XSEDE User Portal	1/3/11	6/30/16	0.00%							
Redesign dock at the top and apply theme	1/2/12	3/2/12	100.00%							
Update profile portlet - expand with picture, publications, etc.	1/2/12	6/29/12	100.00%							
Link checker for XUP & XSEDE web site	1/2/12	3/30/12	0.00%							
Add forget username feature	1/2/12	4/30/12	100.00%							
Integrate future grid status in XUP system monitor	1/2/12	6/29/12	0.00%							
Expand system status beyond up/down/etc.	1/2/12	6/29/12	0.00%							
Implement new News categories	1/2/12	3/30/12	100.00%							
GridShib/InCommon integration	1/2/12	9/28/12	0.00%							
Enable dynamic feedback on each page	1/2/12	3/30/12	100.00%							
Migrate TGU staff queries to XUP staff area	1/2/12	3/30/12	75.00%							
Compete guest homepage redesign	1/2/12	6/29/12	0.00%							
Disable/gray out login link when resources are down	4/2/12	6/29/12	0.00%							
Merge add/remove user page and allocation page	7/2/12	6/30/16	0.00%							
Look at giving gateways a different 'view' for their community allocations	7/2/12	6/30/16	0.00%							
Merge DN listing with user profile	7/2/12	6/30/16	0.00%							
Integrate new ticketing system	7/2/12	12/31/12	0.00%							
Chat for help	7/2/12	6/30/16	0.00%							
Make hot links/bookmarking feature	7/2/12	6/30/16	0.00%							
Screen sharing with support staff	7/2/12	6/30/16	0.00%							
Expand allocations/usage/job history with graphs	7/2/12	6/30/16	100.00%							
Integrate XDMoD services	7/2/12	6/30/16	9.00%							
Custom views for communities: campus champions, gateways, fields	1/3/11	1/1/15	8.00%							
Network connectivity monitor	7/2/12	6/30/16	0.00%							
Dynamic visualization information in system monitor	7/2/12	6/30/16	0.00%							
Remote visualization services	7/2/12	6/30/16	0.00%							

Name	Start	Finish	% Completed - XSEDE
RSS feeds and SMS notifications for user portal functions	7/2/12	6/30/16	50.00%
OSG documentation integration	7/2/12	6/30/16	100.00%
Videos & interactive guide on features of XUP	7/2/12	6/30/16	0.00%
Online training for XUP	7/2/12	6/30/16	0.00%
Online survey for users on features	7/2/12	6/30/16	0.00%
Workflow management interface	7/2/12	6/30/16	0.00%
Metascheduling, job reservations, and ensemble job submission	7/2/12	6/30/16	0.00%
Change community account form to enter information in XCDB and	7/2/12	6/30/16	10.00%
Ongoing: Prioritize and integrate SD&I configuration items in US UII	1/3/11	12/31/15	20.00%
User Engagement	4/1/11	7/1/16	0.00%
Annual User Surveys	1/2/12	6/30/16	0.00%
Develop and Implement Y1 Annual User Survey	1/2/12	3/30/12	100.00%
Milestone: Y1 Annual User Survey report	7/2/12	7/2/12	0.00%
Develop and Implement Y2 Annual User Survey	1/1/13	4/1/13	0.00%
Milestone: Y2 Annual User Survey report	7/1/13	7/1/13	0.00%
Develop and Implement Y3 Annual User Survey	1/1/14	3/31/14	0.00%
Milestone: Y3 Annual User Survey report	6/30/14	6/30/14	0.00%
Develop and Implement Y4 Annual User Survey	1/1/15	3/31/15	0.00%
Milestone: Y4 Annual User Survey report	6/30/15	6/30/15	0.00%
Develop and Implement Y5 Annual User Survey	1/1/16	3/31/16	0.00%
Milestone: Y5 Annual User Survey report	6/30/16	6/30/16	0.00%
XSEDE CRM	1/2/12	7/2/12	0.00%
Document current XSEDE activities athat act as CRM	1/2/12	1/2/12	0.00%
Design and Implemnt initial CRM system	1/2/12	6/29/12	0.00%
Milestone: Initial CRM system deployed	7/2/12	7/2/12	0.00%
Data Mining	10/3/11	7/1/16	0.00%
Q1Y1 Data mining operations	10/3/11	12/30/11	100.00%
Milestone: Q1Y1 Data mining report	12/30/11	12/30/11	100.00%
Q2Y1 Data mining operations	1/2/12	3/30/12	100.00%
Milestone: Q2Y1 Data mining report	4/2/12	4/2/12	100.00%

Name	Start	Finish	% Completed - XSEDE
Q3Y1 Data mining operations	4/3/12	7/2/12	100.00%
Milestone: Q3Y1 Data mining report	7/2/12	7/2/12	0.00%
Q4Y1 Data mining operations	7/3/12	10/1/12	0.00%
Milestone: Q4Y1 Data mining report	10/2/12	10/2/12	0.00%
Q1Y2 Data mining operations	10/2/12	12/31/12	0.00%
Milestone: Q1Y2 Data mining report	1/1/13	1/1/13	0.00%
Q2Y2 Data mining operations	1/2/13	4/2/13	0.00%
Milestone: Q2Y2 Data mining report	4/2/13	4/2/13	0.00%
Q3Y2 Data mining operations	4/1/13	6/28/13	0.00%
Milestone: Q3Y2 Data mining report	7/1/13	7/1/13	0.00%
Q4Y2 Data mining operations	7/1/13	9/30/13	0.00%
Milestone: Q4Y2 Data mining report	10/1/13	10/1/13	0.00%
Q1Y3 Data mining operations	10/1/13	12/31/13	0.00%
Milestone: Q1Y3 Data mining report	1/1/14	1/1/14	0.00%
Q2Y3 Data mining operations	1/1/14	3/31/14	0.00%
Milestone: Q2Y3 Data mining report	4/1/14	4/1/14	0.00%
Q3Y3 Data mining operations	4/1/14	6/30/14	0.00%
Milestone: Q3Y3 Data mining report	7/1/14	7/1/14	0.00%
Q4Y3 Data mining operations	7/1/14	9/30/14	0.00%
Milestone: Q4Y3 Data mining report	10/1/14	10/1/14	0.00%
Q1Y4 Data mining operations	10/1/14	12/31/14	0.00%
Milestone: Q1Y4 Data mining report	1/1/15	1/1/15	0.00%
Q2Y4 Data mining operations	1/1/15	3/31/15	0.00%
Milestone: Q2Y4 Data mining report	4/1/15	4/1/15	0.00%
Q3Y4 Data mining operations	4/1/15	6/30/15	0.00%
Milestone: Q3Y4 Data mining report	7/1/15	7/1/15	0.00%
Q4Y4 Data mining operations	7/1/15	9/30/15	0.00%
Milestone: Q4Y4 Data mining report	10/1/15	10/1/15	0.00%
Q1Y5 Data mining operations	10/1/15	12/31/15	0.00%
Milestone: Q1Y5 Data mining report	1/1/16	1/1/16	0.00%



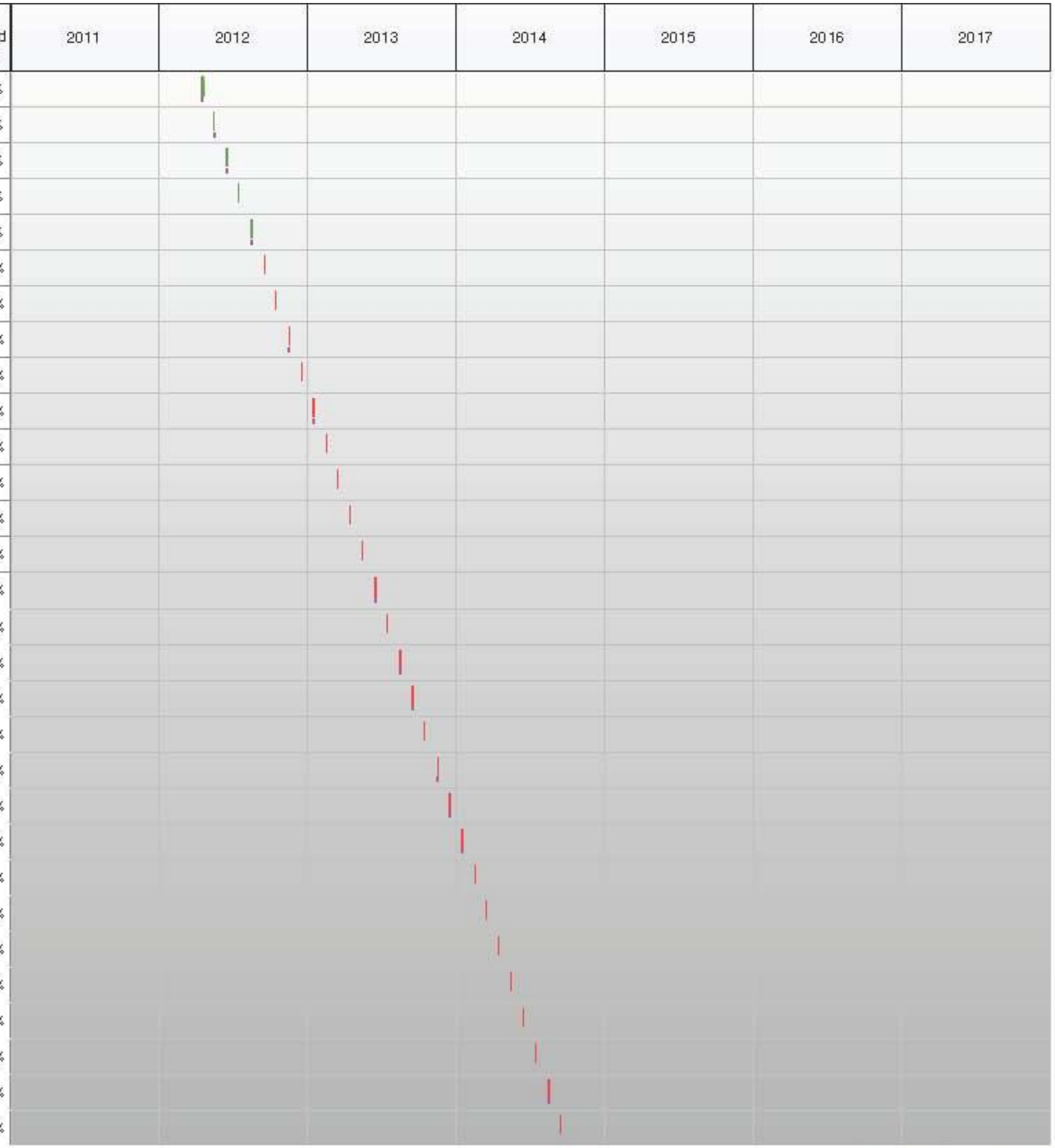
Name	Start	Finish	% Completed - XSEDE
Q2Y5 Data mining operations	1/1/16	3/31/16	0.00%
Milestone: Q2Y5 Data mining report	4/1/16	4/1/16	0.00%
Q3Y5 Data mining operations	4/1/16	6/30/16	0.00%
Milestone: Q3Y5 Data mining report	7/1/16	7/1/16	0.00%
Develop Focus Group Topics and conduct focus groups	9/30/11	6/30/16	0.00%
Milestone: Q1Y1 Focus Group Report	9/30/11	9/30/11	100.00%
Milestone: Q2Y1 Focus Group Report	1/2/12	1/2/12	100.00%
Milestone: Q3Y1 Focus Group Report	4/2/12	4/2/12	100.00%
Milestone: Q4Y1 Focus Group Report	7/2/12	7/2/12	0.00%
Milestone: Q1Y2 Focus Group Report	10/1/12	10/1/12	0.00%
Milestone: Q2Y2 Focus Group Report	12/31/12	12/31/12	0.00%
Milestone: Q3Y2 Focus Group Report	4/1/13	4/1/13	0.00%
Milestone: Q4Y2 Focus Group Report	7/1/13	7/1/13	0.00%
Milestone: Q1Y3 Focus Group Report	9/30/13	9/30/13	0.00%
Milestone: Q2Y3 Focus Group Report	12/31/13	12/31/13	0.00%
Milestone: Q3Y3 Focus Group Report	3/31/14	3/31/14	0.00%
Milestone: Q4Y3 Focus Group Report	6/30/14	6/30/14	0.00%
Milestone: Q1Y4 Focus Group Report	9/30/14	9/30/14	0.00%
Milestone: Q2Y4 Focus Group Report	12/31/14	12/31/14	0.00%
Milestone: Q3Y4 Focus Group Report	3/31/15	3/31/15	0.00%
Milestone: Q4Y4 Focus Group Report	6/30/15	6/30/15	0.00%
Milestone: Q1Y5 Focus Group Report	9/30/15	9/30/15	0.00%
Milestone: Q2Y5 Focus Group Report	12/31/15	12/31/15	0.00%
Milestone: Q3Y5 Focus Group Report	3/31/16	3/31/16	0.00%
Milestone: Q4Y5 Focus Group Report	6/30/16	6/30/16	0.00%
Develop and Conduct BoF Sessions	9/30/11	12/31/15	0.00%
Milestone: Conduct Y1 XSEDE BoF and Report	9/30/11	9/30/11	100.00%
Milestone: Conduct Y1 SC BoF and Report	12/31/11	1/2/12	100.00%
Milestone: Conduct Y2 XSEDE BoF and Report	9/30/12	10/1/12	0.00%
Milestone: Conduct Y2 SC BoF and Report	12/31/12	12/31/12	0.00%



Name	Start	Finish	% Completed - XSEDE
Milestone: Conduct Y3 XSEDE BoF and Report	9/30/13	9/30/13	0.00%
Milestone: Conduct Y3 SC BoF and Report	12/31/13	12/31/13	0.00%
Milestone: Conduct Y4 XSEDE BoF and Report	9/30/14	9/30/14	0.00%
Milestone: Conduct Y4 SC BoF and Report	12/31/14	12/31/14	0.00%
Milestone: Conduct Y5 XSEDE BoF and Report	9/30/15	9/30/15	0.00%
Milestone: Conduct Y5 SC BoF and Report	12/31/15	12/31/15	0.00%
Conduct Usability Panels and Testing (as needed)	4/1/11	3/31/16	20.00%
User Engagement General Operations	7/1/11	6/30/16	0.00%
Monitor tickets	7/1/11	6/30/16	20.00%
Resolve XSEDE wide tickets	7/1/11	6/30/16	20.00%
Bi-weekly User Engagement status meetings	7/1/11	6/30/16	20.00%
Quarterly reporting	7/1/11	6/30/16	20.00%
XSEDE Ticket System	4/1/11	6/29/12	0.00%
Transition existing TG ticket system to production in XSEDE	4/1/11	7/7/11	100.00%
Milestone: Release Production Ticketing System	7/7/11	7/8/11	100.00%
Deploy new/improved XSEDE Ticket System	7/1/11	6/29/12	0.00%
Define requirements for ticket system	7/1/11	8/12/11	100.00%
Evaluate candidate ticket systems	8/15/11	9/30/11	100.00%
Milestone: Select new/improved ticket system	10/1/11	10/3/11	100.00%
Develop and deploy new/improved ticket system	10/3/11	3/30/12	0.00%
Milestone: Release new/improved ticket system	3/31/12	4/2/12	0.00%
Integrate new/improved ticket system with CRM	4/2/12	6/29/12	0.00%
Milestone: Release Integrated CRM interface	6/29/12	6/29/12	0.00%
Integrate new/improved ticket system with XSEDE User Portal	4/2/12	6/29/12	0.00%
Milestone: Release Integrated XUP interface	6/29/12	6/29/12	0.00%
Consulting Policies	10/3/11	7/2/12	0.00%
Create consulting policies, procedures, and support guide	10/3/11	6/29/12	50.00%
Milestone: Deploy consulting policies, procedures, and support guide	7/2/12	7/2/12	0.00%
Contact PI at allocation start	7/15/11	4/15/16	0.00%
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	7/15/11	4/15/16	100.00%

Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	7/15/11	7/18/11	100.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	10/17/11	10/18/11	100.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	1/16/12	1/17/12	100.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	4/16/12	4/17/12	100.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	7/16/12	7/16/12	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	10/15/12	10/15/12	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	1/15/13	1/15/13	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	4/15/13	4/15/13	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	7/15/13	7/15/13	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	10/15/13	10/15/13	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	1/15/14	1/15/14	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	4/15/14	4/15/14	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	7/15/14	7/15/14	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	10/15/14	10/15/14	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	1/15/15	1/15/15	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	4/15/15	4/15/15	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	7/15/15	7/15/15	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	10/15/15	10/15/15	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	1/15/16	1/15/16	0.00%							
Contact "new"/"renewal" PIs at the beginning of each XRAC allocation	4/15/16	4/15/16	0.00%							
Contact "startup" PIs each month	7/15/11	6/15/16	0.00%							
1	7/15/11	7/18/11	100.00%							
2	8/15/11	8/16/11	100.00%							
3	9/15/11	9/16/11	100.00%							
4	10/17/11	10/18/11	100.00%							
5	11/15/11	11/16/11	100.00%							
6	12/15/11	12/16/11	100.00%							
7	1/16/12	1/17/12	100.00%							
8	2/15/12	2/16/12	100.00%							
9	3/15/12	3/16/12	100.00%							

Name	Start	Finish	% Completed - XSEDE
10	4/16/12	4/17/12	100.00%
11	5/15/12	5/16/12	100.00%
12	6/15/12	6/15/12	100.00%
13	7/16/12	7/16/12	100.00%
14	8/15/12	8/15/12	100.00%
15	9/17/12	9/17/12	0.00%
16	10/15/12	10/15/12	0.00%
17	11/15/12	11/15/12	0.00%
18	12/17/12	12/17/12	0.00%
19	1/15/13	1/15/13	0.00%
20	2/15/13	2/15/13	0.00%
21	3/15/13	3/15/13	0.00%
22	4/15/13	4/15/13	0.00%
23	5/15/13	5/15/13	0.00%
24	6/17/13	6/17/13	0.00%
25	7/15/13	7/15/13	0.00%
26	8/15/13	8/15/13	0.00%
27	9/16/13	9/16/13	0.00%
28	10/15/13	10/15/13	0.00%
29	11/15/13	11/15/13	0.00%
30	12/16/13	12/16/13	0.00%
31	1/15/14	1/15/14	0.00%
32	2/17/14	2/17/14	0.00%
33	3/17/14	3/17/14	0.00%
34	4/15/14	4/15/14	0.00%
35	5/15/14	5/15/14	0.00%
36	6/16/14	6/16/14	0.00%
37	7/15/14	7/15/14	0.00%
38	8/15/14	8/15/14	0.00%
39	9/15/14	9/15/14	0.00%

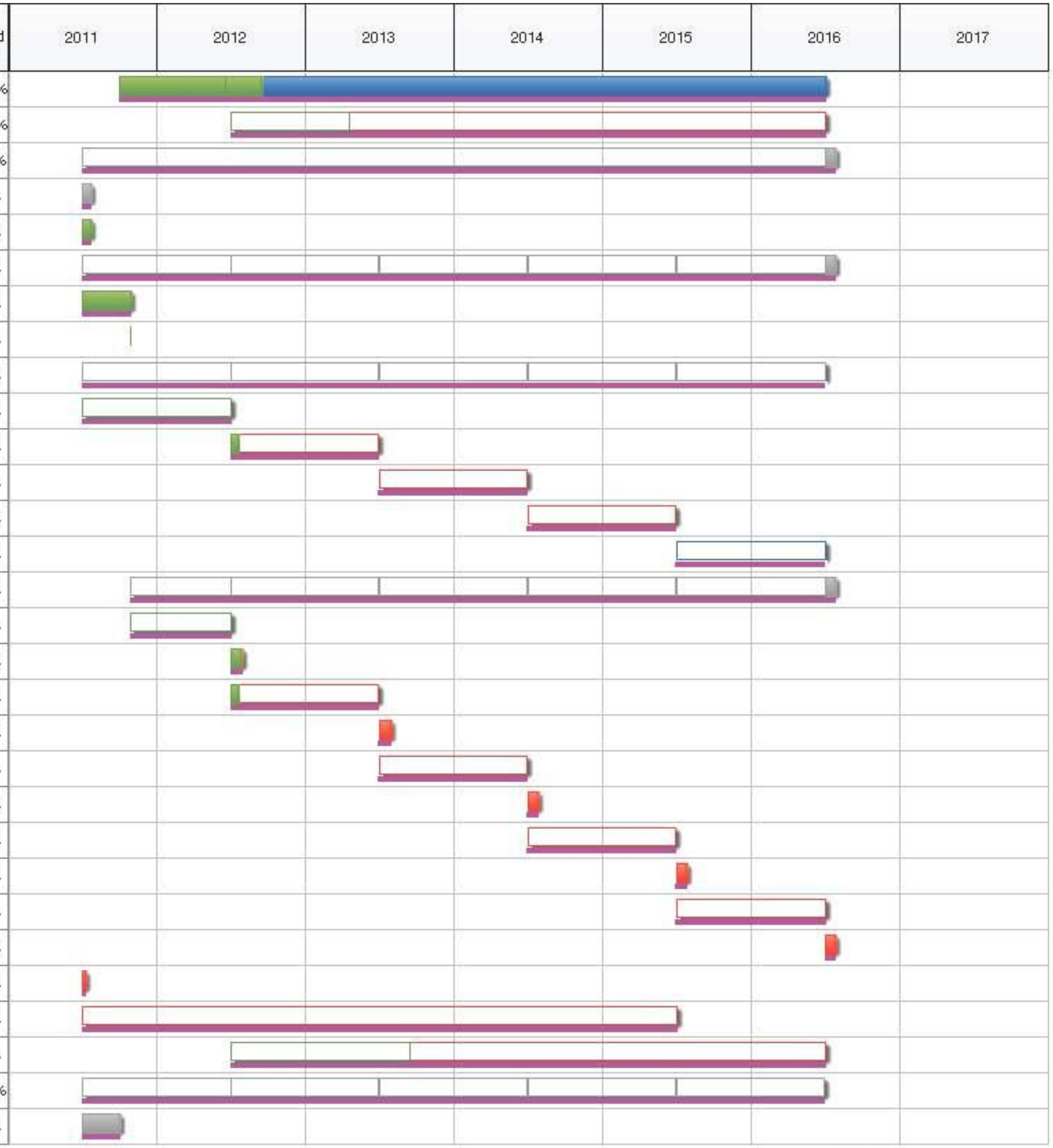


Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
40	10/15/14	10/15/14	0.00%							
41	11/17/14	11/17/14	0.00%							
42	12/15/14	12/15/14	0.00%							
43	1/15/15	1/15/15	0.00%							
44	2/16/15	2/16/15	0.00%							
45	3/16/15	3/16/15	0.00%							
46	4/15/15	4/15/15	0.00%							
47	5/15/15	5/15/15	0.00%							
48	6/15/15	6/15/15	0.00%							
49	7/15/15	7/15/15	0.00%							
50	8/17/15	8/17/15	0.00%							
51	9/15/15	9/15/15	0.00%							
52	10/15/15	10/15/15	0.00%							
53	11/16/15	11/16/15	0.00%							
54	12/15/15	12/15/15	0.00%							
55	1/15/16	1/15/16	0.00%							
56	2/15/16	2/15/16	0.00%							
57	3/15/16	3/15/16	0.00%							
58	4/15/16	4/15/16	0.00%							
59	5/16/16	5/16/16	0.00%							
60	6/15/16	6/15/16	0.00%							
Allocations	1/3/11	5/2/16	0.00%							
Allocations policy in place	4/1/11	7/1/11	100.00%							
Host Quarterly Allocations Meetings Annually	9/1/11	6/1/15	0.00%							
Host Year 1 Quarterly Allocations Meetings	9/1/11	6/1/12	0.00%							
Host Quarterly Allocations Meeting	9/1/11	9/2/11	100.00%							
Host Quarterly Allocations Meeting	12/1/11	12/2/11	100.00%							
Host Quarterly Allocations Meeting	3/1/12	3/2/12	100.00%							
Host Quarterly Allocations Meeting	6/1/12	6/1/12	100.00%							
Host Year 2 Quarterly Allocations Meetings	9/3/12	6/3/13	0.00%							

Name	Start	Finish	% Completed - XSEDE
Host Quarterly Allocations Meeting	9/3/12	9/3/12	0.00%
Host Quarterly Allocations Meeting	12/3/12	12/3/12	0.00%
Host Quarterly Allocations Meeting	3/4/13	3/4/13	0.00%
Host Quarterly Allocations Meeting	6/3/13	6/3/13	0.00%
Host Year 3 Quarterly Allocations Meetings	9/2/13	6/2/14	0.00%
Host Quarterly Allocations Meeting	9/2/13	9/2/13	0.00%
Host Quarterly Allocations Meeting	12/2/13	12/2/13	0.00%
Host Quarterly Allocations Meeting	3/3/14	3/3/14	0.00%
Host Quarterly Allocations Meeting	6/2/14	6/2/14	0.00%
Host Year 4 Quarterly Allocations Meetings	9/1/14	6/1/15	0.00%
Host Quarterly Allocations Meeting	9/1/14	9/1/14	0.00%
Host Quarterly Allocations Meeting	12/1/14	12/1/14	0.00%
Host Quarterly Allocations Meeting	3/2/15	3/2/15	0.00%
Host Quarterly Allocations Meeting	6/1/15	6/1/15	0.00%
Conduct How to Write a Successful Proposal Webcasts Annually	1/3/11	5/2/16	0.00%
Conduct Year 1 How to Write a Successful Proposal Webcast	8/1/11	6/1/12	0.00%
Conduct How to Write a Successful Proposal Webcast	8/1/11	9/1/11	100.00%
Conduct How to Write a Successful Proposal Webcast	11/1/11	12/1/11	100.00%
Conduct How to Write a Successful Proposal Webcast	2/1/12	3/1/12	100.00%
Conduct How to Write a Successful Proposal Webcast	5/1/12	6/1/12	100.00%
Conduct Year 2 How to Write a Successful Proposal Webcast	8/1/12	5/1/13	0.00%
Conduct How to Write a Successful Proposal Webcast	8/1/12	8/1/12	0.00%
Conduct How to Write a Successful Proposal Webcast	11/1/12	11/1/12	0.00%
Conduct How to Write a Successful Proposal Webcast	2/1/13	2/1/13	0.00%
Conduct How to Write a Successful Proposal Webcast	5/1/13	5/1/13	0.00%
Conduct Year 3 How to Write a Successful Proposal Webcast	1/3/11	2/3/14	0.00%
Conduct How to Write a Successful Proposal Webcast	8/1/13	8/1/13	0.00%
Conduct How to Write a Successful Proposal Webcast	11/1/13	11/1/13	0.00%
Conduct How to Write a Successful Proposal Webcast	2/3/14	2/3/14	0.00%
Conduct How to Write a Successful Proposal Webcast	1/3/11	1/3/11	0.00%

Name	Start	Finish	% Completed - XSEDE
Conduct Year 4 How to Write a Successful Proposal Webcast	8/1/14	5/4/15	0.00%
Conduct How to Write a Successful Proposal Webcast	8/1/14	8/1/14	0.00%
Conduct How to Write a Successful Proposal Webcast	11/3/14	11/3/14	0.00%
Conduct How to Write a Successful Proposal Webcast	2/2/15	2/2/15	0.00%
Conduct How to Write a Successful Proposal Webcast	5/4/15	5/4/15	0.00%
Conduct Year 5 How to Write a Successful Proposal Webcast	8/3/15	5/2/16	0.00%
Conduct How to Write a Successful Proposal Webcast	8/3/15	8/3/15	0.00%
Conduct How to Write a Successful Proposal Webcast	11/2/15	11/2/15	0.00%
Conduct How to Write a Successful Proposal Webcast	2/1/16	2/1/16	0.00%
Conduct How to Write a Successful Proposal Webcast	5/2/16	5/2/16	0.00%
New POPS Interface	2/13/12	7/2/12	0.00%
Design new interface	2/13/12	4/2/12	0.00%
Implement and test	4/2/12	6/29/12	0.00%
Deploy	7/2/12	7/2/12	0.00%
Add Tier 2 resources to allocation process	1/3/11	12/30/11	0.00%
Implement Allocation Levels and Types	1/3/11	6/28/13	0.00%
Levels	7/2/12	6/28/13	0.00%
Small	7/2/12	6/28/13	0.00%
Standard	7/2/12	6/28/13	0.00%
XRAC	7/2/12	6/28/13	0.00%
Types	1/3/11	6/28/13	0.00%
Storage	1/3/11	12/30/11	0.00%
Visualization	7/2/12	6/28/13	0.00%
Throughput	7/2/12	6/28/13	0.00%
Advanced Support for Research Teams (ECSS)	7/2/12	6/28/13	0.00%
GPU, MIC, other non-heterogeneous x86 compute	7/2/12	6/28/13	0.00%
Extended Collaborative Support Service (ECSS)-Projects	7/1/11	7/22/16	0.00%
Create/test proj. mgmt. framework for ECSS work plans/reporting	7/1/11	10/28/11	100.00%
Add at least 1 external FTE to fill an identified skills gap	7/1/11	6/30/16	100.00%
Fill 1.5 Discretionary Hires (as needed)	7/2/12	6/28/13	33.33%

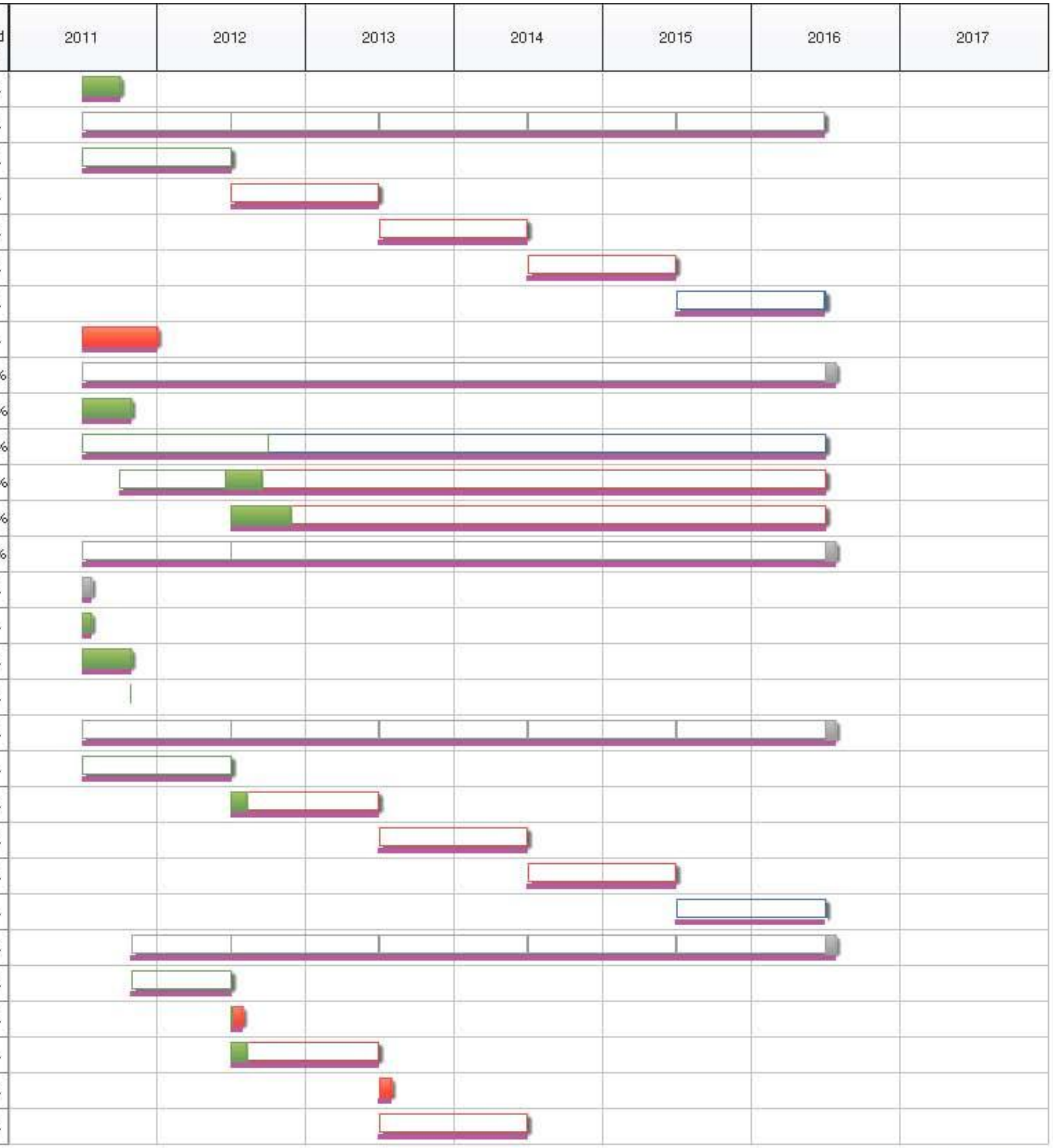
Name	Start	Finish	% Completed - XSEDE
Host monthly symposium open to ECSS, XSEDE staff, and Users	10/3/11	6/30/16	20.00%
Conduct continuing training of ECSS Staff as needed	7/2/12	6/30/16	20.00%
Extended Support for Research Team (ESRT)	7/1/11	7/22/16	0.00%
Establish ESRT group	7/1/11	7/21/11	0.00%
Set up ESRT staff and management teams and communications	7/1/11	7/21/11	100.00%
Milestone: Support 20 ESRT Projects Annually	7/1/11	7/22/16	0.00%
Work w/TG AUS to transition ASTA Projs. To ESRT mgmt.	7/1/11	10/28/11	100.00%
Milestone: All TG ASTA proj. managed as XD ESRT Projs.	10/28/11	10/28/11	100.00%
Work with XD CMS to Generate 20 new XD ESRT projs. Annually	7/1/11	6/30/16	0.00%
Year 1 Generate 20 new XD ESRT projects	7/1/11	6/29/12	100.00%
Year 2 Generate 20 new XD ESRT projects	7/2/12	6/28/13	5.00%
Year 3 Generate 20 new XD ESRT projects	7/1/13	6/30/14	0.00%
Year 4 Generate 20 new XD ESRT projects	7/1/14	6/30/15	0.00%
Year 5 Generate 20 new XD ESRT projects	7/1/15	6/30/16	0.00%
20 ESRT work plans documented and actively managed annually	10/31/11	7/22/16	0.00%
Year 1 - 20 work plans documented	10/31/11	6/29/12	100.00%
milestone: Y1 prepare and complete Final Reports	7/2/12	7/27/12	80.00%
Year 2 - 20 work plans documented	7/2/12	6/28/13	5.00%
milestone: Y2 prepare and complete Final Reports	7/1/13	7/26/13	0.00%
Year 3 - 20 work plans documented	7/1/13	6/30/14	0.00%
milestone: Y3 prepare and complete Final Reports	7/1/14	7/25/14	0.00%
Year 4 - 20 work plans documented	7/1/14	6/30/15	0.00%
milestone: Y4 prepare and complete Final Reports	7/1/15	7/24/15	0.00%
Year 5 - 20 work plans documented	7/1/15	6/30/16	0.00%
milestone: Y5 prepare and complete Final Reports	7/1/16	7/22/16	0.00%
Organize and execute HPC workshop at IEEE E-Science conference	7/1/11	7/7/11	0.00%
Host Annual Workshop on Petascale Computing	7/1/11	7/1/15	0.00%
Conduct XRAC Meetings Adaptive Reviews	7/2/12	6/30/16	30.00%
Novel & Innovative Projects	7/1/11	6/29/16	0.00%
Establish NIP group	7/1/11	9/30/11	0.00%



Sciforma 4.0

Thursday, August 23, 2012

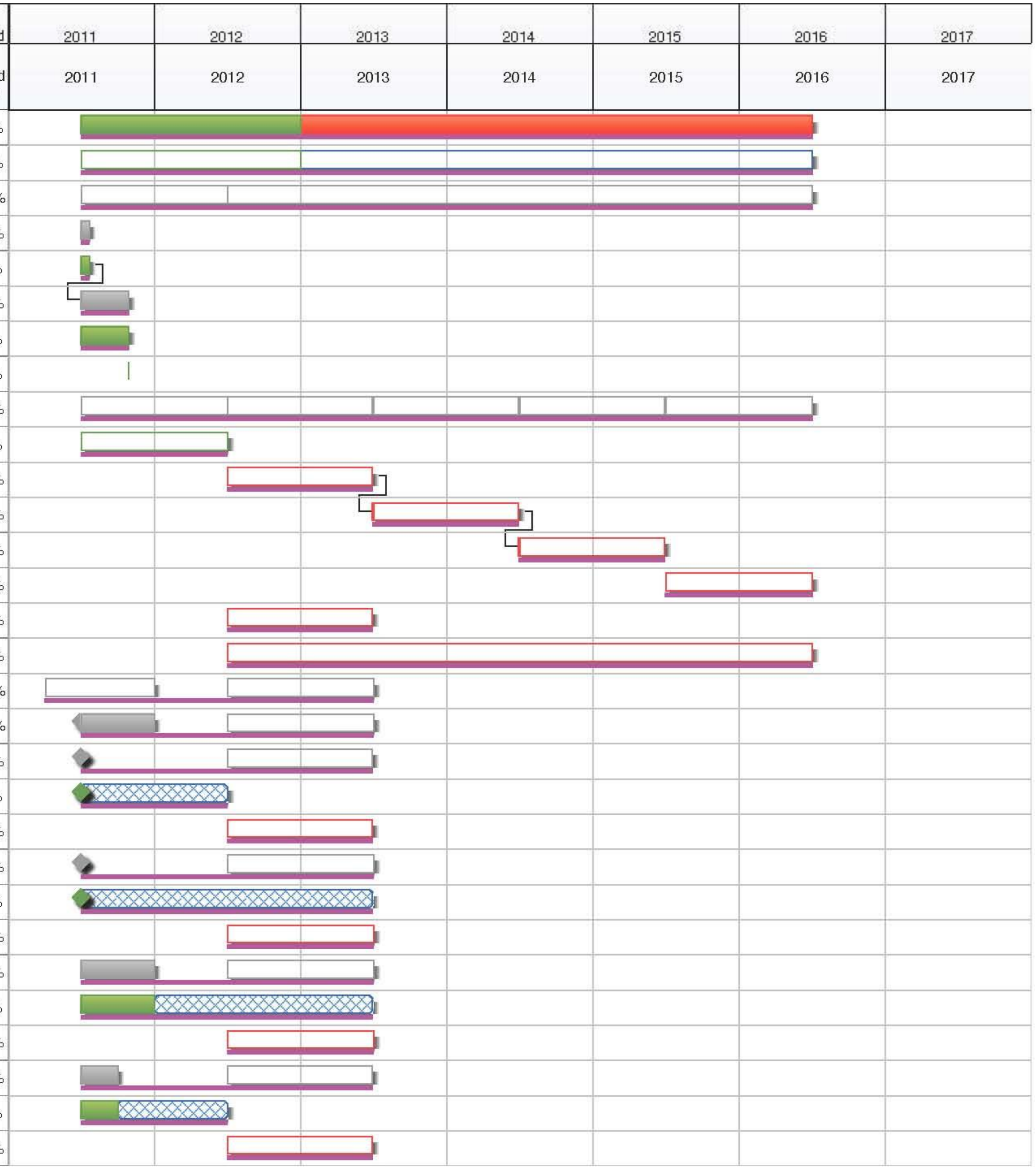
Name	Start	Finish	% Completed - XSEDE
Set up NIP staff and management teams and communications	7/1/11	9/30/11	100.00%
Milestone: Generate 20 new XSEDE+ ECS NIPs Annually	7/1/11	6/29/16	0.00%
Year 1- work w/XD CMS,TEOS,TIS to generate 20 XD NIP projects	7/1/11	6/29/12	100.00%
Year 2- work w/XD CMS,TEOS,TIS to generate 20 XD NIP projects	7/2/12	6/28/13	0.00%
Year 3- work w/XD CMS,TEOS,TIS to generate 20 XD NIP projects	7/1/13	6/30/14	0.00%
Year 4- work w/XD CMS,TEOS,TIS to generate 20 XD NIP projects	7/1/14	6/30/15	0.00%
Year 5- work w/XD CMS,TEOS,TIS to generate 20 XD NIP projects	7/1/15	6/29/16	0.00%
Milestone: Create ECSS Project work plans	7/1/11	12/30/11	0.00%
Extended Collaborative Support Service - Communities	7/1/11	7/22/16	0.00%
Create/test proj. mgmt. framework for ESCC work plans/reporting	7/1/11	10/28/11	100.00%
Add at least 1 external FTE to fill an identified skills gap	7/1/11	6/30/16	25.00%
Host monthly symposium for ECSS and XSEDE staff	10/3/11	6/30/16	20.00%
Conduct Continuing Training of ECSS Staff as needed	7/2/12	6/30/16	10.00%
Extended Support for Community Codes (ESCC)	7/1/11	7/22/16	0.00%
Establish ESCC group	7/1/11	7/21/11	0.00%
Set up ESCC staff and management teams and communications	7/1/11	7/21/11	100.00%
Transition TG ASP proj. to ESCC management	7/1/11	10/28/11	100.00%
Milestone:Active TG ASP proj. managed as XD ESCC proj.	10/28/11	10/28/11	100.00%
Milestone: Support 10 ESCC Projects Annually	7/1/11	7/22/16	0.00%
Year 1- Work w/TG CMS,TEOS,TIS to generate 10 XD ESCC proj.	7/1/11	6/29/12	100.00%
Year 2- Work w/TG CMS,TEOS,TIS to generate 10 XD ESCC proj.	7/2/12	6/28/13	10.00%
Year 3- Work w/TG CMS,TEOS,TIS to generate 10 XD ESCC proj.	7/1/13	6/30/14	0.00%
Year 4- Work w/TG CMS,TEOS,TIS to generate 10 XD ESCC proj.	7/1/14	6/30/15	0.00%
Year 5- Work w/TG CMS,TEOS,TIS to generate 10 XD ESCC proj.	7/1/15	6/30/16	0.00%
Milestone: Create 10 ESCC work plans Annually	10/31/11	7/22/16	0.00%
Y1 - 10 work plans documented	10/31/11	6/29/12	100.00%
milestone: Y1 prepare and complete Final Reports	7/2/12	7/27/12	10.00%
Y2 - 10 work plans documented	7/2/12	6/28/13	10.00%
milestone: Y2 prepare and complete Final Reports	7/1/13	7/26/13	0.00%
Y3 - 10 work plans documented	7/1/13	6/30/14	0.00%



Name	Start	Finish	% Completed - XSEDE	2011	2012	2013	2014	2015	2016	2017
milestone: Y3 prepare and complete Final Reports	7/1/14	7/25/14	0.00%							
Y4 - 10 work plans documented	7/1/14	6/30/15	0.00%							
milestone: Y4 prepare and complete Final Reports	7/1/15	7/24/15	0.00%							
Y5 - 10 work plans documented	7/1/15	6/30/16	0.00%							
milestone: Y5 prepare and complete Final Reports	7/1/16	7/22/16	0.00%							
Work with the TIS group to evaluate and recommend SI2 software projects for	7/2/12	6/30/16	0.00%							
Develop documentation, sample scripts, optimized builds to cover the top community	7/2/12	6/30/16	0.00%							
⊟ Extended Collaborative Support Service - Science Gateways	7/1/11	7/22/16	0.00%							
⊟ Establish ESSGW group	7/1/11	6/30/16	0.00%							
Set up ESSGW staff and management teams and communications	7/1/11	7/29/11	100.00%							
Organize bi-weekly gateway community calls open to all XSEDE users	9/1/11	6/30/16	100.00%							
⊟ Milestone: Support 10 Science Gateways Annually	7/1/11	7/22/16	0.00%							
Year 1- Work w/TG CMS,TEOS,TIS to generate 10 XD ESSGW proj.	7/1/11	6/29/12	100.00%							
Year 2- Work w/TG CMS,TEOS,TIS to generate 10 XD ESSGW proj.	7/2/12	6/28/13	10.00%							
Year 3- Work w/TG CMS,TEOS,TIS to generate 10 XD ESSGW proj.	7/1/13	6/30/14	0.00%							
Year 4- Work w/TG CMS,TEOS,TIS to generate 10 XD ESSGW proj.	7/1/14	6/30/15	0.00%							
Year 5- Work w/TG CMS,TEOS,TIS to generate 10 XD ESSGW proj.	7/1/15	6/30/16	0.00%							
⊟ Milestone: Create ESSGW work plans	10/31/11	7/22/16	0.00%							
⊟ At least 10 ESSGW work plans documented&actively managed	10/31/11	7/22/16	0.00%							
Y1 - 10 work plans documented	10/31/11	6/29/12	50.00%							
milestone: Y1 prepare and complete Final Reports	7/2/12	7/27/12	10.00%							
Y2 - 10 work plans documented	7/2/12	6/28/13	0.00%							
milestone: Y2 prepare and complete Final Reports	7/1/13	7/26/13	0.00%							
Y3 - 10 work plans documented	7/1/13	6/30/14	0.00%							
milestone: Y3 prepare and complete Final Reports	7/1/14	7/25/14	0.00%							
Y4 - 10 work plans documented	7/1/14	6/30/15	0.00%							
milestone: Y4 prepare and complete Final Reports	7/1/15	7/24/15	0.00%							
Y5 - 10 work plans documented	7/1/15	6/30/16	0.00%							
milestone: Y5 prepare and complete Final Reports	7/1/16	7/22/16	0.00%							
Gateway Outreach: Constantly reach out to new potential gateways independently and	7/1/11	6/30/16	30.00%							

Name Start Finish % Completed 2011 2012 2013 2014 2015 2016 2017

Name	Start	Finish	% Completed - XSEDE
XSEDE Requirements: Work with Gateway community in analyzing the XSEDE	7/1/11	6/30/16	30.00%
XSEDE Architecture Test Cases: Provide Test Cases to SD&I teams in nature of tests	7/1/11	6/30/16	30.00%
Extended Support for Training Education and Outreach (ESTEO)	7/1/11	6/30/16	0.00%
Establish ESTEO group	7/1/11	7/21/11	0.00%
Set up ESTEO staff and management teams and communications	7/1/11	7/21/11	100.00%
Milestone: Contribute content for TEO modules	7/1/11	10/28/11	0.00%
Work w/TG AUS to transition ASEOT projs. to XD ESTEO mgmt.	7/1/11	10/28/11	100.00%
Milestone: All TG ASEOT projs. Managed as XD ESTEO projs.	10/28/11	10/28/11	100.00%
Milestone: 50 ESTEO projects/activities supported Annually	7/1/11	6/30/16	0.00%
Year 1-work w/XD CMS, TEOS, TIS to generate 50 XD ESTEO	7/1/11	6/29/12	100.00%
Year 2- work w/XD CMS, TEOS, TIS to generate 50 XD ESTEO	7/2/12	6/28/13	0.00%
Year 3 - work w/XD CMS, TEOS, TIS to generate 50 XD ESTEO	7/1/13	6/30/14	0.00%
Year 4 - work w/XD CMS, TEOS, TIS to generate 50 XD ESTEO	7/1/14	6/30/15	0.00%
Year 5 - work w/XD CMS, TEOS, TIS to generate 50 XD ESTEO	7/1/15	6/30/16	0.00%
Test and Document initial work assignments for UCB CS class using XSEDE resouces	7/2/12	6/28/13	0.00%
Arrange ECSS Internal Training Seminars Annually	7/2/12	6/30/16	0.00%
Education and Outreach	4/4/11	7/1/13	0.00%
Education	7/1/11	7/1/13	0.00%
Milestone: 2 HPC Graduate level summer schools annually	7/1/11	6/28/13	0.00%
Milestone: 2 HPC Graduate level summer schools annually	7/1/11	7/1/11	100.00%
Milestone: 2 HPC Graduate level summer schools annually	7/2/12	6/28/13	0.00%
Milestone: 5 summer workshops annually	7/1/11	7/1/13	0.00%
Milestone: 5 summer workshops annually	7/1/11	7/1/11	100.00%
Milestone: 5 summer workshops annually	7/2/12	7/1/13	0.00%
Milestone: Add certificate programs at specific universities	7/1/11	7/1/13	0.00%
Milestone: ID univ to work with to dev vert pgm	7/1/11	12/30/11	100.00%
Milestone:Add cert and/or deg pgm @ univ and cont to ID univs for cert pgm	7/2/12	7/1/13	0.00%
Milestone: Provide online educational services	7/1/11	6/28/13	0.00%
Milestone: Provide online educational services	7/1/11	9/30/11	100.00%
Milestone: Provide online educational services	7/2/12	6/28/13	0.00%



Name	Start	Finish	% Completed - XSEDE
Collect Community Requirements	4/4/11	6/28/13	0.00%
Annual collection and analysis of community needs and requirements	4/4/11	11/16/11	100.00%
Collect Community Requirements	7/2/12	6/28/13	0.00%
Infratructure	**4/4/11**	**6/28/13**	**0.00%**
Curation of TEOS information on public web and XSEDE portal	4/4/11	7/26/11	100.00%
E&O Curation	7/2/12	6/28/13	75.00%
E&O Infrastructure Lead	7/2/12	6/28/13	0.00%
Campus Bridging	**4/4/11**	**6/28/13**	**0.00%**
Lead Campus Bridging Effort	4/4/11	6/28/13	0.00%
Ongoing: Share information with campuses interested in campus bridging	4/4/11	7/26/11	100.00%
Lead Campus Bridging Effort	7/2/12	6/28/13	0.00%
Campus Bridging Travel to Pilot Program Sites	7/2/12	6/28/13	0.00%
Pilot program, software packaging, documentation and suppport	7/2/12	6/28/13	0.00%
Rocks Roll test cluster 10GbE at Cornell	7/2/12	6/28/13	0.00%
Rocks Roll test cluster Infiniband at IU	7/2/12	6/28/13	0.00%
Lead Campus Bridging Effort	7/2/12	6/28/13	0.00%
External Evaluation	**4/4/11**	**6/28/13**	**0.00%**
External Evaluator Quarterly Reports	4/4/11	7/26/11	100.00%
External Evaluation	7/2/12	6/28/13	0.00%

2011 2012 2013 2014 2015 2016 2017

XSEDE

Project (Gantt Bar Styles) Default

Task (Gantt Bar Styles) Default

Critical Filter

Parent Filter

Buffers

Buffer Incursion 0% - 25%

Buffer Incursion 25% - 75%

Buffer Incursion 75% - 100%

Task (Gantt Symbols) Default

Gantt Legend

Completed Scheduled Baseline 1

Completed Scheduled Free Float Total Float Negative Float Start Delay Baseline 1

Scheduled

Scheduled

Scheduled

Buffer Incursion

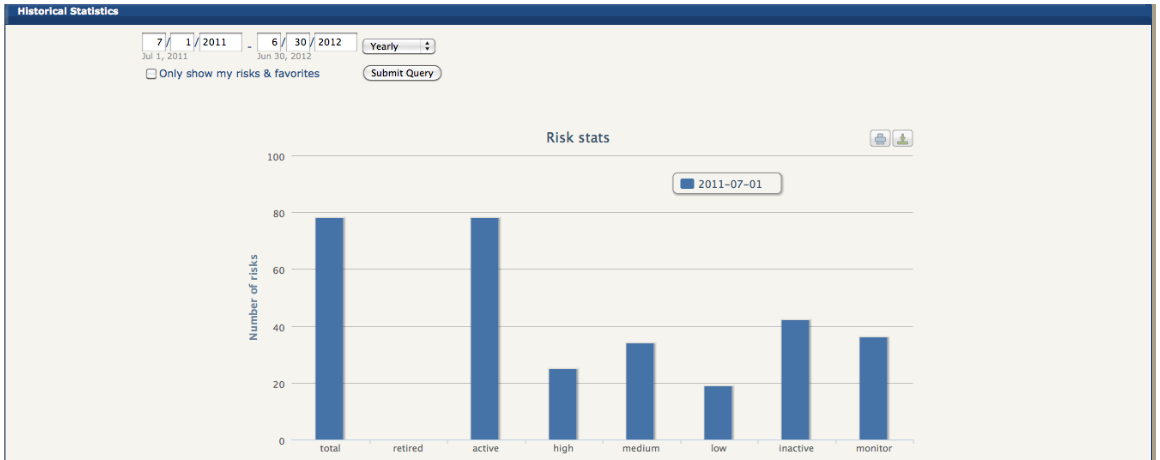
Buffer Incursion

Buffer Incursion


Objective Start
Objective Finish
Gantt Baseline2 Finish
Gantt Baseline3 Finish
Gantt Baseline4 Finish
Gantt Baseline5 Finish

C XSEDE Risk Register

Annual Summary of Risk



	Number of Risks			Risk Level			Risk Status		
	Total	Retired	Active	High	Medium	Low	Inactive	Monitor	Triggered
2011-07-01	78	0	78	25	34	19	42	36	
Changes	0	0	0	0	0	0	0	0	0
Most Recent Changes	78	0	78	25	34	19	42	36	0

Risk Id	Risk	Risk Level	Probability	Impact	Status	Subproject	Monitor Date	Owner
301	Opportunity: Leverage new/novel resources	High	Medium	High	Monitor	1.1 Project Office	Jul 01, 2011	John Towns
331	Program Plan	High	Medium	High	Monitor	1.1.1 Project Management and Reporting	Feb 16, 2012	Timothy Cockerill
279	Remote root compromise	High	Medium	High	Monitor	1.2.1 Security	Jul 01, 2011	Randal Butler
320	Security Compromise of XSEDE Infrastructure	High	Medium	High	Monitor	1.2.1 Security	Jul 01, 2011	Christopher Jordan
250	Suitable project management framework and process is not available.	High	Medium	High	Monitor	1.4.1 Extended Collaborative Research Teams Support	Jul 01, 2011	Mark Fahey

Risk Id	Risk	Risk Level	Probability	Impact	Status	Subproject	Monitor Date	Owner
274	Usage of deployed software and services	High	High	Medium	Monitor	1.2.4 Software Testing & Deployment	Jul 01, 2011	Troy Baer
242	XRAC Meeting Costs	High	Medium	High	Monitor	1.3.4 Allocations	Jul 01, 2011	Ken Hackworth
298	XSEDE archival storage gap	High	Medium	High	Monitor	1.2.2 Data Services	Jul 01, 2011	Christopher Jordan
252	Campus Infrastructure	Medium	Medium	Medium	Monitor	1.6.5 Campus Bridging	Jul 01, 2011	Craig Stewart
318	Communication Breakdown	Medium	Medium	Medium	Monitor	1.2.6 Systems Operational Support	Jul 01, 2011	Stephen McNally
325	Delay in program implementation	Medium	Medium	Medium	Monitor	1.6.1 Education	Sep 19, 2011	Steven Gordon
264	Failure of Security Systems and Procedures	Medium	Medium	Medium	Monitor	1.2.1 Security	Jul 01, 2011	Randal Butler
305	Federated identity management does not catch on	Medium	Medium	Medium	Monitor	1.1 Project Office	Jul 01, 2011	John Towns
271	Insufficient Service Provider Integration into XSEDE activities	Medium	Medium	Medium	Monitor	1.1 Project Office	Jan 01, 2011	John Towns
251	Mentoring Program	Medium	Medium	Medium	Monitor	1.6.2 Outreach	Jul 01, 2011	Laura McGinnis
277	Noncompliant service provider	Medium	Medium	Medium	Monitor	1.1 Project Office	Apr 01, 2011	Timothy Cockerill
253	Remote User Support	Medium	Medium	Medium	Monitor	1.6.5 Campus Bridging	Jul 01, 2011	Craig Stewart
322	Suitable project management framework and process is not available	Medium	Medium	Medium	Monitor	1.5.2 Extended Collaborative Science Gateways Support	Jul 01, 2011	Suresh Marru
255	Suitable project management framework and process is not available.	Medium	Medium	Medium	Monitor	1.5.1 Extended Support for Community Codes	Jul 01, 2011	John Cazes

Risk Id	Risk	Risk Level	Probability	Impact	Status	Subproject	Monitor Date	Owner
273	Grid Software Scaling	High	High	High	Monitor	1.1.3 Architecture & Design	Apr 01, 2011	Andrew Grimshaw
272	Grid Software System Integration	High	High	High	Monitor	1.1.3 Architecture & Design	Apr 01, 2011	Andrew Grimshaw
328	Campus Bridging expectation management	High	High	Medium	Monitor	1.6.5 Campus Bridging	Oct 01, 2012	Craig Stewart
296	Capacity of XSEDE reources less than demand	High	Medium	High	Monitor	1.1 Project Office	Jul 01, 2011	John Towns
278	Implementation delays and inconsistencies	High	High	Medium	Monitor	1.2.4 Software Testing & Deployment	Jul 01, 2011	Troy Baer
280	Local root compromise	High	Medium	High	Monitor	1.2.1 Security	Jul 01, 2011	Randal Butler
287	Loss of Project PI or co-PI	High	Medium	High	Monitor	1 XSEDE	Jan 01, 2011	John Towns
282	Malicious Social Engineering	High	Medium	High	Monitor	1.2.1 Security	Jul 01, 2011	Randal Butler
258	Mismatch between research teams' needs, XD resources, and AUSS staff availability.	High	Medium	High	Inactive	1.4 Extended Collaborative Support - Projects	Apr 01, 2011	Sergiu Sanielevici
312	Network Disruption Disables Critical Services	High	Medium	High	Monitor	1.2.3 XSEDEnet	Jul 01, 2011	Linda Winkler

Quarterly Report Risk Information

Non-retired Risks

72 risks

268	Software Partner Failure	Low	Low	Medium	Monitor	1.1.3 Architecture & Design	Apr 01, 2011	Andrew Grimshaw
249	Student Internships	Low	Medium	Low	Monitor	1.6.2 Outreach	Jul 01, 2011	Laura McGinnis
245	Training -AUSS support	Low	Low	Medium	Monitor	1.3.1 Training	Jul 01, 2011	Dan Stanzione
261	UNICORE project is cancelled	Low	Low	Medium	Monitor	1.1.3 Architecture & Design	Jan 01, 2011	Andrew Grimshaw
244	XD Architecture not fully implemented at an XD Service Provider	Low	Low	Medium	Inactive	1.1 Project Office	Apr 01, 2011	John Towns
262	XD Service Provider has insufficient resources to implement XSEDE Architecture	Low	Low	Medium	Monitor	1.2.4 Software Testing & Deployment	Jul 01, 2011	Troy Baer
329	Delay in ticketing system interface release	Low	Low	Low	Monitor	1.3.2 User Information & Interfaces	Jan 01, 2012	Maytal Dahan
247	Education Workshops	Low	Low	Low	Monitor	1.6.1 Education	May 01, 2011	Steven Gordon
269	Lack of Qualified System Administration Personnel	Low	Low	Low	Monitor	1.1.2 Systems and Software Engineering	Mar 01, 2011	Janet Brown
267	Software Complexity	Low	Low	Low	Monitor	1.2.4 Software Testing & Deployment	Jul 01, 2011	Troy Baer

	SPs							Cockerill
321	non-traditional users' needs / Mismatch between XSEDE services, and NIP Data allocation.	Low	Medium	Low	Monitor	1.4.2 Novel & Innovative Projects	Jul 01, 2011	Sergiu Sanielevici
254	Mismatch between research teams' needs, XD resources, and ECSS staff availability.	Medium	Low	High	Monitor	1.5.1 Extended Support for Community Codes	Jul 01, 2011	John Cazes
326	NLR Business Focus Changes	Low	Low	Medium	Monitor	1.2.3 XSEDEnet	Sep 01, 2011	Linda Winkler
256	Mismatch between research teams' needs, XD resources, and ECSS staff availability.	Medium	Low	High	Monitor	1.5.3 Extended Collaborative EOT Support	Jul 01, 2011	Jay Alameda
327	Namespace collisions between XSEDE software and local software	Low	Medium	Low	Monitor	1.2.4 Software Testing & Deployment	Oct 03, 2011	Troy Baer
319	Overloading XSEDE Operations Center Staff	Low	Medium	Low	Monitor	1.2.6 Systems Operational Support	Jul 01, 2011	Stephen McNally
248	Mismatch maintained between research teams' needs, XD resources, and ECSS staff availability.	Medium	Low	High	Monitor	1.4.1 Extended Collaborative Research Teams Support	Jul 01, 2011	Mark Fahey
306	Plan for long term service scaling may not meet short term needs	Low	Medium	Low	Monitor	1.1 Project Office	Jul 01, 2011	John Towns
311	NLR Services Are Inadequate	Medium	Low	High	Monitor	1.2.3 XSEDEnet	Jul 01, 2011	Linda Winkler
Risk Id	Risk	Risk Level	Probability	Impact	Status	Subproject	Monitor Date	Owner
283	Network Security and Operation	Medium	Low	High	Monitor	1.2.5 Security	Jul 01, 2011	Randal Butler
239	Potential equipment failure can disrupt database services.	Low	Low	Medium	Monitor	1.2.1 Accounting/Account Mgmt	Jul 01, 2011	Steven Quinn
316	Prohibitive Cost for Required Software	Medium	Low	High	Monitor	1.2.4 Software Testing & Deployment	Jul 01, 2011	Troy Baer
241	Prohibitive Operating Costs for Hardware or Software	Low	Low	Medium	Monitor	1.2.6 Systems Operational Support	Jul 01, 2011	Stephen McNally
286	Reduced overall budget for the project	Medium	Low	High	Monitor	1 XSEDE	Sep 15, 2010	John Towns
Risk Id	Risk	Risk Level	Probability	Impact	Status	Subproject	Monitor Date	Owner

D XSEDE Change Control Report

Nothing to report this quarter.

E Metrics

To demonstrate its success and help focus management attention on areas in need of improvement, XSEDE monitors a wide range of metrics in support of different aspects of "success" for the program. Since Q2 2012 marks the end of XSEDE's first program year, this appendix has a modified format. First, it discusses notable trends and changes in XSEDE's metrics over the course of the year (§E.1). Note that most tables and charts throughout this appendix cover all four quarters. After this section, the appendix follows the same format used in other quarterly reports, providing a consistent structure and continuous data points for historical purposes. The metrics presented in the quarterly reports provide a view into XSEDE's user community, including its success at expanding that community, the projects and allocations through which XSEDE manages access to resources, and the use of the resources by those projects (§E.2). In addition, XSEDE has identified metrics describing the program's success at delivering centralized services to this community, including operations, user support, advanced user support, and education and outreach activities (§E.3). Together, these metrics provide perspectives on how XSEDE works to ensure that the XSEDE-associated services and resources deliver science impact for the science and engineering research community.

E.1 Annual Metrics Highlights

In almost every metric, XSEDE experienced considerable growth over its first program year. From the size of the user community to the number of projects to the number of compute cycles delivered, XSEDE continued to see greater interest and associated use in its computational offerings. Similarly, XSEDE program activity and services grew to support the larger community. This growth included greater volumes of data movement via Globus Online, more Extended Collaborative Support projects, greater use of the user portal and web user interfaces, and more Campus Champions. At the same time, XSEDE managed to reduce the time it took to create user accounts, process allocation requests, and respond to user help requests.

E.1.1 XSEDE Resource and Service Usage Highlights

User community. The XSEDE user community absorbed significant growth in the program's first year. From 6,056 in Q3 2011, the number of open user accounts grew by 9.6% to 6,636 at the end of Q2 2012. An even larger increase of 14.2% was seen in the number of active user accounts, from 1,965 to 2,245. Gateway users consistently increase the number of active users each quarter by 33% or more, peaking at 1,580 gateway users in Q2 2012. Similarly, XSEDE's institutional breadth grew by 11.2% from 337 distinct institutions with active users to 375 institutions worldwide, with increasing participation from EPSCoR states and minority-serving institutions. During the year, 3,813 distinct users from 499 different institutions charged compute jobs, and 9,387 distinct user accounts were open at some point in the year.

XSEDE also succeeding in expanding its presence among a number of targeted communities. The number of active Campus Champion projects per quarter increased from 56 in Q3 2011 to 62 in Q2 2012 (+11%), the number of active institutions in EPSCoR states increased from 57 to 67 (+17.5%), and the number of active MSIs increased from 12 to 18 (+50%). XSEDE's reach included usage from 173 collaborating users at 105 institutions in 31 countries outside the U.S. (amounting to 4% of total usage).

Projects and allocations. The level of demand for XSEDE computational resources continued to outstrip the available resources. A number of metrics dipped in Q4 2011 after several TeraGrid resources were fully retired, but most numbers rebounded again. While NUs available at XRAC meetings rose from 13.4 billion to 17.9 billion (+33%), due primarily to the arrival of the Gordon system at SDSC, the resources were more than two times over-requested at every meeting except Q4 2012, where resources were only 166% over-requested. While the number of open projects

rose slightly, the number of active projects hit 1,027 in Q2 2012, breaking the thousand mark for the first time. The activity by field of science (Figure 32) remained relatively similar across the quarters, with nine areas responsibly for 94% of NUs delivered; however, the number of users in the 23 other fields of science accounted for nearly 30% of PIs and 35% of open user accounts. The top four "other" disciplines, each accounting for about 1% of usage, were Mathematical Sciences, Environmental Biology, Ocean Sciences, and Earth Sciences. In Q2 2012, usage by the "other" fields of science exceeded 6% of total NUs for the first time.

The XRAC faced a continued high level of requests—reviewing 157, 138, 123, and 147 requests at their four quarterly meetings—but still managed to make at least partial awards to an average of 91.5% of those requests. New awards accounted for a third of the XRAC awards from each meeting.

Resources and usage. At the end of its first program year, XSEDE encompassed 13 resources and 2.92 peak petaflops of computing performance. A total of 5.49 million jobs were reported, delivering 61.48 billion NUs over the course of the year. The decline in the number of resources was due to a number of TeraGrid systems retiring early in the XSEDE year, while the increased availability was due to the arrival of SDSC Gordon in early 2012. Despite the dip due to retiring systems, usage increased by 10% in XSEDE's first year over TeraGrid's final year.

The overall usage of XSEDE resources continues to balance between many projects running at smaller job scales (with "smaller" being fewer than 2,048 cores) and fewer projects running at larger scales. Consistent with the large-scale systems deployed at XSEDE Service Providers, most of the system usage is delivered to projects that need to run at large scale for at least some of their work. Figure 33 shows that the balance point for the program year is at 76/24.

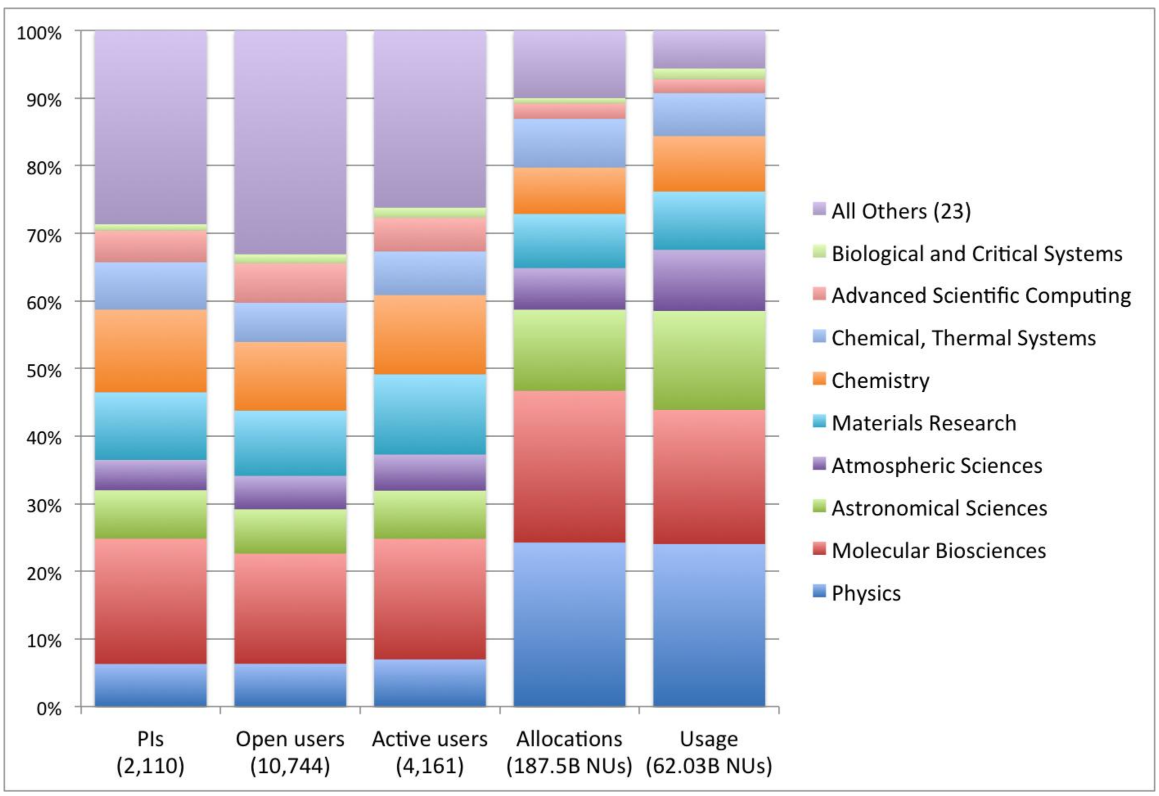


Figure 32. XSEDE Program Year 1 computing activity by field of science. PIs and users may appear in more than one field of science, so totals are higher than counts of unique PIs and users during the year.

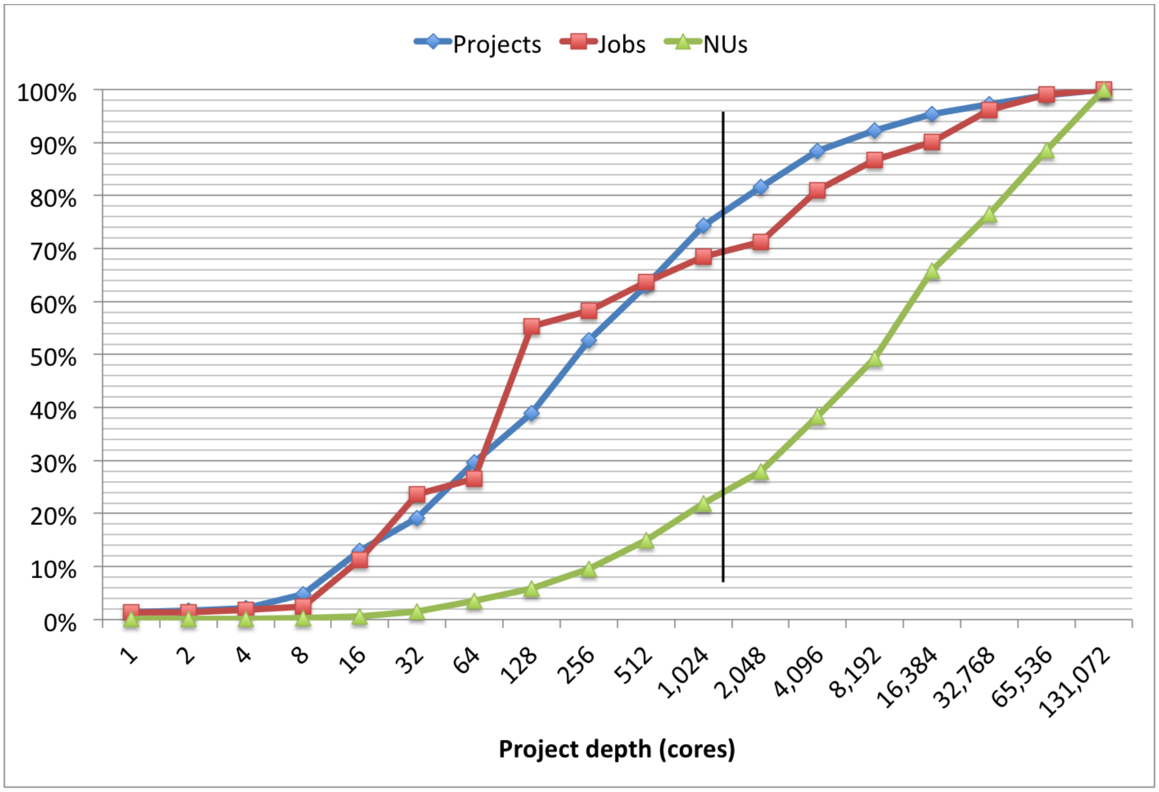


Figure 33. XSEDE project depth shows a 76/24 distribution: 76% of XSEDE projects need just over 1,024 cores at most to complete their work, but consume only 24% of the resources. The converse also holds true: 24% of projects use more than 1,024 for at least some of their work, and they consume 76% of the NUs delivered by XSEDE.

Delivering a growing number of NUs to a growing user base presents challenges to XSEDE Service Providers in terms of maintaining rapid turnaround for jobs and, in turn, high levels of user satisfaction. To that end XSEDE is monitoring some coarse metrics that provide a glimpse into overall system responsiveness across XSEDE. These include unweighted and weighted (by job core-hours) averages for wait time, response time, and slow down (or expansion factor). The weighted averages better reflect the reality that a small number of large-core-count jobs consume most of the delivered NUs (see Table 14). In general, these metrics show that the average job, in terms of resource consumption, ran for more than 20 hours and waited about the same amount of time on average. In Q2 2012, the weighted slow down increased and the weighted run time decreased to just over 20 hours, suggesting that XSEDE systems got busier in the last quarter of the program year as the systems delivered more NUs to a larger number of users.

E.1.2 <u>XSEDE Program Highlights</u>

In addition to tracking its user community and computing resource utilization, the XSEDE program is tracking a number of metrics about the program itself and its operational services in order to monitor how well XSEDE itself is functioning and identify situations requiring management attention.

Of note, the metrics indicate several highlights from around the program:

- XSEDE's security team needed to respond to very few security incidents in the first program year—none in 2012 so far.
- Globus Online was deployed as an official XSEDE service, and a growing number of users are moving hundreds of terabytes and millions of files each quarter.

- The Allocations team has worked to improve the response time for the various types of ongoing allocation requests. Notably, the time to process the average Startup request decreased from 13 days on average in Q3 2011 to nine days in Q1 and Q2 2012 (-31%).
- Several POPS and XSEDE User Portal components were deployed in a continuing effort to automate the user account creation process, which decreased the average time for creating an XSEDE user account from a high of 4.6 days in Q4 2011 to a 49 *minutes* in Q2 2012—*a 99% reduction*—due to the elimination of an entire manual processing step. This number is the time needed to add a user to an allocated project—from when the PI submits the request in the portal to when the authorized request is sent from POPS to the XDCDB (from which the request is subsequently distributed to Service Provider sites). The increase is such that XSEDE is considering a new metric or metrics to provide a finer-grained look at this key user support activity.
- While the number of tickets received at the XSEDE Operations Center each quarter remained relatively flat, the XSEDE support teams increased the number of tickets that received a response within 2 business days from 71% to 85% and the number that were closed within 2 business days from 33% to 39%.
- The Systems Operations team was responsible for transitioning more than 30 TeraGrid services to the XSEDE environment and domain in program year 1. Even though there were several planned and unplanned outages during the year, the SysOps team maintained high overall uptime, which ensured data integrity and availability. By leveraging failover resources where appropriate, downtime was greatly minimized. As such, no central service experienced any less than 98.48% uptime for PY1; of the 33 reporting services 31 had better than 99% uptime and 22 had better than 99.5% uptime.
- XSEDE's User Information and Interfaces group managed significant growth over the course of the year (140% more web hits, 95% more portal hits, 47% more portal users) as users made the transition from the TeraGrid to XSEDE's web site and portal. The number of Knowledgebase documents increased by 42% during the course of the year. Several new web and portal applications were added during the year—User News, Training Registration, and POPS Submission—and have been among the most frequently used parts of the XSEDE sites.
- The Extended Collaborative Support Service (ECSS) began tracking project metrics at the start of the program year and built on the set of projects underway at the end of TeraGrid. ECSS received 176 project requests over the year and was able to initiate 106 projects (60% of requests); 49 ECSS projects were completed during the year. Starting with 41 projects in Q3 2011, ECSS expanded to 100 active projects in Q2 2012—50 Research Team, 30 Community Code, and 20 Science Gateways. The Novel and Innovative Projects (NIP) staff conducted a number of outreach activities and grew to 9 active projects.

E.2 XSEDE Resource and Service Usage Metrics

Table 5 highlights a few key XSEDE measures that summarize the user community, the projects and allocations, and resource utilization for Q2 2012. Expanded information and five-year historical trends are shown in three corresponding subsections.

In Q2 2012, the XSEDE user community continued to grow, with 2,245 individuals representing 375 institutions charging compute jobs. XSEDE added 760 new users to its ranks and 29 fields of science report use. More details are in §E.2.1.

Project and allocation activity remained strong, with XSEDE resources requested at 209% of what was available. The quality of requests was also strong, with XRAC recommending support

for 129% of the available resources. During the quarter, the number of active projects climbed to 1,027, surpassing the 1,000 threshold for the first time. More details are in §0.

Table 5. Quarterly activity summary

User Community	Q3 2011	Q4 2011	Q1 2012	Q2 2012
Open user accounts	6,056	5,829	6,313	6,636
Active individuals	1,965	1,819	2,165	2,245
Gateway users	1,158	1,389	1,039	1,580
New user accounts	471	545	1,063	760
Active fields of science	28	26	26	29
Active institutions	337	316	340	375
Projects and Allocations				
NUs available at XRAC	13.398B	16.780B	17.377B	17.939B
NUs requested at XRAC	33.164B	28.012B	36.959B	37.539B
NUs recommended by XRAC	21.928B	16.268B	20.562B	23.194B
NUs awarded at XRAC	13.973B	16.268B	18.230B	17.502B
Open projects	1,588	1,560	1,545	1,600
Active projects	971	917	979	1,027
Active gateways	21	16	17	16
New projects	187	188	216	247
Closed projects	253	254	219	207
Resources and Usage				
Resources open (all types)	32	23	23	24
Total peak petaflops	2.55	2.50	2.92	2.92
Resources reporting use	19	15	14	13
Jobs reported	1.62M	1.13M	1.04M	1.70M
NUs delivered	14.84B	15.21B	14.76B	16.67B
Avg wtd run time (hrs)	22.6	22.5	23.3	20.9
Avg wtd wait time (hrs)	27.7	26.8	23.2	32.6
Avg wtd slow down	3.1	3.5	3.3	4.7

XSEDE computing resources held steady at 2.9 Pflops (peak) at the end of the quarter. The central accounting system showed 13 resources reporting activity, and together they delivered 16.67 billion NUs of computing. This represents an increase of approximately 13% over the previous quarter, due largely to Gordon being in production for the full quarter. However, XSEDE users experienced longer wait times, on average, according to several metrics. More details are in §E.2.3.

E.2.1 User community metrics

Figure 34 shows the five-year trend in the XSEDE user community, including open user accounts, total active XSEDE users, active individual accounts, active gateway users, and the number of new accounts during the quarter. New with this quarter's report, we have included the total number of unvetted user accounts (i.e., portal-only accounts) at the end of each quarter since that capability was added. Such accounts can be used for training course registration and other functions. The numbers of individual users, both open and active accounts, rebounded significantly and set new record levels after two quarters of decline. The number of gateway users also jumped to 1,580, an increase of more than 500. The 760 new user accounts created represents a strong but not record level.

Figure 35 shows the activity on XSEDE resources according to field of science, including the relative fraction of PIs, open accounts, active users, allocations, and NUs used according to discipline. For consistency across quarters, we show (and will show in future quarters) the nine fields of science that typically consume ~2% or more of delivered NUs per quarter. PIs and users are counted more than once if they are associated with projects in different fields of science. The Q2 data show that the percentages of PIs and accounts associated with the "other" disciplines represent nearly 30% of all PIs and more than 30% of user accounts, and nearly 25% of active

users. Collectively the "other" fields of science represent 6.2% of total quarterly usage, a modest increase from the previous quarter and the high point thus far during the XSEDE program.

Table 6 and Table 7 highlight aspects of the broader impact of XSEDE. The former shows that graduate students, post-doctoral researchers, and undergraduates make up 65% of the XSEDE user base. The latter table shows XSEDE's reach into targeted institutional communities, including a 39% increase in usage EPSCoR states over the previous quarter. Institutions with Campus Champions represent a large portion of XSEDE's usage (this table shows all users at Campus Champion institutions, not just those on the champion's project). The table also shows XSEDE's reach into EPSCoR states, the MSI community, and internationally. Also of note for Q2, the number of institutions and users in the EPSCoR and MSI categories all increased.

Table 6. End of quarter XSEDE open user accounts by type, excluding XSEDE staff.

Category	**Q3 2011**	**Q4 2011**	**Q1 2012**	**Q2 2012**
Graduate Student	2,368	2,275	2,466	2,574
Faculty	1,336	1,299	1,293	1,344
Postdoctorate	1,008	1,002	1,059	1,075
Undergraduate Student	505	416	587	639
University Research Staff (excluding postdocs)	506	501	509	535
High school	5	5	5	4
Others	328	331	394	465
TOTALS	***6,056***	***5,829***	***6,313***	***6,636***

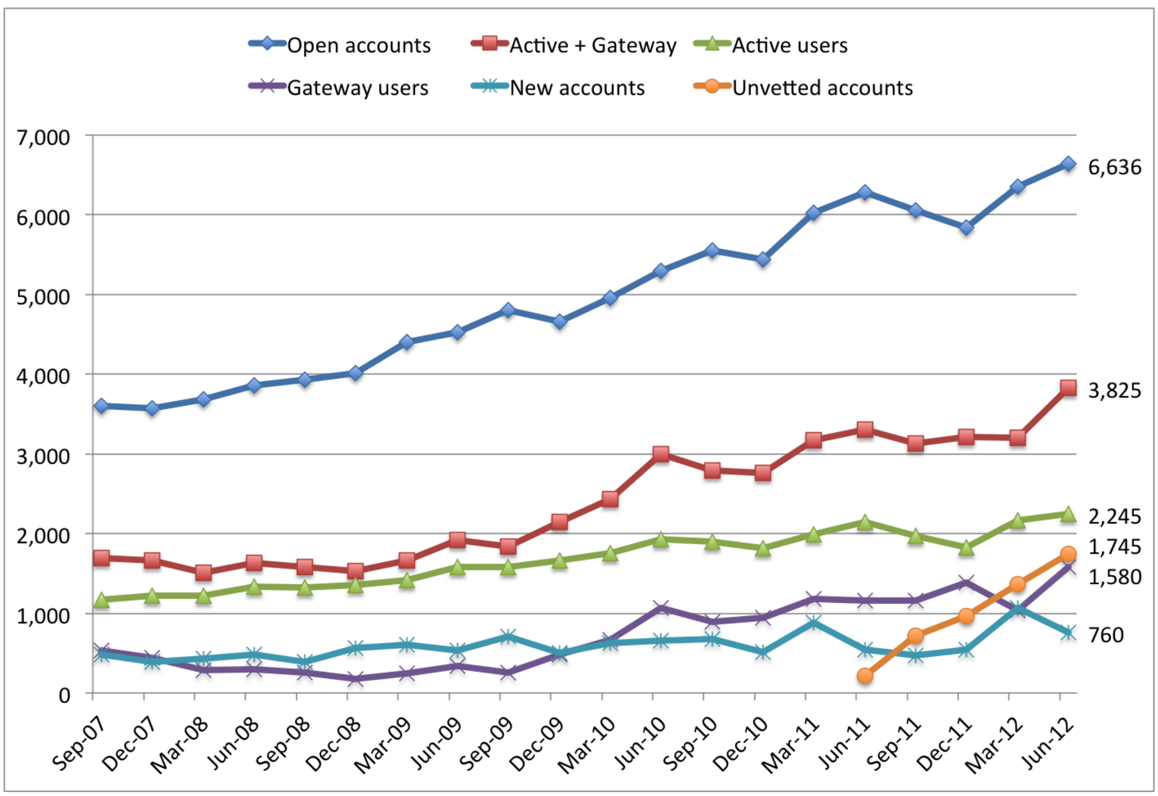


Figure 34. XSEDE user census, excluding XSEDE staff.

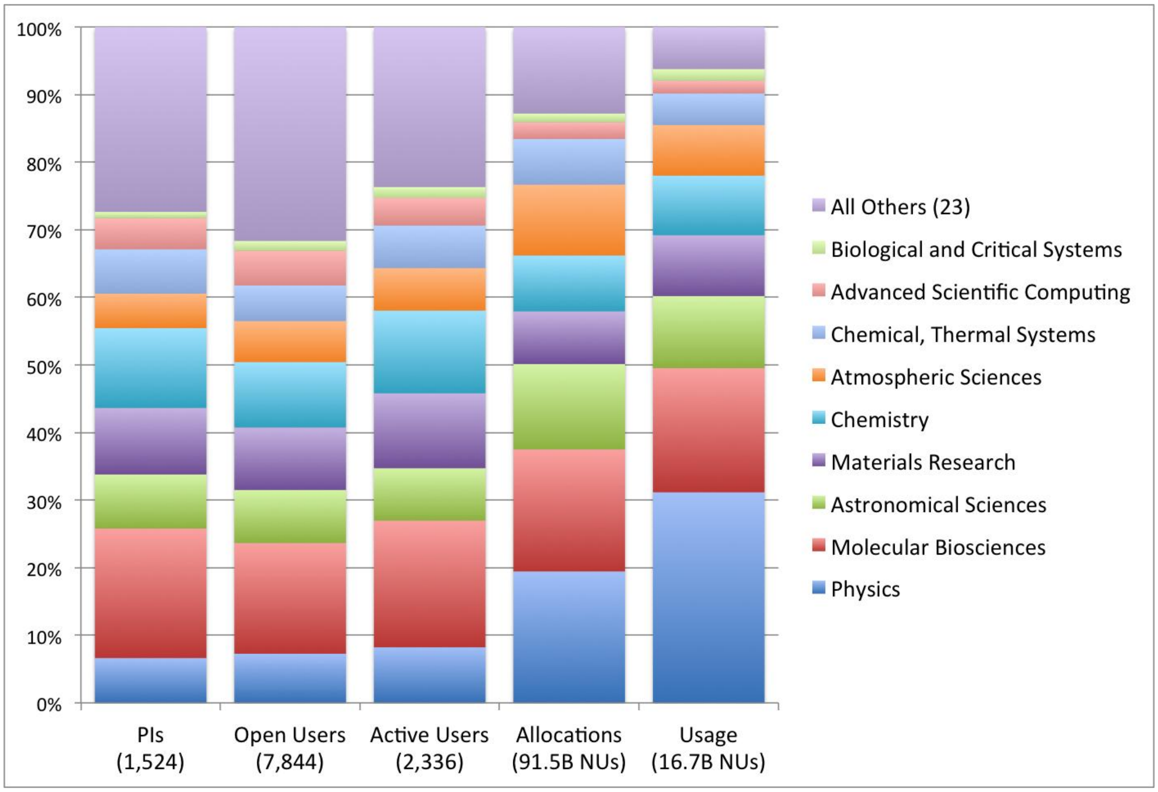


Figure 35. Quarterly XSEDE user, allocation, and usage summary by field of science, in order by usage, excluding staff projects. Note: PIs, users may appear under more than one field of science.

Table 7. Active institutions, overall and in selected categories. Notes: "Total" reflects institutions not in any specially designated category. Institutions may be in more than one category.

Category	Institutions	Users	NUs	% NUs
Campus Champion site	62	795	7,526,458,642	45%
EPSCoR state	67	318	3,306,768,444	20%
MSI	18	40	114,577,771	1%
International	63	86	543,798,072	3%
Total	375	2,245	16,670,381,984	100%

E.2.2 Project and allocation metrics

Figure 36 shows the historical trends for requests and awards at XSEDE quarterly allocation meetings. The figure shows the continued strong growth in demand against the relative plateau in available resources; NUs requested were 209% of NUs available, while the XRAC recommended awarding almost 30% more NUs than were available to the 147 requests for resources. Table 8 presents a summary of overall project activity. Notably, 90% of XRAC requests received an award, and 39% were new awards.

Table 9 and Table 10 show projects and activity in key project categories—Campus Champion, Staff, and Science Gateways—and by allocation board type. (Science Gateways may appear under any board.) Table 9 also shows new and closed projects for the quarter, while Table 10 shows the number of open and active user accounts with each type of project. Table 11 shows detailed information about allocations activity for the various request types available for the different classes of projects. Notably, XSEDE had 147 Research (XRAC) requests, of which 133 (90%) received awards, including 52 new projects. There were also 206 Startup requests, of

which 169 (82%) received awards; 12 Education requests with 12 awards; and 22 Campus Champion requests with 21 awards.

As a special class of projects, science gateway activity is detailed in Figure 37 showing continued high levels of usage and users from these projects. Table 12 shows gateway activity supported by specific XSEDE resources.

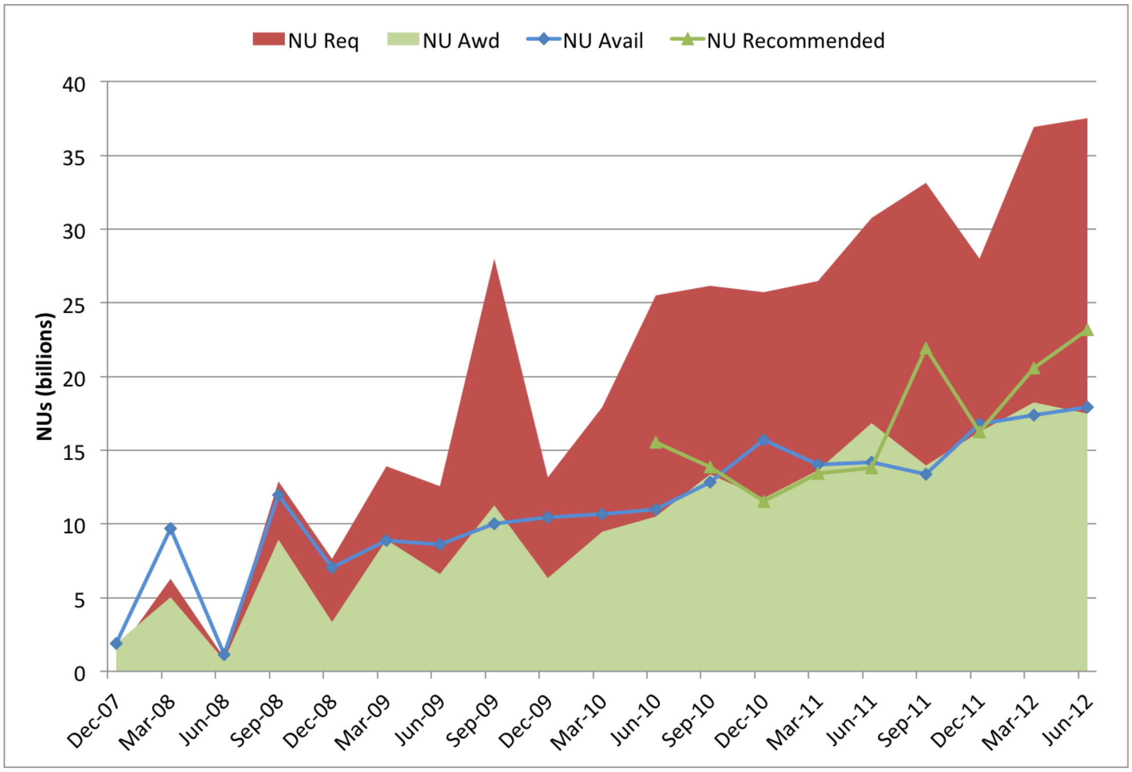


Figure 36. Allocation meeting history, showing NUs requested, awarded, available, and recommended. December 2007 and June 2008 were the last MRAC meetings, i.e. only had "medium" requests.

Table 8. Project summary metrics

Project metric	Q3 2011	Q4 2011	Q1 2012	Q2 2012
XRAC requests	157	138	123	147
XRAC request success	91%	88%	96%	90%
XRAC new awards	37%	29%	36%	39%
Startup requests	137	129	168	206
Startup request success	86%	89%	89%	82%
Projects open	1,588	1,560	1,545	1,600
Projects new	187	188	216	247
Projects active	971	917	979	1,027
Projects closed	252	254	219	207
Resource diversity (wtd)	1.5 (2.3)	1.4 (2.1)	1.4 (1.9)	1.5 (2.0)
SP diversity (wtd)	1.4 (1.8)	1.3 (1.7)	1.3 (1.6)	1.3 (1.7)

Table 9. Project activity in designated categories.

Type	Open	New	Closed	Active	NUs	% NUs
Campus Champion	96	15	10	36	62,083,592	0.4%
Science Gateway	21		1	18	809,537,474	4.8%
TG Staff Project	24		2	14	64,210,311	0.4%
Other	1,459	232	194	959	15,798,780,462	94.4%
Totals	1,600	247	207	1,027	16,734,611,839	100.0%

Table 10 Project activity by allocation board type.

Board	Open projects	Open users	Active projects	Active users	NUs
XRAC/TRAC	608	4,404	533	1,525	15,905,144,014
Startup	821	1,998	407	534	615,705,137
Discretionary	2	25	3	7	76,285,393
Staff	30	514	16	85	64,210,358
Campus Champions	80	433	29	65	43,498,826
Educational	59	1,095	39	299	29,768,115
Totals	1,600	8,469	1,027	2,515	16,734,611,843

Table 11. Allocations activity in POPS, excluding staff and discretionary projects.

	Research				Startup			
	# Req	**SUs Req**	**# Awd**	**SUs Awd**	**# Req**	**SUs Req**	**# Awd**	**SUs Awd**
New	62	162,848,177	52	63,206,036	182	26,989,270	154	18,399,646
Prog. Report	2	1,170,050	1	607,050	n/a	n/a	n/a	n/a
Renewal	83	558,945,553	80	315,617,305	24	2,160,612	15	1,324,928
Advance	37	33,070,136	26	7,419,505	n/a	n/a	n/a	n/a
Justification	0	0	2	24,780,012	0	0	0	0
Supplemental	22	12,706,499	17	6,470,013	25	7,851,845	21	1,751,845
Transfer	94	36,417,738	81	18,715,901	38	2,311,851	35	1,540,987
Extension	62	n/a	57	n/a	33	n/a	37	n/a

	Education				Campus Champions			
	# Req	**SUs Req**	**# Awd**	**SUs Awd**	**# Req**	**SUs Req**	**# Awd**	**SUs Awd**
New	12	4,170,014	11	2,770,004	15	11,066,065	14	9,312,009
Prog. Report	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Renewal	1	200,000	1	200,000	7	4,841,209	7	3,610,017
Advance	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Justification	0	0	0	0	0	0	0	0
Supplemental	2	100,000	1	20,000	7	980,000	7	753,000
Transfer	0	0	0	0	1	2,600,000	1	190,000
Extension	3	n/a	3	n/a	0	n/a	0	n/a

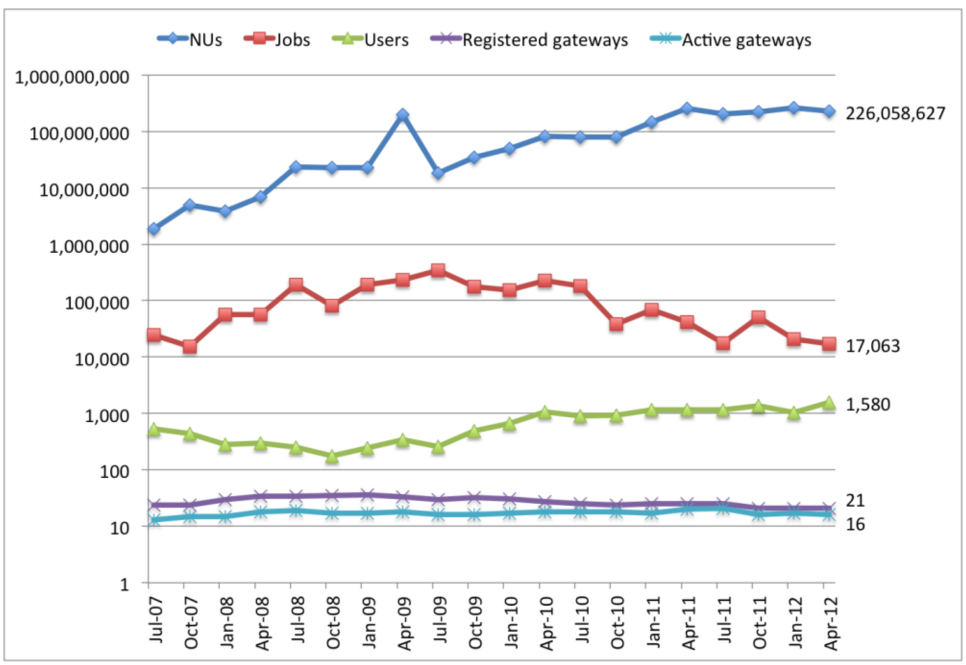


Figure 37. Quarterly gateway usage (NUs), jobs submitted, users (reported by ECSS), registered gateways, and active gateways.

Table 12. Gateway activity by resource.

Resource	Gateways	Jobs	NUs
SDSC Trestles	6	11,438	180,022,632
NICS Kraken	5	3,465	30,762,265
TACC Ranger	9	1,131	11,714,386
SDSC Gordon	3	63	1,680,365
TACC Lonestar	5	349	1,545,852
PSC Blacklight	1	44	192,310
NCSA Forge	1	250	95,310
Purdue Steele	2	323	45,506

E.2.3 Resource and usage metrics

Figure 38 shows the total NUs delivered by XSEDE computing systems, as reported to the central accounting system over the past five years. In Q2, the systems delivered 16.7 billion NUs, an increase of about 13% from the previous quarter, and 7% more NUs than the year-ago quarter. Table 13 breaks out the resource activity according to different resource types.

Figure 39 presents a perspective of the capacity and capability use of XSEDE resources by project. The figure shows the cumulative percentage of projects and resource usage according to each project's largest reported job size (in cores). The point at which the proverbial 80/20 rule holds precisely is at 76/24; that is the 76% of projects whose largest jobs were between 512 and 1,024 cores consumed only 24% of the delivered NUs, while the remaining 24% of projects consumed the remaining 76% of delivered NUs.

Finally, Table 14 presents some summary metrics to reflect aggregate "usage satisfaction," including the average run time, wait time, response time (run + wait), and slow down (or expansion factor). These values are presented as unweighted averages, which show the impact of small jobs, and as averages weighted by each job's portion of the workload (in core-hours), which show responsiveness to the jobs responsible for most of the delivered NUs. Notably for Q2, while the "average" job is only 2 hours long, the average weighted job is just over 20 hours long. On most of the unweighted usage satisfaction metrics, XSEDE showed modest improvements, suggesting better turnaround for shorter jobs. However, larger jobs did not fare as well, with increases in average weighted wait time, average weighted response time, and average weighted slow down. The weighted average for slow down (4.7) eliminates the skew in the job slow down attributed to small jobs and shows a much more realistic average perceived slowdown for the work delivered.

XSEDE provides central monitoring of GRAM5 and GRAM job submission activity at XSEDE SP sites (Figure 40). GRAM has been deprecated in favor of GRAM5.

Table 13. Resource activity, by type of resource, excluding staff projects.
Note: A user will be counted for each type of resource used.

Type	Resources	Jobs	Users	NUs
High-performance computing	6	1,244,147	1,858	15,200,359,485
Data-intensive computing	2	103,335	458	1,380,167,894
Visualization system	3	10,454	123	71,489,074
High-throughput computing	2	341,696	18	18,385,076
Total	13	1,699,632	2,457	16,670,401,529

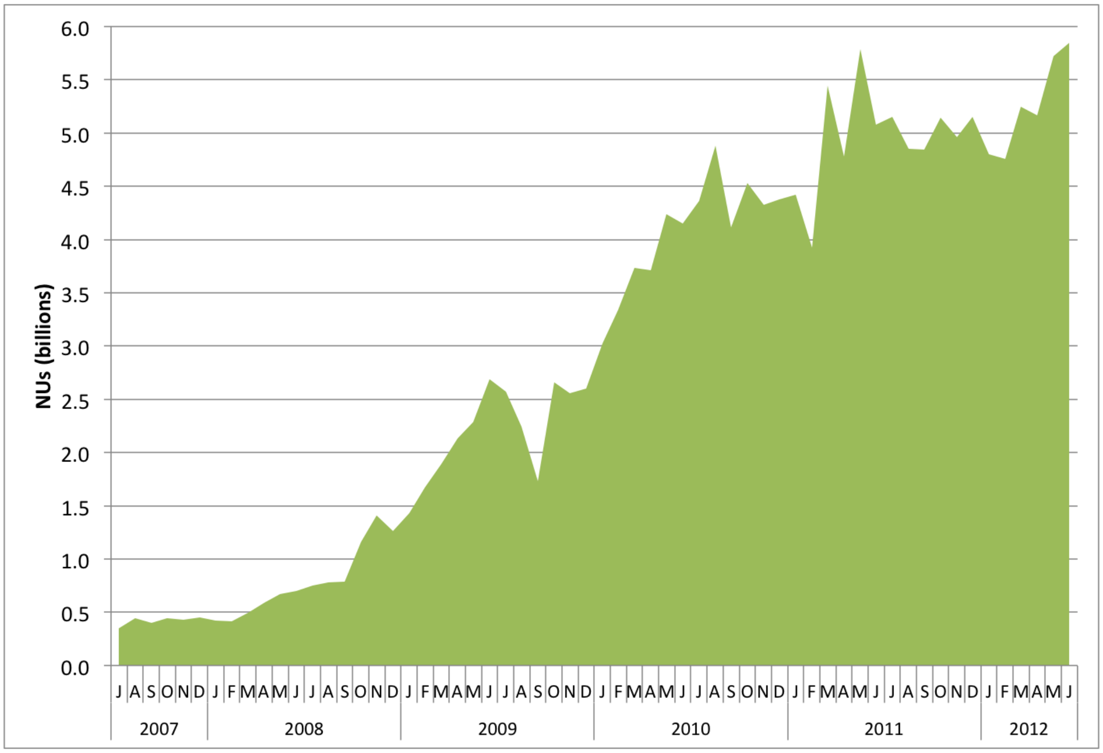


Figure 38. Total XSEDE resource usage in NUs.

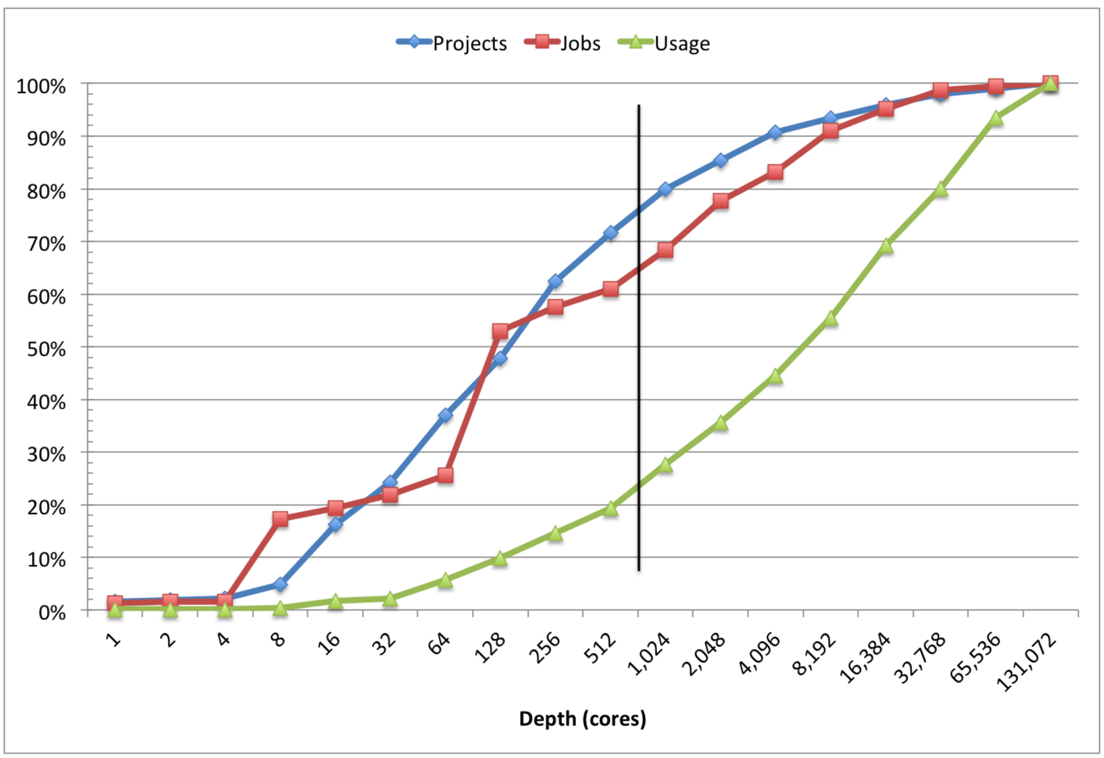


Figure 39. Cumulative distribution of projects, jobs, and usage according to project's maximum job size in cores (excludes staff projects). Vertical line (black) shows "joint ratio" of 76/24 at between 512 and 1,024 cores. I.e., 76% of projects use fewer than 1,024 cores and consume 24% of XSEDE NUs; the other 24% of projects have jobs larger than 1,024 cores and consume the other 76% of XSEDE NUs.

Table 14. Usage satisfaction metrics, for HPC and data-intensive computing resources only, excluding staff projects.

	Job attribute	Q3 2011	Q4 2011	Q1 2012	Q2 2012
Unweighted average	Run time (hrs)	2.0	3.0	5.0	2.0
	Wait time (hrs)	4.3	3.1	6.1	4.4
	Response time (hrs)	7.2	7.0	12.1	7.3
	Slow down	699.4	252.3	334.4	324.7
Weighted average	Wtd run time (hrs)	22.6	22.5	23.3	20.9
	Wtd wait time (hrs)	27.7	26.8	23.2	32.6
	Wtd response time (hrs)	50.3	49.4	46.5	53.5
	Wtd slow down	3.1	3.5	3.3	4.7

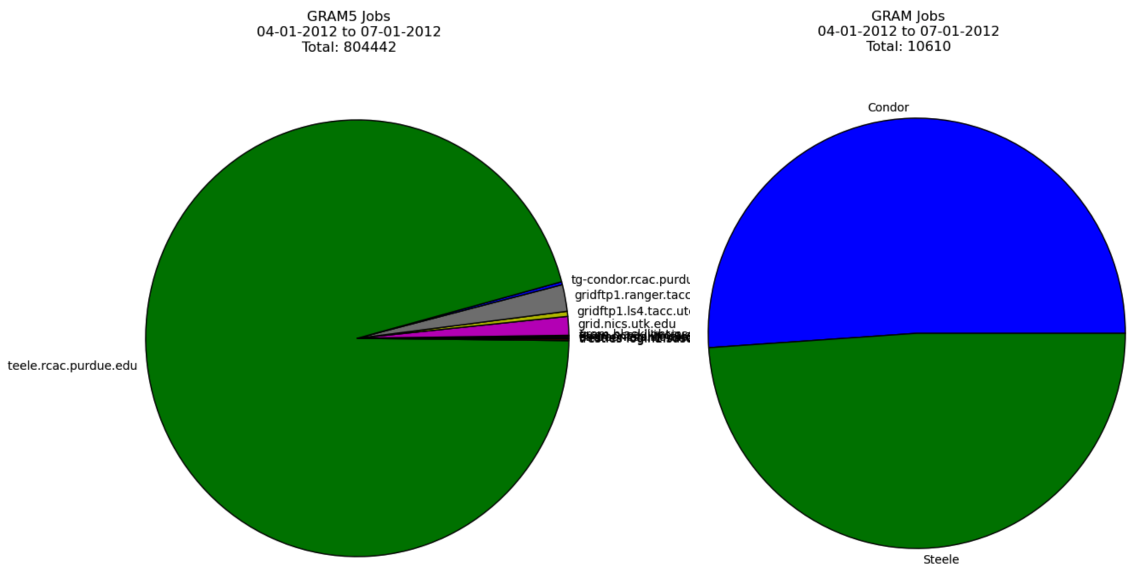


Figure 40. GRAM5 and GRAM jobs recorded per resource.

E.3 XSEDE Program Metrics

E.3.1 Project Office 1.1

1.1.4 External Relations

The XSEDE External Relations team reported the following 33 media hits for the quarter.

Table 15. XSEDE media hits.

Date	Source	Item
6/20/12	Medicalxpress.com	Husband-wife team set out to improve breast cancer exams
6/18/12	StorageReview.com	TACC Stampede supercomputer will deploy Mellanox InfiniBand interconnects
6/18/12	HPCwire	Stampede supercomputer to run 56G IB interconnect
6/18/12	PR Newswire	Appro deployed an aggregate performance of nearly five petaflops of HPC in last ten months
6/18/12	redOrbit	Appro deployed an aggregate performance of nearly five petaflops of HPC in last ten months
6/18/12	MarketWatch	Mellanox FDR 56Gb/s InfiniBand to Accelerate World-Class Supercomputer at TACC
6/18/12	HPCwire	Appro ships over 5 Pflops of E5 systems
6/14/12	HPCwire	Six days remain for XSEDE12 Early Bird Registration
6/11/12	Datanami	Understanding data-intensive analysis on large-scale HPC compute systems

Date	Source	Item
6/5/12	HPCwire	San Diego Supercomputer Center completes Data Oasis
6/1/12	genomeweb/Bioinform	Pittsburgh Supercomputing Center uses $1.5M NSF grant to link Galaxy to XSEDE network
5/29/12	NSF website	Pittsburgh Supercomputing Center provides direct link from Galaxy to the XSEDE backbone
5/29/12	HPCWire	Optical exchange connects XSEDE, Penn State networks
5/24/12	ScienceBlog	How ion bombardment reshapes metal surfaces
5/17/12	Appro	SDSC director talks shop about Appro supercomputing; Gordon Supercomputer surpasses every initial benchmark set forth by the SDSC
5/14/12	HPCWire	Registration now open for XSEDE12
5/10/12	HPCWire	National Computational Science Institute, XSEDE offer workshops to educators
5/7/12	HPCWire	XSEDE12 introduces new sponsor levels aimed at nonprofits, education groups
5/2/12	HPCWire	Researcher at Ohio Supercomputer Center named Campus Champion Also: http://phys.org/news/2012-05-osc-tomko-champion-high-performance.html
4/30/12	HPCWire	Appro to sponsor XSEDE12
4/25/12	ISGTW	Containing multitudes
4/23/12	HPCwire	Harvard team to study green technology on Trestles supercompter
4/23/12	Phys.org	SDSC's Trestles supercomputer speeds clean energy research
4/19/12	HPCwire	XSEDE12 extends deadline for papers
4/18/12	HPCWire	New analytics facility to launch at San Diego Supercomputer Center
4/17/12	HPCWire	University of Illinois Urbana-Champaign awarded time on Blacklight supercomputer
4/12/12	HPC Advisory Council	Distinguished mathematician, professor, diversity advocate Richard Tapia to deliver keynote at XSEDE12
4/11/12	HPCwire	Cornell Center for Advanced Computing to present workshop on TACC supercomputers
4/10/12	HPC in the Cloud	Live from GlobusWorld 2012
4/2/12	Appro Supercomputer Solutions website	Appro Xtreme-X, SDSC's Gordon Supercomputer Ready for Researchers

E.3.2 Operations 1.2

1.2.1 Security

The XSEDE security team has identified the following metrics for tracking security incidents and response. Table 12 summarizes the metrics, and details on any incidents are provided in the main body of the report. The security enhancement rollout for Q2 was a new XSEDE CA tarball, announced April 12, 2012, that included the NCSA 2-factor SLCS CA, the DigiCert CA, and CILogon CA.

1.2.2 Data Services

XSEDE supports monitoring for two central data movement services: the gridFTP service connecting the XSEDE service providers and the Globus Online service for connecting XSEDE service providers as well as external sites. Table 17 shows quarterly summary metrics for the past several quarters, while Figure 41 and Figure 42 show Globus Online and GridFTP activity, respectively, by SP site.

Table 16. XSEDE security metrics and incident response

	Q4 2011	Q1 2012	Q2 2012
XSEDE-wide notice of vulnerability	0	0	0
Compromised user accounts	5	0	0
Other incident response	1	0	0
Critical rollout of vulnerability patches	0	0	0
Security enhancement rollouts	0	0	1

Table 17. Globus Online activity to and from XSEDE endpoints, excluding GO XSEDE speed page user.

		Q3 2011	Q4 2011	Q1 2012	Q2 2012
To/from XSEDE endpoint	Files to XSEDE	12.7 million	8 million	27.6 million	34.8 million
	TB to XSEDE	206	216	232	824
	Files from XSEDE	13.5 million	21.8 million	17.8 million	23.9 million
	TB from XSEDE	186	180	242	185
	Faults detected	732,000	573,000	584,000	971,000
	Users	116	124	149	187
To/from XSEDE via Globus Connect	Files to XSEDE	2.2 million	1.5 million	6.3 million	2.5 million
	TB to XSEDE	41	19	45	18
	Files from XSEDE	1.7 million	1.4 million	0.6 million	2.2 million
	TB from XSEDE	18	20	25	11
	Faults detected	422,000	102,000	242,000	381,000
	Users	71	74	87	97

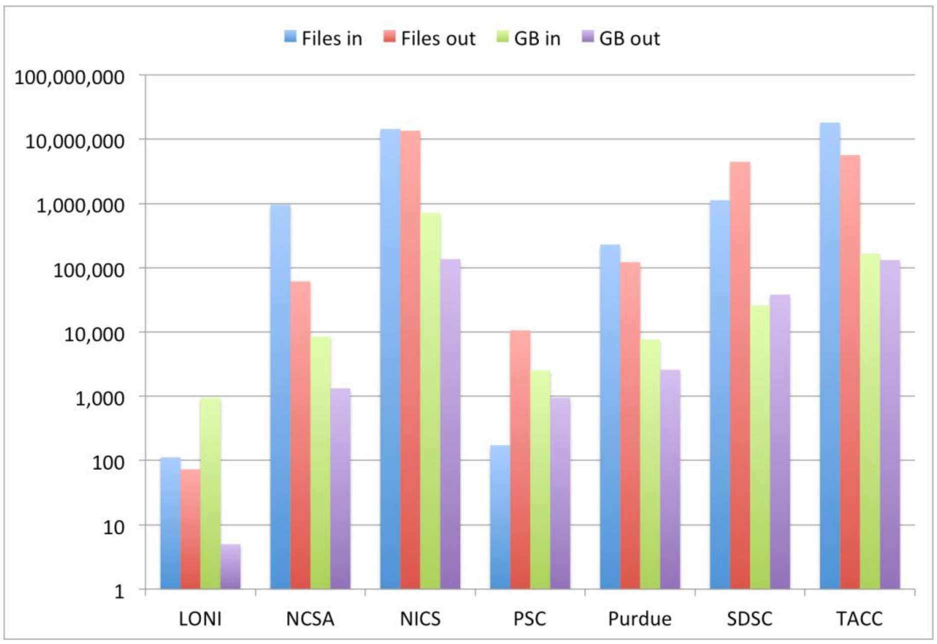


Figure 41. Globus Online activity into and out of XSEDE SP end points

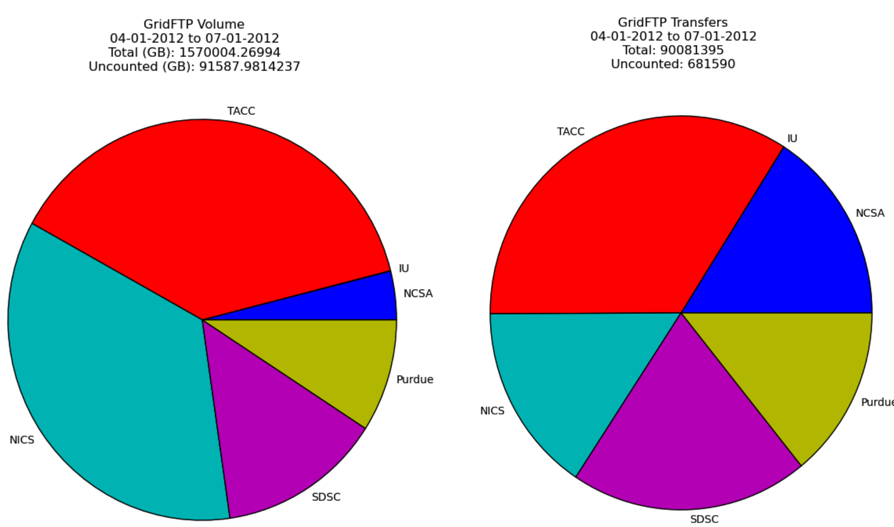


Figure 42. GridFTP volume and file transfers, per SP site.

1.2.3 XSEDEnet

Traffic utilization of the Chicago-Denver XSEDEnet link is shown in two figures below. Figure 43 shows the peak bandwidth across the link for the period. Figure 44 shows link utilization as a percentage. Traffic across all XSEDEnet links is shown in Figure 45. In Q2 2012, NLR changed the provider for their network operations center, which resulted in a few weeks of missing data, as well as changes to the style and format for these charts.

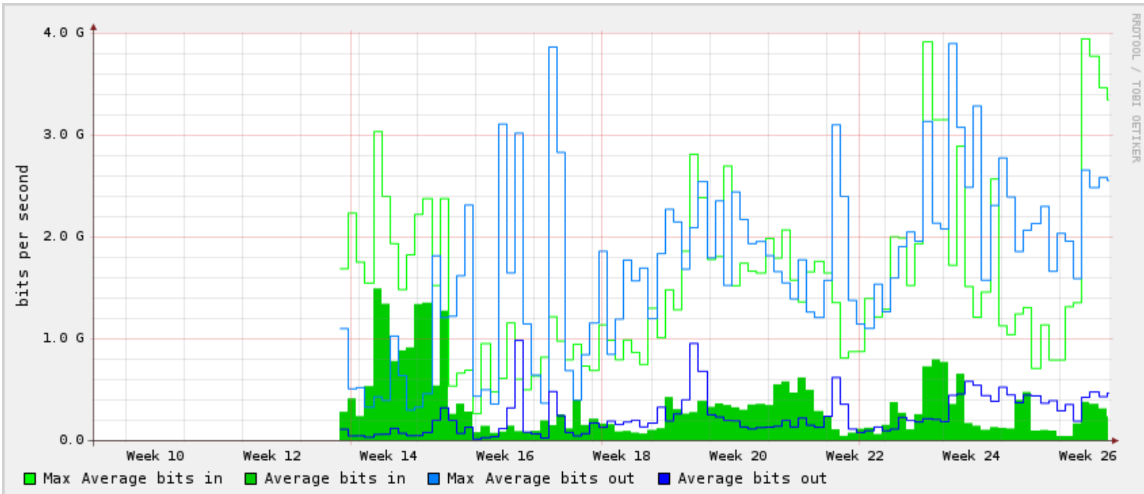


Figure 43. XSEDEnet Chicago-Denver peak bandwidth

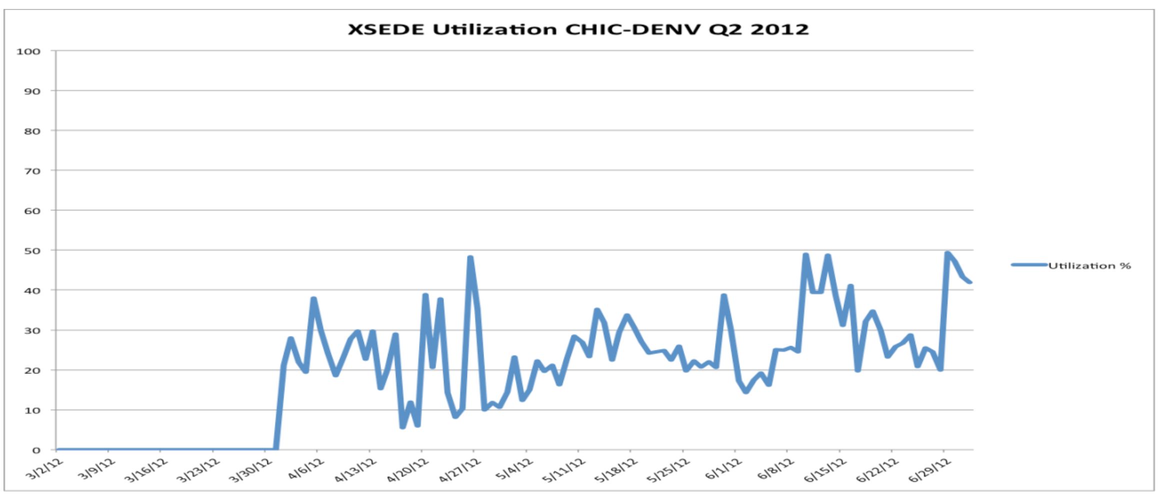


Figure 44. XSEDEnet Chicago-Denver utilization (as a percentage)

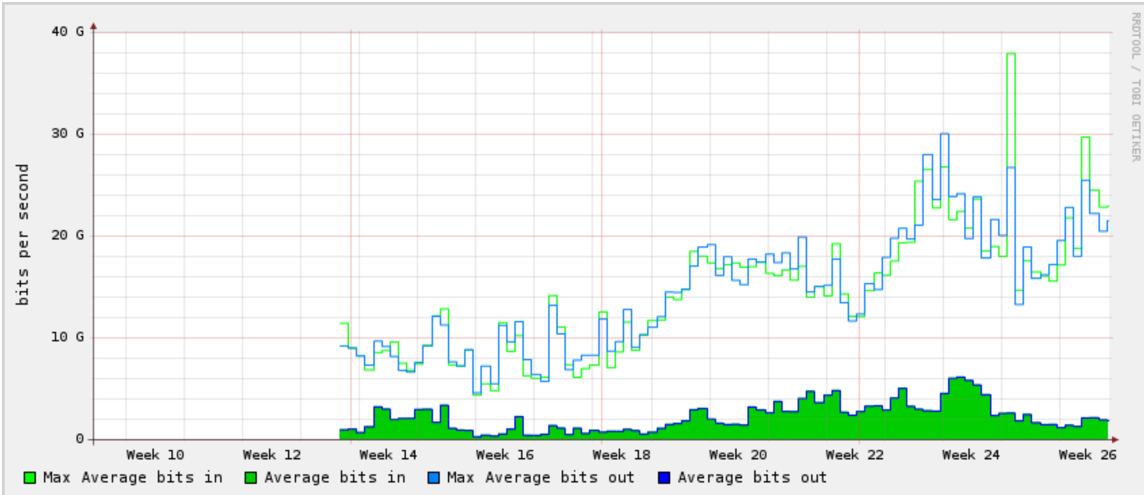


Figure 45. XSEDEnet Q2 2012 aggregate bandwidth across all links

1.2.5 Accounting and Account Management

The Accounting and Account Management group administers and operates the software for the XSEDE allocations system (POPS), the accounting system, and user account management. XSEDE reduced the account creation time by more than 50% by training a Help Desk team member to provide additional support in this area.

Table 18. Average time to process allocation requests and account creation requests, in days. (Excludes quarterly XRAC requests; "n/a" indicates none submitted.)

ALLOCATION REQUESTS		**Q3 2011**	**Q4 2011**	**Q1 2012**	**Q2 2012**
Research	Advance	15	12	6	7
	Transfer	5	4	3	3
	Supplement	23	27	25	14
	Justification	37	54	15	n/a
Startup, Education, Campus Champions, Discretionary	New	13	13	9	9
	Renewal	11	11	10	6
	Transfer	5	4	3	3
	Supplement	10	11	4	4
Account creation requests		3.2	4.6	1.8	0.03

1.2.6 Systems Operational Support

The Systems Operational Support group encompasses the XSEDE Operations Center (XOC), which includes front-line user support and the ticket system, and the system administration of all XSEDE centralized services. In the ticket system, XSEDE tracks total ticket volume, responsiveness, which groups ("resolution centers") field the tickets, and the numbers of tickets in seven common categories. The totals by resolution center and by category do not add up to the total number of tickets opened and closed because some tickets are resolved by staff not in a resolution center and some resolution centers have non-standard categories.

For the central services, XSEDE tracks the uptime reported by system administrators (Table 22) as well as the Inca-reported uptime for seven key user-visible services (Table 23). The Inca-reported uptime better reflects "user-visible outages," that is, what the average user would experience, and typically exceeds the actual system uptime, reflecting the effectiveness of XSEDE's backup systems, failover capabilities, and operational responsiveness.

Table 19. XSEDE Operations Center ticket system metrics.

	Q3 2011	Q4 2011	Q1 2012	Q2 2012
Total tickets opened	2,800	2,140	2,651	2,744
Tickets opened – email	1,788	1,318	2,381	2,448
Tickets opened – portal	812	684	254	26
Tickets opened – phone	200	138	16	270
Total tickets closed	2,965	2,111	2,335	2,394
Tickets, response in 24 hrs	1,985 (71%)	1,550 (72%)	2,263 (85%)	2,326 (85%)
Tickets closed within 2 bus. days	925 (33%)	739 (35%)	1,044 (39%)	1,050 (38%)

Table 20. Ticket breakdown (opened/closed) for each major resolution center.

	Q3 2011		Q4 2011		Q1 2012		Q2 2012	
	Opened	**Closed**	**Opened**	**Closed**	**Opened**	**Closed**	**Open**	**Closed**
NICS	529	506	413	372	670	605	657	640
XOC	518	396	398	278	427	305	415	298
TACC	408	352	336	300	330	295	276	251
Proposal issues	163	136	142	97	343	251	408	376
SDSC	137	120	118	100	218	182	379	327
PSC	143	121	160	147	115	108	89	87
Purdue	129	116	68	64	69	58	118	115
NCSA	134	123	45	36	60	42	89	72
User facing services	93	83	64	51	88	68	73	65
UST	16	12	16	14	12	9	11	11
IU	7	3	7	0	7	1	10	7
OSG	0	0	1	0	2	1	3	3
Others	*667*	*532*	*535*	*430*	*335*	*258*	*215*	*135*

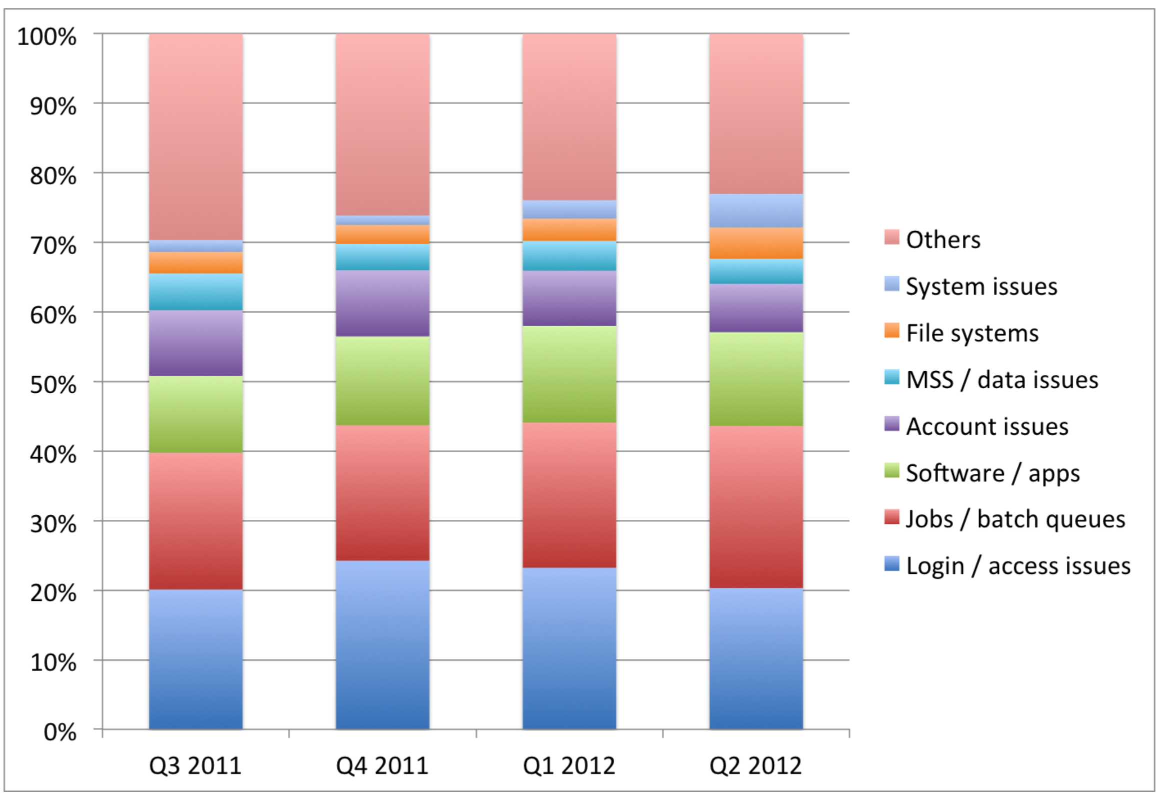


Figure 46. Tickets in seven primary problem categories. This pie chart represents a significant portion of tickets but does not represent the entire range. Tickets largely fit in the seven displayed categories; other categories are not significant enough to visually represent.

Table 21. Ticket counts for the seven primary problem categories shown in Figure 46.

	Q3 2011	Q4 2011	Q1 2012	Q2 2012
Login / access issues	456	416	483	445
Jobs / batch queues	446	334	434	510
Software / apps	250	219	289	295
Account issues	214	163	165	152
MSS / data issues	120	65	89	79
File systems	70	47	67	99
System issues	39	23	55	105
Others	*672*	*448*	*497*	*504*

Table 22. XSEDE centralized service uptime and outages. Empty cells indicate *no outages (100% up)*. “% Up” is percent uptime; “Hrs (P|U)” shows outage hours, planned and unplanned.

Service	Q3 2011		Q4 2011		Q1 2012		Q2 2012	
	% Up	Hrs (P|U)	% Up	Hrs (P|U)	% Up	Hrs (P|U)	% Up	Hrs (P|U)
AMIE	98.87%	0|25					99.91%	0|2
AMIE backup	99.62%	0.5|8	99.77%	5|0	99.98%	0.5|0		
Bugzilla					97.46%	0|56	99.82%	4|0
Build and Test	99.77%	0|5			97.46%	0|56	99.82%	4|0
Certificate Authority					99.62%	8.5|0		
Data Movement Service	99.77%	0|5						
Globus Listener	99.46%	12|0			97.46%	0|56	99.82%	4|0
IIS Metrics	99.77%	0|5			97.46%	0|56	99.82%	4|0
Inca	96.20%	0|84	98.44%	3.75|30.75			99.31%	0|15
Inca backup							99.50%	11|0
Information Services	99.77%	0|5	99.95%	0|1	97.46%	0|56	99.82%	4|0
Karnak	98.28%	0|38	98.60%	0|31	97.40%	0|57.5	99.82%	4|0
Kerberos backup	99.95%	0|1	99.91%	2|0	99.91%	2|0.08		
Knowledgebase	99.77%	0|5						
MyProxy	99.91%	0|2	99.73%	6|0				
Openfire Jabber	99.46%	12|0						
POPS	98.91%	24|0						
RDR	99.62%	0.5|8	99.77%	5|0	98.69%	5|24		
Secure Wiki	99.46%	12|0						
SELS	99.46%	12|0						
Sharepoint	98.87%	0|25						
Software Distribution	99.77%	0|5			97.46%	0|56	99.82%	4|0
Source Repository	99.77%	0|5			97.46%	0|56	99.82%	4|0
Speedpage	99.95%	0|1	99.91%	2|0	99.91%	2|0		
TG Wiki	99.77%	0|5			97.46%	0|56	99.82%	4|0
Ticket System	98.91%	24|0						
Usage Reporting Tools	99.46%	12|0						
User News	98.87%	0|25	*n/a. User news became part of User Portal in Q4 2011.*					
User Portal			99.99%	0|0.17	99.97%	0|0.75		
User Portal backup					99.62%	8.5|0	99.95%	1|0
User Profile Service	99.77%	0|5			97.46%	0|56		
XDCDB	99.57%	0|9.5	99.99%	0.25|0			99.91%	0|2
XDCDB backup			99.77%	5|0	99.98%	0.5|0	99.91%	2|0

Table 23. Inca-monitored XSEDE central services, Inca-detected uptimes, and outages.

Service	Outage definition. Test frequency.	Q3 2011	Q4 2011	Q1 2012	Q2 2012
Inca	Inca status pages unavailable or test details page fails to load. Every 5 min.	97.1%	98.4%	100%	99.3%
Information Services	Information web pages unavailable. Every 15 min.	100%	99.95%	99.97%	99.9%
Karnak	Karnak front page fails to load. Every 30 min.	99.2%	98.6%	99.93%	99.9%
MyProxy	MyProxy server does not respond to credential check. Every hour.	99.9%	99.7%	100%	100%

Service	Outage definition. Test frequency.	Q3 2011	Q4 2011	Q1 2012	Q2 2012
User Portal	Portal home page fails to load correctly. Every 30 min.	100%	100%	99.98%	100%
XDCDB	Connection to database refused or slow (using check_postgres.pl script). Every 5 min.	99.9%	100%	100%	99.99%

E.3.3 User Services 1.3

The User Information and Interfaces group provides XSEDE users with central information and services via the XSEDE User Portal (XUP), web site, XUP mobile, and knowledgebase. Table 24 shows increasing activity on most user information interfaces, as well as increases in the numbers of logged-in users accessing these interfaces. Table 25 shows the most popular XUP applications, by visits in Q1 2012. After being integrated into XUP last quarter, User News received the most hits this quarter, but Allocations/Usage had the most users.

Table 24. XSEDE web site, user portal and XUP Mobile activity. (Note: "Users" indicates logged-in users.)

UII Activity	Q3 2011	Q4 2011	Q1 2012	Q2 2012
Web hits	1,114,358	1,924,648	2,667,120	2,676,532
Web visitors	11,693	24,208	38,502	33,362
XUP hits	751,526	1,164,994	1,340,704	1,465,537
XUP visitors	6,290	10,346	14,913	13,787
XUP users	2,694	3,024	3,919	3,976
XUP Mobile hits	886	1,702	1,307	1,693
XUP Mobile users	41	48	37	25
KB docs retrieved	17,139	64,679	68,619	57,451
Total KB docs	348	405	478	497
New KB docs	24	25	73	19

Table 25. XUP and Web site application visits. "Users" indicates logged-in users.

	Q3 2011		Q4 2011		Q1 2012		Q2 2012	
Application	**Visits**	**Users**	**Visits**	**Users**	**Visits**	**Users**	**Visits**	**Users**
Allocations/Usage	27,589	2,686	69,290	2,991	49,261	3,803	50,998	3,724
User News			4,312	25	62,784	526	47,194	2,013
GSI-SSH	11,438	919	34,132	1,083	24,714	1,343	28,317	1,354
Resource Listing	64,688	652	100,544	864	36,244	1,104	22,483	1,265
File Manager	49,537 (transfers)	81 users (7.3 TB)	60,009 (transfers)	60 (3.9 TB)	31,952 (transfers)	79 (2.5 TB)	21,375 (transfers)	78 (20 TB)
System Accounts	10,949	1,515	13,457	1,609	12,897	1,841	12,235	1,813
Training Registration			12,648	345	19,535	828	11,642	631
Knowledge Base	4,882	285	17,454	261	11,436	377	10,964	338
System Monitor	5,861	1,047	13,350	1,050	8,213	1,257	9,401	1,034
Science Gateways List	10,803	177	51,684	180	16,338	229	8,914	234
Software Search	5,249	150	14,812	176	8,204	265	8,879	316
POPS Submit							5,639	1,204
User Profile	870	398			1,954	833	5,476	1,262
Help Desk (Consulting)	2,520	434	6,206	557	4,659	748	4,815	733
Add User Form	3,483	533	8,406	611	5,751	626	4,628	617
My Jobs	2,465	834	5,640	908	3,500	960	3,195	900
Ticketing System	1,503	559	3,708	609	2,366	682	2,600	654
Online Training Listing	1,111	175	2,720	215	2,093	347	1,668	437
SU Calculator (web)	1,001	141	4,376	180	2,048	243	1,466	239
Karnak Q Prediction	751	215	1,770	212	1,108	222	871	264
DN Listing	498	371	1,092	399	554	371	450	326
Community Accounts							436	311
Feedback form			1,174	57	791	53	426	39
Gateway Registration							243	16

E.3.4 Extended Collaborative Support Service 1.4, 1.5

The Extended Collaborative Support Service pairs members of the XSEDE user community with expert staff in projects lasting up to a year to solve challenging science and engineering problems. Table 26 shows project and staffing metrics and, notably, the increase to 100 active projects in for Q2 2012. Table 27 shows metrics for Extended Support for Training, Education, and Outreach.

Table 26. Extended Collaborative Support project and staffing activity

		Q3 2011	Q4 2011	Q1 2012	Q2 2012
Project requests	XRAC	22	16	22	20
	Supplemental/Startups	5	21	41	28
	ECSS In-house project			1	0
Projects initiated	Research Team	6	18	19	14
	Community Codes	2	9	8	7
	Science Gateways		4	6	6
	Unassigned			7	0
Projects cancelled/no-go	XRAC		1	4	5
	Supplemental/Startups		9	20	16
Projects active	Research Team (XRAC)		7	27	29
	Research Team (S/S)		4	17	20
	Research (TG)		29	12	1
	Subtotal	**37**	**40**	**56**	**50**
	Community Codes (XRAC)		3	6	10
	Community Codes (S/S)		3	9	17
	Community Codes (TG)		2	2	2
	In-house			1	1
	Subtotal	**2**	**8**	**18**	**30**
	Science Gateways (XRAC)			4	9
	Science Gateways (S/S)			7	10
	Science Gateways (TG)	2	10	4	1
	Subtotal	**2**	**10**	**15**	**20**
	Total Projects Active	**41**	**58**	**89**	**100**
Work plans			6	14	10
Projects completed		7	11	16	15
Novel, Innovative Projects (NIP)	User groups engaged	27	15	20	47
	NIP-led ECS planning efforts	8	8	9	9
	NIP ECS requests (prospective user groups)	5	14	23	38
	NIP ECS projects active	5	5	7	9
	NIP outreach events	2	9	9	16
ECSS staffing (FTE)	Research Team	13.62	13.72	13.28	12.8
	Community Codes	6.98	6.73	7.11	8.24
	Science Gateways	4.83	4.83	4.92	4.73
	NIP	4.7	5.30	5.79	5.64
	TEO	4.07	4.02	4.42	3.53
	Total	**34.20**	**34.60**	**35.52**	**34.94**

Table 27. Extended Support for Training, Education and Outreach 1.5.3

	Q3 2011		Q4 2011		Q1 2012		Q2 2012	
Description	**#**	**Staff**	**#**	**Staff**	**#**	**Staff**	**#**	**Staff**
Requests for service	1*		1	1	4	4	2	2
User meetings and BOFs	5	15	7	9	2	2	18	20
Mentoring	2	4	3	3	3	2	12	12
Talks and presentations	6	8	7	8	12	9	20	20
Tutorials	33	58	3	3	21	25	31	32
Online tutorials and webinars	2	5	1	3	4	6	0	0
Online tutorial reviews	8	10	5	4	6	6	17	17

F XSEDE Publications Listing

F.1 XSEDE Staff Publications

XSEDE staff identified the following publications and presentations and being supported in full or in part by the XSEDE award. In prior quarters of the program year, another 40 publications were also supported by XSEDE: 17 (Q3 2011), 13 (Q4 2011), and 10 (Q1 2012). The definition and collection process for XSEDE staff publications is being refined for Q3 2012.

F.1.1 Project Office 1.1

1.1.2 System & Software Engineering, 1.1.3 Architecture & Design, 1.1.6 Software development & Integration

1. Allen, B., Bresnahan, J., Childers, L., Foster, I., Kandaswamy, G., Kettimuthu, R., Kordas, J., Link, M., Martin, S., Pickett, K. and Tuecke, S., “Software as a Service for Data Scientists”, Communications of the ACM, 55(2):81-88, 2012.
2. Kettimuthu, R., Lacinski, L., Link, M., Pickett, K., Tuecke, S. and Foster, I., “Instant GridFTP”, 9th Workshop on High Performance Grid and Cloud Computing, 2012.
3. P. Mhashilkar, Z. Miller, R. Kettimuthu, G. Garzoglio, B. Holzman, X. Duan, C. Weiss, and L. Lacinski, “End-To-End Solution for Integrated Workload and Data Management using glideinWMS and Globus Online”, International Conference on Computing in High Energy and Nuclear Physics (CHEP), 2012.
4. Vas Vasiliadis, Panel on SaaS-Based Research: The Path to Reality for “Research in the Cloud”, SC’11, Seattle, WA, November 17, 2011
5. *Steve Tuecke, Vas Vasiliadis, GlobusWORLD conference, Argonne, IL, April 10-12 2012.*
6. Ravi Madduri, Panel on SaaS-Based Translational Research: The Path to Reality for “Research in the Cloud”, American Medical Informatics Association (AMIA) Conference, San Francisco, March 19, 2012
7. R. Ananthakrishnan, “Globus Online and Identity Management”, MAGIC Meeting, August 3, 2011.
8. R. Ananthakrishnan, “Globus Online: Towards a hosted solution for Research Data Management”, SE Wisconsin HPC Data Symposium, March 1, 2012.
9. Andrew Grimshaw, “The XSEDE Architecture, TeraGrid'11”, Salt Lake City, July 2011.
10. Andrew Grimshaw, “The Role of Standards for Risk Reduction and Inter-operation in XSEDE”, European Grid Initiative Annual Meeting Keynote, Lyon, FR, September, 2011.
11. Andrew Grimshaw, “The Role of Standards for Risk Reduction and Inter-operation in XSEDE”, Large Scale Network (LSN) Middleware And Grid Infrastructure Coordination (MAGIC) Team, February, 2012.
12. Andrew Grimshaw, “XSEDE Execution Management Services and UNICORE 6”, UNICORE 6 Summit, Munich, Germany, March, 2012.
13. Andrew Grimshaw, “Activity Endpoint Profile”, OGF 33, Lyon, FR, September, 2011.
14. Andrew Grimshaw, “BES Directory Profile”, OGF 33, Lyon, FR, September, 2011.
15. Andrew Grimshaw, “Activity Endpoint Profile”, OGF 34, Oxford, UK, March, 2012.
16. Andrew Grimshaw, “XSEDE Interoperability Architecture”, OGF 34, Oxford, UK, March, 2012.
17. Andrew Grimshaw, “BES Directory Profile”, OGF 34, Oxford, UK, March, 2012.
18. Andrew Grimshaw, “Activity Endpoint Profile”, OGF 35, Delft, Netherlands, June 2012.
19. Andrew Grimshaw, “BES Directory Profile”, OGF 35, Delft, Netherlands, June 2012.
20. Lindsey, S.; M. Dahan, J.L. Fischer, C.A. Stewart, J. Boisseau. 2012. XSEDE system description template. Available from: http://hdl.handle.net/2022/14521
21. Warren Smith, “An Information Architecture Based on Publish/Subscribe Messaging”, Proceedings of the TeraGrid'11 Conference, July 2011, https://www.xsede.org/wwwteragrid/archive/c/document_library/get_file%3Fuuid=daedc319-7f5e-400e-9333-c376129cfb42&groupId=334534.
22. Stewart, C.A., Knepper, R., Grimshaw, A., Foster, I., Bachmann, F., Lifka, D., Riedel, M. and Tuecke, S. Campus Bridging Use Case Quality Attribute Scenarios, 2012. [Accessed 13 June 2012]; Available from: http://hdl.handle.net/2022/14476.
23. Stewart, C.A., Knepper, R., Grimshaw, A., Foster, I., Bachmann, F., Lifka, D., Riedel, M. and Tuecke, S. XSEDE Campus Bridging Use Cases, 2012. [Accessed 13 June 2012]; Available from: http://hdl.handle.net/2022/14475.
24. Stewart, C.A. 2012. *Campus Bridging*. (Presentation) XSEDE (eXtreme Enviroment for Science and Engineering Discovery) Advisory Board Meeting (Chicago, IL, 23 Apr). Available from: http://hdl.handle.net/2022/14443

25. Stewart, C.A., S. Marru, R. Knepper, D.Y. Hancock, J. Wernert, C. Aikman, J. Bolte, P. Brown and T.M. Miller. 2012. *Indiana University collected XSEDE update*. (Presentation) XSEDE (eXtreme Enviroment for Science and Engineering Discovery) Quarterly Meeting (Austin, TX, 6-7 Mar). Available from: http://hdl.handle.net/2022/14444
26. Stewart, C.A. 2012. “What is campus bridging and why should XSEDE Campus Champions Care?” Presentation. Presented to XSEDE Campus Champions 8 June 2012. Presented to campus champions virtual meeting from Michigan City, IN.
27. Stewart, C.A. 2012. “WBS 1.6.5 Campus Bridging Program Plan” Presentation. Presented to NSF review panel 14 June 2012. Arlington VA.
28. S. Tuecke, “Using Globus Online (XSEDE XUAS) for Reliable, Secure File Transfer”, TeraGrid’11 presentation, July 19, 2011.
29. S. Tuecke, “Software-as-a-Service for Research Data Management”, University of Utah seminar, July 20, 2011.
30. S. Tuecke, “Globus Toolkit Update”, GlobusEurope Conference presentation, September 19, 2011.
31. S. Tuecke, “Globus Online Update”, GlobusEurope Conference presentation, September 19, 2011.
32. S. Tuecke, “Simplifying Large-Scale Data Movement with Globus”, EGI Technical Forum Data Workshop presentation, September 19, 2011.
33. S. Tuecke, “Globus Online: Reliable File Transfer. No IT Required”, EGI Technical Forum presentation, September 20, 2011.
34. S. Tuecke, “Globus related standards: Past, present and future”, MAGIC Forum webcast, February 1, 2012.
35. S. Tuecke, “The Research Cloud”, CASC quarterly meeting presentation, February 29, 2012.
36. S. Tuecke, “Delegation and Authorization Challenges with SaaS”, NSF Workshop on Security for Cloud Computing presentation and panel, March 16, 2012.
37. S. Tuecke, “Accelerating data-driven discovery by outsourcing the mundane”, European Globus Community Forum keynote, March 26, 2012.
38. S. Tuecke, “Delivering a Scalable Service”, European Grid Initiative community forum presentation, March 27, 2012.
39. I. Foster, “Globus State of the Union”, GlobusWORLD, April 11, 2012.
40. S. Tuecke, “Integrating with Globus Online”, GlobusWORLD, April 11, 2012.
41. S. Tuecke, “Globus Online Product Roadmap”, GlobusWORLD, April 12, 2012.
42. V. Vasiliadis, “Hybrid Strategies for Biomedical Research Data Management”, Bio IT World, April 25, 2012.
43. S. Tuecke, “Cloud Based Services for Science”, Condor Week conference presentation, May 2, 2012.
44. I. Foster, "So long, computer overlords: How Cloud (and Grid) can liberate research IT – and transform discovery", National Science Foundation, Washington DC, August 10, 2011.
45. I. Foster, "Accelerating data-driven discovery by outsourcing the mundane," Keynote, ISC Cloud Conference, Mannheim, Germany, September 26, 2011.
46. I. Foster, "Accelerating data-driven discovery by outsourcing the mundane," RPI-NSF Workshop on Multiscale Modeling of Complex Data, Rensselaer Polytechnic Institute, Troy, NY, September 12, 2011.
47. I. Foster, "Powering Collaboration with Cloud-Hosted Data Movement", Internet2 Fall Members meeting, Raleigh, NC, October 5, 2011.
48. I. Foster, "High-performance, collaborative computing for the long tail of GIS", Keynote, ACM HPDGIS 2011: the Second International ACM SIGSPATIAL Workshop on High Performance and Distributed Geographic Information Systems (HPDGIS), Chicago, IL, November 1, 2011.
49. I. Foster, "Accelerating data-driven discovery by outsourcing the mundane," Keynote, Workshop on Network-aware Data Management, Seattle, WA, November 14, 2011.
50. I. Foster, “My Life in a Time Machine,” Panel on SaaS-Based Research: The Path to Reality for “Research in the Cloud”, SC’12, Seattle WA, November 17, 2011.
51. I. Foster, "Rethinking cyberinfrastructure for massive data but modest budgets", International Workshop on Grid Computing: The Next Decade, Zakopane, Poland, January 4, 2012.
52. I. Foster, "Rethinking cyberinfrastructure for massive data but modest budgets", Keynote, Data Symposium 2012, Milwaukee, WI, March 1, 2012.
53. I. Foster, "Accelerating data-driven discovery by outsourcing the mundane," Keynote, ISUM 2012 Congreso Internacional de Supercómputo, Guanajuato, México, March 13, 2012.
54. I. Foster, "How software as a service can transform research," presentation to cross-campus faculty summit, University of Chicago, April 10, 2012.
55. I. Foster, "Big process for big data: Process automation for data-driven science", Keynote, NIST Big Data Workshop, Gaithersburg, Maryland, June 13, 2012.
56. I. Foster, "Process automation for data-driven science," Keynote, Materials Genome Initiative Workshop, NIST, Washington DC, May 14-15, 2012.

57. I. Foster, "My life in a time machine," Panel on SaaS-Based Translational Research: The Path to Reality for "Research in the Cloud", American Medical Informatics Association (AMIA) Conference, San Francisco, March 19, 2012.
58. I. Foster, "20 years of grid computing," Keynote, IEEE International Symposium on High-Performance Distributed Computing, Delft, The Netherlands, June 22, 2012.
59. I. Foster, "Process automation for data-driven science," HPC Conference, Cetraro, Italy, June 25, 2012.

F.1.2 Operations 1.2

F.1.3 User Services 1.3

F.1.4 Extended Collaborative Support Service –1.4, 1.5

60. Potluri S., Wang H., Bureddy D., Singh A., Rosales C., Panda D "Optimizing MPI Communications on Multi- GPU Systems using CUDA Inter-process Communication" *The Second International Workshop on Accelerators and Hybrid Exascale Systems (AsHES), IPDPS*, Shanghai, China, May 25, 2012.
61. Chathura Herath, Fang Liu, Suresh Marru, Lahiru Gunathilake, Masha Sosonkina, James P. Vary, Pieter Maris, Marlon Pierce, "Web Service and Workflow Abstractions to Large Scale Nuclear Physics Calculations", (SCC2012-2101). Proceedings of ICWS 2012, Honolulu, HI June 24-29, 2012.

F.1.5 Education and Outreach 1.6

F.2 Publications from XSEDE Users

Over the course of XSEDE's first year, users identified hundreds of publications and conference papers that were published, in press, accepted, submitted, or in preparation each quarter. The totals from the quarterly reports were 662 (Q3 2011), 535 (Q4 2011), 638 (Q1 2012), and 709 (Q2 2012, below). Thus, in the past year, more than 2,500 publications and manuscripts have benefitted from access to XSEDE resources.

The following publications were gathered from Research submissions to the June 2012 XSEDE Resource Allocations Committee (XRAC) meeting. Renewal submissions are required to provide a file specifically to identify publications resulting from the work conducted in the prior year. The publications are organized by the proposal with which they were associated. This quarter, users identified 709 publications and conference papers and presentations that were published, in press, accepted, submitted, or in preparation.

ACS110028

1. T. Kim and D. You, Large-eddy simulation of turbulent flow over a two-blade horizontal wind turbine rotor, in preparation.

AST060032

2. "Cyberinfrastructure to Support Science and Data Management for the Dark Energy Survey" Ngeow, C. Mohr, J.J., Alam, T., Barkhouse, W.A., Beldica, C., Cai, D., Daues, G., Plante, R. Annis, J., Lin, H., Tucker, D. and Smith, C. 2006, SPIE, 6270, 68-79.
3. "The Dark Energy Survey Data Management System" Mohr, J.J., Adams, D., Barkhouse, W., Beldica, C., Bertin, E., Cai, Y.D., da Costa, L.A.N., Darnell, J.A., Daues, G.E., Jarvis, M., Gower, M., Lin, H., Martelli, L., Neilsen, E., Ngeow, C., Ogando, R.L.C., Parga, A., Sheldon, E., Tucker, D., Kuropatkin, N. & Stoughton, C. for the Dark Energy Survey Collaboration 2008, SPIE, 7016, 17.

AST080028

4. Efficiency of Magnetic to Kinetic Energy Conversion in a Monopole Magnetosphere, Tchekhovskoy, A., McKinney, J. C., Narayan, R., ApJ, 699, 1789 (2009)
5. Numerical Studies of Relativistic Jets, Narayan, R., Tchekhovskoy, A., McKinney, J., in Accretion and Ejection in AGNs: A Global View, ASP Conf. Series, Vol. 427, p127 (2010)
6. Simulations of Magnetized Discs Around Black Holes: Effects of Black Hole Spin, Disc Thickness and Magnetic Field Geometry, Penna, R. F., McKinney, J. C., Narayan, R., Tchekhovskoy, A., Shafee, R., McClintock, J. E., MNRAS, 408, 752 (2010)
7. Magnetohydrodynamic Simulations of Gamma-Ray Burst Jets: Beyond the Progenitor Star, Tchekhovskoy, A., Narayan, McKinney, J. C., New Astron., 15, 749 (2010)

8. Black Hole Spin and the Radio Loud/Quiet Dichotomy of Active Galactic Nuclei, Tchekhovskoy, A., Narayan, R., McKinney, J. C., ApJ, 711, 50 (2010)
9. Efficient Generation of Jets from Magnetically Arrested Accretion on a Rapidly Spinning Black Hole, Tchekhovskoy, A., Narayan, R., McKinney, J. C. 2011, MNRAS, 418, L79 (2011)
10. Measuring Black Hole Spin by the Continuum-Fitting Method: Effect of Deviations from the Novikov-Thorne Disc Model, Kulkarni, A. K., Penna, R. F., Shcherbakov, R. V., Steiner, J. E., Narayan, R., Sadowski, A., Zhu, Y., McClintock, J. E., Davis, S. W., McKinney, J. C., MNRAS, 414, 1183 (2011)
11. General Relativistic Modeling of Magnetized Jets from Accreting Black Holes, Tchekhovskoy, A., McKinney, J. C., Narayan, R., Proc. Conf. "The Central Kiloparsec in Galactic Nuclei: Astronomy at High Angular Resolution 2011", J Phys: Conf. Series (JPCS), in press (arXiv1202.2864)
12. The Eye of the Storm: Light from the Inner Plunging Region of Black Hole Accretion Discs, Zhu, Y., Davis, S. W., Narayan, R., Kulkarni, A. K., Penna, R. F., McClintock, J. E., Apj, submitted (arXiv1202.1530)

AST090007

13. Dolence, J. Dynamical and Radiative Models of Low-Luminosity Active Galactic Nuclei. Ph.D. Thesis, University of Illinois at Urbana-Champaign.
13. Moscibrodzka, M., Shiokawa, H., Gammie, C. F., Dolence, J. C. 2012. The Galactic Center Weather Forecast. The Astrophysical Journal, in press. ArXiv e-prints arXiv:1204.1371.
14. Dolence, J. C., Gammie, C. F., Shiokawa, H., Noble, S. C. 2012. Near-infrared and X-Ray Quasi-periodic Oscillations in Numerical Models of Sgr A*. The Astrophysical Journal 746, L10.
15. Shiokawa, H., Dolence, J. C., Gammie, C. F., Noble, S. C. 2012. Global General Relativistic Magnetohydrodynamic Simulations of Black Hole Accretion Flows: A Convergence Study. The Astrophysical Journal 744, 187.
16. Mo´scibrodzka, M., Gammie, C. F., Dolence, J. C., Shiokawa, H. 2011. Pair Production in Low-luminosity Galactic Nuclei. The Astrophysical Journal 735, 9.
17. Mo´scibrodzka, M., Gammie, C. F., Dolence, J., Shiokawa, H., Leung, P. K. 2011. Numerical Models of Sgr A*. The Galactic Center: a Window to the Nuclear Environment of Disk Galaxies 439, 358.
18. Guan, X., Gammie, C. F. 2011. Radially Extended, Stratified, Local Models of Isothermal Disks. The Astrophysical Journal 728, 130.
19. Dolence, J., Gammie, C. F., Moscibrodzka, M., Shiokawa, H., & Leung, P.-K. Self-Consistent Light Curves from 3D GR-MHD Models of Sgr A*. In Preparation.
20. Dolence, J., Gammie, C. F., Shiokawa, H. Self-Consistent Light Curves from 3D GR-MHD Models of Sgr A*. In Preparation.
21. Moscibrodzka, M., Gammie, C. F. Steady State Pair Winds from Black Holes. In Preparation. Shiokawa, H., & Gammie, C. F. GRMHD Models of Radiatively Inefficient Disks: Variation of the Equation of State. In Preparation.

AST100004

22. Hallman, E. J., & Jeltema, T. E. 2011. Structure and turbulence in simulated galaxy clusters and the implications for the formation of radio haloes. MNRAS, 418(Dec.), 2467–2480.
23. Shull, J. M., Harness, A., Trenti, M., & Smith, B. D. 2012. Critical Star Formation Rates for Reionization: Full Reionization Occurs at Redshift z _ 7. ApJ, 747(Mar.), 100.
24. Skillman, S. W., Xu, H., Hallman, E. J., O'Shea, B. W., Burns, J. O., Li, H., Norman, M. L., & Collins, D. 2012. Cosmological MHD Simulations of Galaxy Cluster Radio Relics. In preparation, Apr.
25. Syphers, D., Anderson, S. F., Zheng, W., Smith, B., Pieri, M., Kriss, G. A., Meiksin, A., Schneider, D. P., Shull, J. M., & York, D. G. 2011. He II Ly_ Gunn-Peterson Absorption: New HST Observations and Theoretical Expectations. ApJ, 742(Dec.), 99.

AST100018

26. LSST Data Products and User Interfaces, Shaw, R. and {Axelrod}, T. and {Becker}, A.~C. and {Bickerton}, S. and {Juric}, M. and {Kantor}, J. and {Krughoff}, S. and {Lupton}, R.~H. and {Van Dyk}, S. and {Data Management}, L. and {Simulations Teams}American Astronomical Society Meeting Abstracts, 2012, Vol. 219, Series American Astronomical Society Meeting Abstracts, January
27. End-to-end Tests of LSST Science Cases with Image Simulations: Rare Astrometric Targets and Ultra-faint Dwarf Galaxies Author Juric, M. and Monet, D. and Gizis, J. E. and Sesar, B. and Willman, B. and Geha, M. and Fadley, R. and Krughoff, K. S. and Gibson, R. R. and Connolly, A. J. and Lupton, R. H. and Peterson, J. R. and Jernigan, G. J. and Silvestri, N. M. and LSST Data Management Team and LSST Image Simulation Team, 2012, Vol. 219, LSST Corporation. AAS Meeting #219, American Astronomical Society, January.

AST110034

28. J. E. Sta_, et al. 2012 \Do R Coronae Borealis stars form from Double White Dwarf Mergers? ", Submitted to the Astrophysical Journal.

AST120042

29. The Voronoi Tessellation Cluster Finder in 2+1 Dimensions, Soares-Santos, M et al., 2011, ApJ, 717, 45.
30. Cosmological Constraints from Galaxy Clustering and the Mass-to-number Ratio of Galaxy Clusters, Tinker, J et al., 2011, ApJ, 745, 16.
31. Sample variance in photometric redshift calibration: cosmological biases and survey requirements, Cunha, C.E., Huterer, D., Busha, M.T., Wechsler, R.H., 2012, submitted to ApJ, (arXiv:1109.5691).
32. A Measurement of the Correlation of Galaxy Surveys with CMB Lensing Convergence Maps from the South Pole Telescope, Bleem et al., 2012, submitted to ApJ, (arXiv:1203.4808).
33. A High-throughput Workflow for Cosmological Simulations, Erickson, B.M.S., et al., 2012, in prep.

ATM40001

34. El-Alaoui, M., M. Ashour-Abdalla, R. L. Richard, M. L. Goldstein, J. M. Weygand, and R. J. Walker (2010), Global magnetohydrodynamic simulation of reconnection and turbulence in the plasma sheet, *J. Geophys. Res.*, 115, A12236, doi:10.1029/2010JA015653.
35. El-Alaoui, M., R. L. Richard, M. Ashour-Abdalla, R. J. Walker, and M. L. Goldstein, Turbulence in a Global Magnetohydrodynamic Simulation of the Earth's Magnetosphere during Northward and Southward Interplanetary Magnetic Field, *Nonlinear Process in Geophysics, submitted,* 2011. Nonlin. Processes Geophys., 19, 165–175, 2012 www.nonlin-processes-geophys.net/19/165/2012/doi:10.5194/npg-19-165-2012
36. El-Alaoui, M., M. Ashour-Abdalla, R. L. Richard, and M. L. Goldstein, Dependence of Magnetotail Reconnection on the IMF: Idealized MHD Simulations Revisited, Yosemite 2010 Workshop on Magnetic Reconnection, 8–12 February 2010, Yosemite, CA.
37. El-Alaoui, M., R.L. Richard, M. Ashour-Abdalla, M.L. Goldstein, J.M. Weygand, and R.J. Walker (2010), Global magnetohydrodynamic simulations of turbulence in the plasma sheet, Abstract SM51C-1835 presented at 2010 Fall Meeting, San Francisco, Calif., 13-17 Dec.

ATM080017

38. Boé J, Hall A, Colas F, McWilliams JC, Qu X, Kurian J, Kapnick SB, Frenzel H (2010) What shapes mesoscale wind anomalies in coastal upwelling zones?, *Clim. Dyn.* 36, 2037-2049, doi: 10.1007/s00382-011-1058-5
39. Pavelsky T, Kapnick S, Hall A (2011) Accumulation and melt dynamics of snowpack from a multiresolution regional climate model in the central Sierra Nevada, California. *JGR--Atmospheres.* Vol. 116: D16115, doi:10.1029/2010JD015479
40. Sun F, Hall A, Walton D, Qu X, Capps S, Huang J, Berg N, Schwartz M, Jousse A, Cerezo-Mota R, Nakamura M, Bunje P (2012): Mid-Century Warming in the Los Angeles Region and its Uncertainty, *In preparation.*
41. Walton, D, Hall A, Frenzel H and Colas F (2012): "Forced and Unforced SST Variability off the California Coast", *In preparation*

ATM090042

42. Bao, X., F. Zhang, and J. Sun, 2011: Diurnal variations of warm-season precipitation east of the Tibetan Plateau over China. *Mon. Wea. Rev.,* **139,** 2790-2810.
43. Fang, J., and F. Zhang, 2012: Beta effect on the development of tropical cyclones. *Mon. Wea. Rev.,* accepted subject to minor revisions.
44. Green, B. W., F. Zhang, and P. M. Markowski, 2011: Multi-scale processes leading to supercells in the landfalling outer rainbands of Hurricane Katrina (2005). *Wea. Forecasting,* **26,** 828-847.
45. Meng, Z., **F. Zhang**, P. Markowski, D. Wu, K. Zhao, 2012: Mechanisms for the Formation of Rear Inflow and Bow Echo within a Squall Line over South China. *Journal of the Atmospheric Sciences*, doi: http://dx.doi.org/10.1175/JAS-D-11-0121.1.
46. Poterjoy, J., and F. Zhang, 2011: Dynamics and structure of forecast error covariance in the core of a developing hurricane. *J. Atmos. Sci.,* **68,** 1586-1606.
47. Qian, T., C. C. Epifanio, and F. Zhang, 2012: Topographic effects on the tropical land and sea breeze. *J. Atmos. Sci.,* accepted subject to revisions.

48. Rozoff, C. M., D. S. Nolan, J. P. Kossin, **F. Zhang**, and J. Fang, 2012: The roles of an expanding wind field and inertial stability in tropical cyclone secondary eyewall formation. *Journal of the Atmospheric Sciences*, in press.
49. Stern, D. P. and **F. Zhang**, 2012: How does the eye warm? A potential temperature budget and trajectory analysis of an idealized tropical cyclone. *Journal of theAtmospheric Sciences*, accepted subjected to revision.
50. Sun, J., and F. Zhang, 2012: Impacts of mountain-plains solenoid on diurnal variations of rainfalls along the Mei-yu front over the East China Plains. *Mon. Wea. Rev.*, **140,** 379-397.
51. Weng, Y., M. Zhang, and F. Zhang, 2011: Advanced data assimilation for cloudresolving hurricane initialization and prediction. *Computing in Science and Engineering*,**13,** 40-49.
52. Weng, Y., and F. Zhang, 2012: Assimilating airborne Doppler radar observations with an ensemble Kalman filter for cloud-resolving hurricane initialization and prediction: Katrina (2005). *Mon. Wea. Rev.*, **140,** 841-859.
53. Xie, B and **F. Zhang**, 2012: Impacts of typhoon track, island topography and monsoon flow on the heavy rainfalls in Taiwan associated with Morakot (2009). Monthly Weather Review, accepted subjected to minor revision.
54. Xie, B. and **F. Zhang,** Q. Zhang, J. Poterjoy and Y. Weng, 2012: Observing Strategy and Observation Targeting for Tropical Cyclones using Ensemble-based Sensitivity Analysis and Data Assimilation. *Journal of Geophysical Research*, submitted.
55. Zhang, F., 2011: The future of hurricane prediction. *Computing in Science and Engineering*, **13,** 9-12.
56. Zhang, M., and F. Zhang, 2012: E4DVar: Coupling an ensemble Kalman filter with four dimensional variational data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, **140,** 587-600.
57. Zhang, M., F. Zhang, X.-Y. Huang, and X. Zhang, 2011: Inter-comparison of an ensemble Kalman filter with three- and four dimensional variational data assimilation methods in a limited-area model during the month of June 2003. *Mon. Wea. Rev.*, **139,** 566-572.
58. Zhang, F., M. Zhang, J. Poterjoy, 2012: E3DVar: Coupling an ensemble Kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar. *Monthly Weather Review*, submitted.

ATM090046

59. Y. Omelchenko, H. Karimabadi, HYPERS: A unidimensional asynchronous framework for multiscale hybrid simulations, J. Comp. Physics, 231, 1766, 2012.
60. V. Roytershteyn, W. Daughton, H. Karimabadi, and F. S. Mozer, Influence of the Lower-Hybrid Drift Instability on Magnetic Reconnection in Asymmetric Configurations, Phys. Rev. Lett., in press, 2012.
61. W. Daughton, V. Roytershteyn, H. Karimabadi, L. Yin, B.J. Albright, B. Bergen and K.J. Bowers, Role of Electron Physics in the Development of Turbulent Magnetic Reconnection in Collisionless Plasmas, Nature Physics 7, 539-542, 2011, doi:10.1038/nphys1965
62. J. Egedal, W. Daughton and A. Le, Large-scale electron acceleration by parallel electric fields during magnetic reconnection, Nature Physics, published online Feb. 26| DOI: 10.1038/NPHYS2249, 2012
63. T.E. Moore, J.L. Burch, W.S. Daughton, S.A. Fuselier, H. Hasegawa, S.M. Petrinec, Zuyin Pu, Multiscale Studies of the Three-Dimensional Dayside X-line, submitted to Journal of Atmospheric and Solar-Terrestrial Physics, 2012
64. H. Ji and W. Daughton, Phase Diagram for Magnetic Reconnection in Heliophysical, Astrophysical and Laboratory Plasmas, Phys. Plasmas 18, 111207, 2011
65. J.D. Scudder, R.D. Holdaway, W. S. Daughton, H. Karimabadi, V. Roytershteyn, C.T. Russell, and J.Y. Lopez, First Resolved Observations of the Demagnetized Electron Diffusion Region, in press, Phys. Rev. Lett., 2012
66. L.-J. Chen, W. Daughton, A. Bhattacharjee, R. B. Torbert, V. Roytershteyn, N. Bessho, In-plane electric fields in secondary and primary plasmoids generated during magnetic reconnection, submitted to Geophys. Res. Lett., 2012
67. H. Karimabadi, V. Roytershteyn, M. Wan, W. H. Matthaeus, W. Daughton, B. Loring, P. Wu, M. Shay, J. Borovsky, E. Leonardis, S. Chapman, T. Nakamura, Coherent Structure, Intermittent Turbulence and Reconnection Mediated Dissipation at Kinetic Plasma Scales, submitted to Nature, 2012
68. A. Le, H. Karimabadi, J. Egedal, V. Roytershteyn, W. Daughton, Electron energization during magnetic island coalescence, submitted, Phys. Plasmas, 2012.
69. Daughton, W., V. Roytershteyn, H. Karimabadi, L. Yin, B. J. Albright, B. Bergen and K. J. Bowers, Role of electron physics in the development of turbulent magnetic reconnection in collisionless plasmas, Nature Physics, DOI: 10.1038/NPHYS1965, 10 April, 2011.
70. Karimabadi, H., V. Roytershteyn, C.G. Mouikis, L.M. Kistler, W. Daughton, Flushing effect in reconnection: Effects of minority species of oxygen ions, Planetary and Space Science, 59, 526, 2011.

71. Karimabadi, H., J. Dorelli, V. Roytershteyn, W. Daughton, and L. Chacon, Flux pileup in collisionless magnetic reconnection: bursty interaction of large flux ropes, Phys. Rev. Lett., 107, 025002, 2011.
72. Karimabadi, H., and J. Dorelli, H. X. Vu, B. Loring, Y. Omelchenko, Is quadrupole structure of out-of-plane magnetic field evidence of Hall reconnection?, to appear in *Modern Challenges in Nonlinear Plasma Physics*, editor D. Vassiliadis, AIP conference, 2010.

ATM100017

73. Maliawco Jr., R.J., G.M. McFarquhar, and B.F. Jewett, 2012: An ensemble approach to examining the role of cloud condensation nuclei on the evolution of idealized tropical cyclones. *Amer. Meteor. Soc. 30th Conf. Hurricanes Trop. Meteor.*, Ponte Vedra Beach, FL, 15-20 April 2012.
74. McFarquhar, G.M., and J. Um, 2011: Quantifying and reducing uncertainty in observed shapes and concentrations of small ice crystals: implications for derived scattering properties. Invited Presentation, *EGU general Assembly 2011*, 13, EGU2011-574.
75. McFarquhar, G.M., and J. Um, 2012: Quantifying and reducing uncertainty in observed sizes, shapes and concentrations of small ice crystals: Implications for derived scattering properties at visible wavelengths. *International Radiation Symposium*, Berlin, Germany, Submitted.
76. McFarquhar, G.M., 2011: Considerations in the development of microphysical and singlescattering representations of ice clouds, *Atmospheric Systems Research Cloud LifetimeWorking Group Fall meeting*, Annapolis, MD.
77. McFarquhar, G.M., B.F. Jewett, S. Nesbitt, M. Gilmore, and T.-L. Hsieh, 2012: Vertical velocity and microphysical distributions related to the rapid intensification of Hurricane Dennis (2005). *J. Atmos. Sci.*, Submitted.
78. Um., J., and G.M. McFarquhar, 2011: Dependence of the single-scattering properties of small ice crystals on idealized models. *Atmos. Chem. Phys.*, 11, 3159-3171.
79. Um. J., and G.M. McFarquhar, 2012: Impacts of shapes and concentrations of small ice crystals on bulk scattering properties of tropical cirrus. *J. Geophys. Res.*, in preparation.
80. Um, J., and G.M. McFarquhar, 2012: Optimal numerical methods for determining the orientation average of single-scattering properties of atmospheric ice crystals. *International Radiation Symposium*, Berlin, Germany, Submitted.
81. Um, J., and G.M. McFarquhar, 2012: Single-scattering properties of small atmospheric ice crystals at absorbing wavelengths: Dependence on idealized models. *International Radiation Symposium*, Berlin, Germany, Submitted.

BCS100001

82. Ascenzi M-G, Kawas N, Lutz A, Kardas D, Nackenhorst U, Keyak J (2012) Individual-specific, multi-scale finite element simulation of the proximal femur cortex. J Comp Physics. Special Issue on Multi-scale Modeling. Guest Editor Ching-Long Lin. In print.

CCR110017

83. J. Clark. Non-linear learning via feature induction. Ph.D. thesis proposal, Carnegie Mellon University, 2012.
84. J. Clark and A. Lavie. Single system, multiple domain adaptation via feature induction. In Proceedings of the Association for Machine Translation in the Americas (In Preparation), 2012.
85. M. Denkowski, G. Hanneman, and A. Lavie. The CMU-Avenue French-English Translation System. In Submitted to the Seventh Workshop on Statistical Machine Translation, 2012.
86. G. Hanneman. Automatically improved category labels for syntax-based statistical machine translation. Ph.D. thesis (in progress), Carnegie Mellon University.
87. K. Heafield, H. Hoang, P. Koehn, T. Kiso, and M. Federico. Left language model state for syntactic machine translation. In Proceedings of the International Workshop on Spoken Language Translation, San Francisco, CA, USA, December 2011.
88. K. Heafield, P. Koehn, and A. Lavie. Language model rest costs and space-efficient storage. In Proceedings of Empirical Methods in Natural Language Processing and Natural Language Learning, Jeju, Korea, July 2012. Under Review.

CHE080006

89. J. Jelic and R. Meyer, "A DFT Study of Pseudomorphic Monolayer Pt Catalysts for Water Gas Shift," *Journal of Catalysis*, **272** (2010) 151-157
90. H. Wei, C. Gomez, and R. J. Meyer, "A Comparative Density Functional Theory Study of Water Gas Shift over PdZn(111) and NiZn(111)" Topics in Catalysis, accepted

CHE080076

91. Azaria S. Eisenberg , Vasiliy Zenamenskiy , Iya Likhtina, and Ronald L. Birke, "Electronic Spectroscopy and Computational Studies of Glutathionylco(III)balamin " J. Chem. Phys. A accepted for pub. (2012)
92. Lombardi, R. L. and Birke, R. L., "The theory of surface-Enhanced Raman Scattering" , J. Chem. Phys. , 136, in press (2012)

CHE090065

93. J. Mondal and A.Yethiraj, "Driving force for the association of amphiphilic molecules" *J. Phys. Chem. Lett.*, 2, 2391(2011)
94. Z. Wu, Q. Cui, and A. Yethiraj, "A new coarse-grained force field for membrane-peptide simulations", *J. Chem. Theory. Comput*. 7, 3793(2011)
95. Z. Wu, Q. Cui, and A. Yethiraj, "The driving force for the association of hydrophobic peptides: The importance of electrostatic interactions in coarse-grained water models", *J. Phys. Chem. Lett*. 2, 1794(2011)
96. J. Mondal, X. Zhu, Q. Cui, and A.Yethiraj, "Sequence-dependent pK(a) shift induced by molecular self-assembly: insights from computer simulation" *J. Phys. Chem. B*, 116, 491(2012)
97. J. Mondal, B. P. Bratton, Y. J. Li, A.Yethiraj, and J. C. Weisshaar "Entropy-Based mechanism of ribosom-nucleoid segregation in E. coli cells" *Biophys. J.*, 100, 2605 (2011)
98. Z. Wu, Q. Cui, and A. Yethiraj, "Coarse-grained model studies o peptide-induced lipid self-assembly phases", To be submitted
99. J. S. Kim, Z. Wu, and A. Yethiraj, "Molecular dynamics study of structure-breaking effect of ions in aqueous salt solutions", To be submitted

CHE100040

100. Kormos, B. L., S. N. Pieniazek, D. L. Beveridge, and A. M. Baranger. 2011. U1A protein-stem loop 2 RNA recognition: prediction of structural differences from protein mutations. Biopolymers **95:591-606.**
101. Dixit, S. B., M. M., and D. L. Beveridge. 2012. Studies of Base Pair Sequence Effects on DNA Solvation Based on All-atom Molecular Dynamics Simulations. Journal of Biosciences In Press.
102. Beveridge, D. L., T. E. Cheatham 3rd, and M. Mezei. 2012. The ABC's of Molecular Dynamics Simulations on B-DNA, c. 2012. Journal of Biosciences In Press 2012.
103. Wheatley, E. G., S. N. Susan N. Pieniazek, I. Mukerji, and D. L. Beveridge. 2011. Molecular Dynamics of a DNA Holliday Junction: The Inverted Repeat Sequence d(CCGGTACCGG)4. Biophysical Chemistry **102**:552-560.
104. Pieniazek, S. N., M. M. Hingorani, and D. L. Beveridge. 2011. Dynamical Allosterism in the Mechanism of Action of the DNA Mismatch Repair Protein MutS. Biophysical Journal **101**:1730-1739.
105. Wheatley, E. G., S. N. Pieniazek, C. I. Vitoc, I. Mukerji, and D. L. Beveridge. 2012. Structure Prediction of a Novel Protein-DNA Complex: HU with a DNA Four-Way Junction. In Innovations in Biomolecular Modeling and Simulation. T.Schlick, editor. In Press 2012.

CHE100082

106. J. McCarty, A. Clark, I. Y. Lyubimov, M. G. Guenza *"Thermodynamic Consistency in HighlyCoarse-Grained Polymer Melts"* (to be submitted).
107. J. McCarty, I.Y. Lyubimov, A. Clark, M.G. Guenza, *"Accurate Determination of theThermodynamic Properties of Polyethylene Melts from Highly Coarse-Grained Simulations: Comparison with United Atom Molecular Dynamics Simulations"* (in preparation)
108. James McCarty and Marina Guenza – APS March Meeting, Boston, Massachusetts, *"How*
109. *Reliable Are Soft Potentials? Ensuring Thermodynamic Consistency Between Hierarchical Models of Polymer Melts"* February 27- March 2, 2012.
110. A. Clark, J. McCarty, I. Y. Lyubimov, M. G. Guenza *"Pressure in soft matter representations of Polymer Melts at varying levels of coarse graining"* (to be submitted).
111. A. Clark, J. McCarty, I. Y. Lyubimov, M. G. Guenza *"Analytical characterization of potentials forsoft matter coarse graining of polymer melts"* (in preparation).
112. Anthony Clark and Marina Guenza – APS March Meeting, Boston, Massachusetts, *"Representation of polymer melts as soft liquids with effective pair interactions"* February 27-March 2, 2012.
113. I. Y. Lyubimov, and M. G. Guenza *"A First Principle Approach to Rescale the Dynamics of Simulated Coarse-Grained Macromolecular Liquids"*, Physical Review E 84, 031801 (2011).
114. I. Y. Lyubimov, and M. G. Guenza *"Theoretical Reconstruction of Realistic Dynamics of Highly-Coarse-Grained cis-1,4-Polybutadiene Melts"* (in preparation).

115. Ivan Y. Lyubimov and Marina Guenza – APS March Meeting, Boston, Massachusetts, *"Reconstructing Polymer Melt Dynamics Sped up due to Large-Scale Coarse-Graining"* February 27- March 2, 2012

CHE110041

116. Christopher Von Bargen, Christopher MacDermaid, Colleen, Murrett, One-Sun Lee, Michael J.Therien, and Jeffery G. Saven. "The Role of Ionic Side Chains in the Helical Wrapping of Phenylene Ethynylene Polymers about Single-Walled Carbon Nanotubes." In preparation.
117. Christopher Von Bargen, Christopher MacDermaid, One-Sun Lee, Michael J. Therien, and Jeffery G. Saven. "Alternate helical wrapping modes of Naphthylene Ethynylene Polymers about Single-Walled Carbon Nanotubes." In preparation.

CHE110045

118. Johannes Hachmann, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S. Sanchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M. Brockway, and Alan Aspuru-Guzik. *"The Harvard Clean Energy Project: Large-Scale ComputationalScreening and Design of Organic Photovoltaics on the World Community Grid."*, J. of Physical Chemistry Letters, 2, 22412251, (2011).
119. Johannes Hachmann, Supriya Shrestha, Anna Brockway, Sule Atahan-Evrenk, Alan AspuruGuzik. "*Organic electronics training set for calibration of electronic energy levels from Density Functional Theory*", in preparation.
120. Sule Atahan-Evrenk**,** Semion K. Saikin, Stephanie Valleau and Alan Aspuru-Guzik, "*Theoretical Modeling of exciton diffusion in di-anthra thieno-thiophene*", in preparation.
121. Sule Atahan-Evrenk, Leon Liu, Anna Brockway and Alan Aspuru-Guzik, "*Correlationbetween reorganization energy and organic field-effect transistor mobility*", in preparation.

CHE110064

122. S. Wen and G. Beran, \Accidental degeneracy in crystalline aspirin: New insights from high-level ab initio calculations" Cryst. Growth Des. in press (2012). DOI: 10.1021/cg300358n
123. S. Wen, K. Nanda, Y. Huang, and G. Beran. \Practical quantum mechanics-based fragment methods for predicting molecular crystal properties." Phys. Chem. Chem. Phys. in press (2012). DOI: 10.1039/c2cp23949c
124. E. Gillis, M. Demireva, K. Nanda, G. Beran, E. Williams, and T. Fridgen. "Structures and energetics of electrosprayed uracilnCa2^{+} clusters (n=14-4) in the gas phase." Phys. Chem. Chem. Phys. 14, 3304{3315 (2012). DOI: 10.1039/c1cp22984b Chosen as the journal cover article.
125. S. Wen and G. Beran. \Accurate molecular crystal lattice energies from a fragment QM/MM approach with on-the-y ab initio force-_eld parameterization." J. Chem. Theory Comput. 7, 3733-3742 (2011). DOI: 10.1021/ct200541h
126. K. Nanda and G. Beran, \Reliable organic molecular crystal structure optimization using a fragment-based QM/MM method." In preparation (2012).

CHE110065

127. X. Xiao, C. K. Hall, P. F. Agris. "Molecular Recognition Mechanism of Peptide Chain Bound to the tRNALys3 Anticodon Loop in vitro". To be submitted
128. X. Xiao, C. K. Hall, P. F. Agris. "The design of peptide sequence as an inhibitor of HIV replication ---- An algorithm combining Monte Carlo and self-consistent mean field techniques". To be submitted.

CHE120051

129. Kinz-Thompson, C.; Conwell, E. J. Phys. Chem. Lett. 2010, 1, 1403.
130. Kravec, S. M.; Kinz-Thompson, C. D.; Conwell, E. M. J. Phys. Chem. B 2011, 115, 6166.

CTS070002

131. J.R. Finn and S.V. Apte, "Characteristics of vortical structures in random and arranged packed beds of spheres." ASME Fluids Engineering Summer Meeting, Submitted for publication(FEDSM2012-72354), 2012.
132. J.R. Finn and S.V. Apte.,"Relative performance of body- tted and ctitious-domain simulations of flow through porous media." ASME Fluids Engineering Summer Meeting, Submitted for publication(FEDSM2012-72355), 2012.
133. K.J. Drost, T. Jackson, R. Haggerty, and S.V. Apte, "RANS predictions of turbulent scalar transport in dead zones of natural streams" ASME Fluids Engineering Summer Meeting, Submitted for publication(FEDSM2012-72380), 2012.

134. J.R. Finn, Apte S.V., B. Wood, V. Patil, and J. Liburdy., "Unsteady vortical structures in porous media flows.," Baltimore, MD, November 20-22 2011. American Physical Society Division of Fluid Dynamics Annual Meeting.

CTS080042

135. Roman S. Voronov, "Local fluid stress and nutrient transport effects via simulations of perfusion through bone tissue engineering scaffolds," PhD Dissertation, University of Oklahoma, Norman (2010).
136. Anjani K.C. Varthy Karna, "Computational investigation of the effects of turbulence structures on transport mechanisms," MS Thesis, University of Oklahoma, Norman (2011).
137. Pham, N.H., Voronov, R.S., VanGordon, S., Sikavitsas, V.I., and *D.V. Papavassiliou,* "Predicting the stress distribution within scaffolds with ordered architecture," submitted to *Biorheology* (2012).
138. Srinivasan, C., and D.V. Papavassiliou, "Comparison of backwards and forwards scalar relative dispersion in turbulent shear flow," submitted to *Int. J. Heat and Mass Transfer* (2012).
139. Voronov, R.S., VanGordon, S., Sikavitsas, V.I., R.L. Shambaugh and D.V. Papavassiliou, "3D Histology of Tissue Engineered Constructs via High Resolution Microcomputer Tomography Without X-Ray Contrast," submitted to *Tissue Engineering* (2012).
140. Bui, K., and D.V. Papavassiliou, "Numerical calculation of the effective thermal conductivity of nanocomposites," submitted to *Numerical Heat Transfer, Part A: Applications* (2012)
141. Pham, N., and D.V. Papavassiliou, "Flow Recovery Downstream from Nanoposts Grown at the Wall of a Microchannel," submitted to *Microfluidics and Nanofluidics* (2012)
142. Karna, A.K., and D.V. Papavassiliou, "Near-wall velocity structures that drive turbulent transport from a line source at the wall," *Physics of Fluids,* **24,** 035102, 21 pages, 2012.
143. VanGordon, S., Voronov, R.S., Blue, T.B., Shambaugh, R.L., Papavassiliou, D.V., and V.I. Sikavitsas, "Effects of scaffold architecture on preosteoblastic cultures under continuous fluid shear," *Ind. Eng. Chem.. Res.,* **50(2),** 620-629, 2011.
144. Srinivasan, C., and D.V. Papavassiliou, "Direction of scalar transport in turbulent channel flow," *Physics of Fluids,* **23(11),** 115105, 21 pages, 2011.
145. Bui, K., Duong, H.M., Striolo, A., and D.V. Papavassiliou, "Effective heat transfer properties of graphene sheet nanocomposites and comparison to carbon nanotube nanocomposites," *Journal of Phys. Chemistry C,* **115(10),** 3872-3880, 2011.
146. Srinivasan, C. and D.V. Papavassiliou, "Prediction of turbulent Prandtl number in wall flows with Lagrangian simulations," *Ind. Eng. Chem. Res.,* **50(15),** 8881-8891, 2011.
147. Konatham, D., Bui, K.N.D., Papavassiliou, D.V., and A. Striolo, "Simulation Insights on Thermally Conductive Graphene-Based Nanocomposites," *Molecular Physics,* **109(1),** 97-111, 2011
148. Voronov, R.S., VanGordon, S., Sikavitsas, V.I., and D.V. Papavassiliou, "Efficient Lagrangian scalar tracking method for reactive local mass transport simulation through porous media," *Int. J. for Numerical Methods in Fluids,* **67,** 501-517, 2011.
149. Voronov, R.S., VanGordon, S., Sikavitsas, V.I., and D.V. Papavassiliou, "Distribution of flow- induced stresses in highly porous media," *Appl. Phys. Lett.,97,* 023701, 2010
150. Voronov, R., VanGordon, S., Sikavitsas, V.I., and D.V. Papavassiliou, "Local velocity and stress fields within 3D porous scaffolds used in perfusion bioreactors for bone tissue growth," *J. of Biomechanics,* **43,** 1279-1286, 2010.
151. Duong, H.M., Yamamoto, N., Bui, K., Papavassiliou, D.V., Maruyama, S., B.L. Wardle, "Morphology effects on non-isotropic thermal conduction of aligned single- and multi-walled carbon nanotubes in polymer nanocomposites," Journal of Phys. Chemistry C, 114 (19), 8851-8860, 2010 16. Srinivasan, C., and D.V. Papavassiliou, "Backwards and forwards dispersion of a scalar in turbulent wall flows," Int. J. Heat Mass Transf. 53, 1023-1035, 2010.
152. Karna, A.K., and D.V. Papavassiliou, D.V., "Direct simulations of line source dispersion in a turbulent channel flow," Proceedings, *2nd NAUN Fluid Mechanics and Heat and Mass Transfer 2011,* Corfu, Greece, July 14-16, 2011
153. Papavassiliou, D.V., and C. Srinivasan, "Turbulent heat and mass transfer simulations using Lagrangian methods and direct numerical simulations," Proceedings, *7th GRACM Congress on Computational Mechanics,* Athens, Greece, June 30-July 2, 2011

CTS090004

154. Reichert, A., Heath, M. T., and Bodony, D. J. (2012) "Energy Stable Numerical Methods for Hyperbolic Partial Di_erential Equations Using Overlapping Domain Decomposition," J. Comput. Phys, accepted.
155. Ostoich, C., Bodony, D. J., and Geubelle, P. H. (2012) "Coupled fluid-thermal response of a spherical dome due to a Mach 6.59 laminar boundary layer," AIAA J., accepted.

156. Zhang, Q. & Bodony, D. J., (2012) “Numerical investigation and modeling of acoustically excited flow through a circular orifice backed by a hexagonal cavity,” J. Fluid Mech., Vol. 693, pp. 367–401.
157. Kim, J., Bodony D.J., Freund, J.B. (2012) “Mechanisms of Jet Noise Reduction and Their Impact on Large-Eddy Simulations,” J. Prop. Power, Vol. 28(2), pp. 259–268.
158. Austin, J. M., & Bodony, D. J. (2011) “Wave propagation in gaseous small-scale channel flows,” Shock Waves, Vol 21(6), pp. 547-557. Submitted archival journal articles
159. Mishra, A., Bodony, D. J., “Evaluation of actuator disk theory for predicting indirect combustion noise,” Submitted to Journal of Sound and Vibration (November, 2011, revision sent February 2012)
160. Kim,J.,Bodony,D.J.,andFreund,J.B.“AeroacousticsControlofaTurbulentMach1.3JetUs- ing Adjoint-Based Optimization,” AIAA Paper 2011-2832, Presented at the 17th AIAA/CEAS Aeroacoustics Conference (32nd AIAA Aeroacoustics Conference), Portland, Oregon, June 5-8, 2011.
161. Zhang, Q., and Bodony D. J. “Impedance Predictions of 3D Honeycomb Liner with Circular Apertures by DNS,” AIAA Paper 2011-2727, Presented at the 17th AIAA/CEAS Aeroacoustics Conference (32nd AIAA Aeroacoustics Conference), Portland, Oregon, June 5-8, 2011.
162. OstoichC.,Bodony,D.J.,Geubelle,P.H.“Predictionofheattransferintoathermallycompliant surface in a Mach 5.73 boundary layer using direct numerical simulation,” AIAA Paper 2011- 3706, Presented at the 41st AIAA Fluid Dynamics Conference and Exhibit, Honolulu, Hawaii, June 27-30, 2011.
163. Zhang,Q.,&Bodony,D.J.“Nearfieldoftransient,acousticallyforcedtransitionalandturbulent jets,” Bulletin of the American Physical Society, Vol. 56(18), 2011.
164. Bodony, D. J., & Natarajan, M. “Validated model of arc-filament plasma actuators for control of wall-bounded flows,” Bulletin of the American Physical Society, Vol. 56(18), 2011.
165. Kim, J., Bodony, D. J., & Freund, J. B. “Analysis of the near-field of an adjoint-based noise controlled Mach 1.3 turbulent jet,” Bulletin of the American Physical Society, Vol. 56(18), 2011.
166. Ostoich, C., Bodony, D. J., & Geubelle, P. H. “Thermal loads on a domed protuberance under a Mach 5.7 boundary layer,” Bulletin of the American Physical Society, Vol. 56(18), 2011.
167. 8. Natarajan, M. and Bodony, D. J. “Validation of an effective computational model for localized arc-filament plasma actuators and application to high subsonic boundary layer separation con- trol,” AIAA Paper 2012-189, Presented at the 50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, Nashville, Tennessee, Jan. 9-12, 2012.
168. Mishra, A. and Bodony, D. J. “Evaluation of actuator disk theory to predict indirect com- bustion noise from a linear turbine cascade,” AIAA Paper 2012-????, To be presented at the 18th AIAA/CEAS Aeroacoustics Conference (33rd AIAA Aeroacoustics Conference), Colorado Springs, CO, June, 4-6, 2012.
169. Zhang, Q., Jambunathan, R., and Bodony, D. J. “Numerical investigation of acoustically-excited flow through a circular orifice backed by a hexagonal cavity with and without grazing flow,” AIAA Paper 2012-????, To be presented at the 18th AIAA/CEAS Aeroacoustics Conference (33rd AIAA Aeroacoustics Conference), Colorado Springs, CO, June, 4-6, 2012.

CTS100002

170. Mittal, R., Erath, B., and Plesniak, M., "Fluid Dynamics of Human Phonation and Speech," *AnnualReview of Fluid Mechanics*. To appear 2013**.**
171. Bhardwaj, R. and Mittal, R., “Benchmarking a coupled Immersed-Boundary-Finite-Element Solver for large scale flow-induced deformation”, *AIAA Journal*. To appear, 2012.
172. Zheng, X., Seo, J. H., Vedula, V., Abraham, T., and Mittal, R., “Computational Modeling and Analysis of Intracardiac Flows in Simple Models of the Left Ventricle”, *European Journal of Mechanics- B/Fluids*. To appear, 2012.
173. Zheng, X., Mittal, R. and Bielamowicz, S. "A Computational Study of Asymmetric Glottal Jet Deflection during Phonation," *J. Acoustical Society of America*, Vol. 129(4), pp. 2133-2143, 2011.
174. Zheng, X., Mittal, R., Xue. Q. and Bielamowicz, "Direct-Numerical Simulation of the Glottal Jet and Vocal-Fold Dynamics in a Three-Dimensional Laryngeal Model," *J. Acoustical Society of America,* Vol. 130(1) pp. 404-415, 2011.
175. Seo, J. H. and Mittal, R., "A High-order Immersed Boundary Method for Acoustic Wave Scattering and Low Mach number Flow Induced Sound in Complex Geometries, " *J. of Computational Physics,* Vol. 230(4) pp. 1000-1019, 2011.
176. Mittal, R., Zheng, X., Bhardwaj, R., Seo, J. H., Xue, Q. and Beilamowicz, S., “Towards a simulation- based tool for the treatment of vocal fold paralysis”, *Frontiers in Computational Physiology and Medicine,* 2011, Vol. 2 (19).
177. Xue, Q., Zheng, X., Beilamowicz, S. and Mittal, R., “Sensitivity of vocal fold vibratory modes to their three-layer structure: Implications for computational modeling of phonation”, *J. Acoustical Society of America.,* 2011, Vol. 130 (2), pp: 965-976

178. Seo, J. H. and Mittal, R., "A sharp interface immersed boundary method with improved mass conservation and reduced spurious pressure oscillations", *Journal of Computational Physics*, 2011, Vol. 230, Issue 19, pp. 7347-7363
179. Seo, J. H. and Mittal, R., "A Coupled Flow-Acoustic Computational Study of Bruits from a Modeled Stenosed Artery," *Medical & Biological Engineering & Computing*, **in review.**
180. Xue, Q., Mittal, R., Zheng, X., and Beilamowicz, S., "Computational Modeling of Phonatory Dynamics in a Realistic Model of the Human Larynx," *J. Acoust. Soc. Am.*, **in review.**
181. Zheng, L., Hedrick, T., and Mittal, R., "A Multi-Fidelity Modeling Approach for Evaluation and Optimization of Wing Stroke Aerodynamics in Flapping Flight," *Journal of Fluid Mechanics*, **in review.**
182. Xue, Q., "Computational Modeling and Analysis of Fluid-Structure Interaction During Phonation in Healthy and Diseased Larynges", Department of Mechanical Engineering, Johns Hopkins University, Aug. 2011.
183.

CTS100041

184. Damilola A. Daramola and G.G. Botte. "Theoretical study of ammonia oxidation on platinum clusters - Adsorption of intermediate nitrogen dimer molecules (N2Hz: z = 0 - 4)", *in preparation*.
185. Damilola A. Daramola and G.G. Botte. "Theoretical study of ammonia oxidation on platinum clusters - Adsorption of ammonia and water fragments", *Computational and Theoretical Chemistry* http://dx.doi.org/10.1016/j.comptc.2012.02.032 (2012).
186. Damilola A. Daramola. "Theoretical Characterization of Ammonia Oxidation Species on Platinum Clusters" Ph.D. Dissertation, Ohio University, Athens, OH (2011).
187. Distributed Power from Wastewater. ERDC/CERL. Gerardine G. Botte. Final Report, September 2011.

CTS110028

188. C. Daley, M. Vanella, A. Dubey, K. Weide and E. Balaras, Optimization of multigrid based elliptic solver for large scale simulations in the FLASH code. *Concurrency Computat.: Pract. Exper.* In Press, (2012)
189. M. de Tullio, J. Nam, G. Pascazio, E. Balaras and R. Verzicco Computational prediction of mechanical hemolysis in aortic valved prostheses, *European Journal of Mechanics - B/Fluids*, In Press (2012)
190. *H. Ezzeldin, N. Beraltis*, M. de Tullio, *S. Solares and E. Balaras* Multiscale modeling of mechanical hemolysis in prosthetic devices, *J. Biomechanics*, Submitted, (2012)
191. J. Wu, J. Jaja, N. Beratlis and E. Balaras An Optimized FFT-based Direct Poisson Solver on multicore architectures, *IEEE Transactions on Parallel and Distributed Computing*, Submitted, (2012).
192. Marcos Vanella and Elias Balaras An immersed-boundary method with adaptive mesh refinement optimized for petascale, *11th US National Congress on Computational Mechanics*, Minneapolis, Minnesota, July 25-28, (2011)
193. Dubey, E. Balaras, M. Vanella, Incompressible Navier-Stokes and Other New Capabilities in FLASH-4, Third International Conference 'Turbulent Mixing and Beyond' Trieste, Italy, 21 - 28 August (2011).

CTS110037

194. L Chau, MS Thesis, University of Texas, San Antonio, Understanding the turbulent flow over rough surfaces with idealized and random roughness elements
195. K. Bhaganagar, Direct numerical simulations of density currents over rough surfaces, 64th Annual meeting APS Division of Fluid Dynamics, APS-DFD, Baltimore, MD, 2011
196. L. Chau and K.Bhaganagar, Direct numerical simulations of flow over ridges in the presence of waves and current, 63rd Annual meeting APS Division of Fluid Dynamics, APS-DFD, Long Beach, CA, 2011
197. K. Bhaganagar and V.Juttijudata, Understanding flow over rough-wall using proper orthogonal decomposition, Theoretical and Computational Fluid Dynamics, DOI 10.1007/s00162-011-0242-x, 2011
K.Bhaganagar and R.Leighton, Three level decomposition for the analysis of turbulent flow over rough-walls, Journal of Applied Fluid Mechanics , 6(3), 2013 (Accepted 2011)
198. K. Bhaganagar, C. Veeramachaneni*, K. Das, Significance of plaque morphology in modifying flow characteristics in a diseased coronary artery: Numerical simulation using plaque measurements from intravascular ultrasound imaging, Applied Mathematical Modeling, Submitted 09/2011
199. L. Chau*, K. Bhaganagar, Characterization of turbulent flow over random roughness,*Physics of Fluids*, Submitted 02/2012
200. K. Bhaganagar, Density currents over rough walls, *Physics of Fluids*, Submitted 01/2012
201. L. Chau*, E. Ruiz*, K. Bhaganagar, Coherent structures over idealized and random surface, *Journal of Visualization*, Submitted 02/2012

DEB090011

202. Miller, M. A., Pfeiffer, W. and Schwartz, T.(2011) “The CIPRES science gateway: a community resource for phylogenetic analyses”. In *Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery*, 2011 (Salt Lake City, July 17, 2011). 41:1-8.
203. Arribas, P., Velasco, J., Abellán, P., Sánchez-Fernández, D., Andújar, C., Calosi, P., Millán, A., Ribera, I. and Bilton, D.T. (2012) "Dispersal ability rather than ecological tolerance drives differences in range size between lentic and lotic water beetles (Coleoptera: Hydrophilidae)." Journal of Biogeography *in press* doi: 10.1111/j.1365-2699.2011.02641.x
204. Arribas, P., Abellán, P., Velasco, J., Bilton, D.T., Millán, A. and Sánchez-Fernández, D. (2012) "Evaluating drivers of vulnerability to climate change: a guide for insect conservation strategies" Global Change Biology *in press* 10.1111/j.1365-2486.2012.02691.x
205. Bhullar, K., Waglechner, N., Pawlowski, A., Koteva, K., Banks, E.D., Johnston, M.D., Barton, H.A. and Wright, G.D. (2012) PLoS ONE *in press*
206. Avis, P.G. (2012) "Ectomycorrhizal iconoclasts: The ITS rDNA diversity and nitrophilic tendencies of foetid Russula" Mycologia *in press*
207. Blum, M., Gamper, H.A., Waldner, M., Sierotzki, H., Gisi, U. (2012) "The cellulose synthase 3 (CesA3) gene of oomycetes: structure, phylogeny and influence on sensitivity to carboxylic acid amide (CAA) fungicides." Fungal Biology *in press* http://dx.doi.org/10.1016/j.funbio.2012.02.003
208. Chatrou, L.W., M.D. Pirie, Erkens, R.H.J., Couvreur, T.L.P., Neubig, K.M., Abbott, J.R., Mols, J.B., Maas, J.W., Saunders, R.M.K and Chase, M.W. (2012) "A new subfamilial and tribal classification of the pantropical flowering plant family Annonaceae informed by molecular phylogenetics". Bot. J. Linn. Soc. 169 *in press*
209. Thomas, D.C., J. Wang, Su, Y.C.F., Meinke, S., Chatrou, L.W. and R.M.K. Saunders (2012) "A plastid DNA phylogeny of Dasymaschalon (Annonaceae) and allied genera: evidence for generic non-monophyly and the parallel evolutionary loss of inner petals." Taxon *in press*
210. Hugall, A.F. and Stuart-Fox, D. (2012) "Accelerated speciation in colour polymorphic birds." Nature *in press*
211. Inda, L.A.I., Pimentel, M. and Chase, M.W. (2012) "Phylogenetics of tribe Orchideae (Orchidaceae; Orchidoideae) based on combined DNA matrices: inferences regarding timing of diversification and evolution of pollination syndromes." Annals of Botany *in press*
212. Martínez, J.J., Ferro, L.I., Mollerach, M.I. and Barquez, R.M. (2012) "The phylogenetic relationships of the Andean Swamp Rat genus Neotomys (Rodentia, Cricetidae, Sigmodontinae) based on mitochondrial and nuclear markers." Acta Theriologica *in press*
213. Jabaily, R.S., Shepherd, K.A., Gustafsson, M.H.G., Sage, L.W., Krauss, S.L., Howarth, D.G. and Motley, T.J. (2012) "Systematics of the Austral-Pacific Family Goodeniaceae - Establishing a Taxonomic and Evolutionary Framework." Taxon *in press*
214. Romero, A.I., Robledo, G., LoBuglio, K.F. and Pfister, D.H. (2012) "Rickiella edulis and its phylogenetic relationships within Sarcoscyphaceae." Kurtziana *in press*
215. Ferrer, A., Miller, A.N., Sarmiento, C. and Shearer, C.A. (2012) "Three new genera representing novel lineages of Sordariomycetidae (Sordariomycetes, Ascomycota) from tropical freshwater habitats in Costa Rica." Mycologia *in press*
216. Vis, M.L., Necchi, O. Jr., Chiasson, W.B. and Entwisle, T.J. (2012) "Molecular phylogeny of the genus Kumanoa (Batrachospermales, Rhodophyta)." Journal of Phycology *in press*
217. Wellehan, J.F.X., Pessier, A.P., Archer, L.L., Childress, A.L., Jacobson, E.R. and Tesh, R.B. (2012) "Initial sequence characterization of the rhabdoviruses of squamate reptiles, including a novel rhabdovirus from a caiman lizard (Dracaena guianensis)." Veterinary Microbiology *in press*
218. Magain, N., Goffinet, B. and Sérusiaux, E. (2012) "Further photomorphs in the lichen family Lobariaceae from Reunion (Mascarene archipelago) with notes on the phylogeny of Dendriscocaulon cyanomorphs." The Bryologist *in press*
219. Drew, B.T. and Sytsma, K.J. (2012) "Phylogenetics, Biogeography, and Staminal Evolution in the Tribe Mentheae (Lamiaceae)." American Journal of Botany *in press*
220. Chen, H.X., Sun, S.C., Sundberg , P., Ren, W.C. and Norenburg, J.L. (2012) "A comparative study of nemertean complete mitochondrial genomes, with description of two new ones for Nectonemertes cf. mirabilis and Zygeupolia rubens, elucidates fundamental pattern for the phylum Nemertea." BMC Genomics *in press*
221. Tomlinson et al. (2012) "Biomass partitioning and root morphology of savanna trees across climate types." J. Ecol *in press*
222. Wellehan, J.F.X., Pessier, A.P., Archer, L.L., Childress, A.L., Jacobson, E.R. and Tesh, R.B. (2012) "Initial sequence characterization of the rhabdoviruses of squamate reptiles, including a novel rhabdovirus from a caiman lizard (Dracaena guianensis)." Veterinary Microbiology *in press*

223. Sharma, B., and Kramer, E.M. (2011) The MADS box gene family of the lower eudicot Aquilegia coerulea E. James. Annals of the Missouri Botanical Garden, *in press*
224. Fishbein, M., Chuba, D., Ellison, C., Mason-Gamer, R., and Lynch, S.P. (2011) Phylogenetic relationships of Asclepias (Apocynaceae) estimated from non-coding cpDNA sequences. Systematic Botany *in press.*

DMR060011

225. K. Park, "First-principles study of surface states of Bi2Te3," AIP Conf. Proc. **1416**, 14 (2011).
226. K. Park, "Antiferromagnetic coupling between the single-molecule magnet Mn12 and a ferromagnetic substrate," Phys. Rev. B 83, 064423 (2011).

DMR060061

227. **(Barely) Solid Li(NH3)4: The Electronics of an Expanded Metal.** E. Zurek, X.-D. Wen, and R. Hoffmann. *J. Am. Chem. Soc.* 133 (2011), 3535-3547.
228. **Graphane Sheets and Crystals under Pressure.** X.-D. Wen, L. Hand, V. Labet, T. Yang, R. Hoffmann, N. W. Ashcroft, A. R. Oganov and A. O. Lyakhov. *Proc. Natl. Acad. Sci.* 108 (2011), 6833-6837.
229. Benzene under High Pressure: A Story of Molecular Crystals Transforming to Saturated Networks, with a Possible Intermediate Metallic Phase. X.-D. Wen, R. Hoffmann, and N.W. Ashcroft. *J. Am. Chem. Soc.* 133 (2011), 9023-9035.
230. High Pressure Ices. A. Hermann, N. W. Ashcroft, and R. Hoffmann, *Proc. Natl. Acad. Sci.* 109 (2012), 745-750.
231. A Fresh Look at Dense Hydrogen under Pressure: 1. An Introduction to the Problem, and an Index Probing Equalization of H-H Distances. V. Labet, P. Gonzalez-Morelos, R. Hoffmann, and N. W. Ashcroft. *J. Chem. Phys.*, 136 (2012), 074501.
232. A Fresh Look at Dense Hydrogen under Pressure: 2. Chemical and Physical Models Aiding Our Understanding of Evolving H-H Distances. V. Labet, R. Hoffmann, and N. W. Ashcroft. *J. Chem. Phys*. 136 (2012) 074502
233. A Fresh Look at Dense Hydrogen under Pressure: 3. Two Competing Effects and the Resulting Intramolecular H-H Separation in Solid Hydrogen under Pressure. V. Labet, R. Hoffmann, and N. W. Ashcroft. *J. Chem. Phys.* 136 (2012) 074503.
234. A Fresh Look at Dense Hydrogen under Pressure: 4. Two Structural Models on the Road from Paired to Monatomic Hydrogen, Via a Possible Non-Crystalline Phase. V. Labet, R. Hoffmann, and N. W. Ashcroft. *J. Chem. Phys.* 136 (2012) 074504.
235. WH_n under Pressure. P. Zaleski-Ejgierd, V. Labet, T. A. Strobel, R. Hoffmann, and N. W. Ashcroft. *J. Physics, Condensed Matter* 24, (2012), 155701.

DMR070008

236. Role of Codeposited Impurities During Growth: I. Explaining Distinctive Experimental Morphology on Cu(0 0 1), Ajmi BH. Hamouda, Rajesh Sathiyanarayanan, A. Pimpinelli, and T.L. Einstein, Phys. Rev. B 83, 035423 [8pp] (2011) **DOI:** 10.1103/PhysRevB.83.035423.
237. Role of Codeposited Impurities During Growth: II. Dependence of Morphology on Binding and Barrier Energies, Rajesh Sathiyanarayanan, Ajmi BH. Hamouda, A. Pimpinelli, and T.L. Einstein, Phys. Rev. B 83, 035424 [6pp] (2011) **DOI:** 10.1103/PhysRevB.83.035424.

DMR100007

238. The response to dynamical modulation of the optical lattice for fermions in the Hubbard model, Zhaoxin Xu, Simone Chiesa, Shuxiang Yang, Shi-Quan Su, Daniel E. Sheehy, Juana Moreno, Richard T. Scalettar, and Mark Jarrell, Phys. Rev. A 84, 021607(R) (2011).
239. • Role of the van Hove Singularity in the Quantum Criticality of the Hub- bard Model, K.-S. Chen, S. Pathak, S.-X. Yang, S.-Q. Su, D. Galanakis, K. Mikelsons, M. Jarrell and J. Moreno, Phys. Rev. B 84, 245107 (2011).
240. Optimized Confinement of Fermions in Two Dimensions, J.D. Cone, S. Chiesa, V.R. Rousseau, G.G. Batrouni, and R.T. Scalettar, Phys. Rev. B 85, 075418 (2012).
241. Solving the Parquet Equations for the Hubbard Model beyond Weak Coupling, Ka-Ming Tam, H. Fotso, S.-X. Yang, Tae-Woo Lee, J. Moreno, J. Ramanujam, M. Jarrell, preprint arXiv: 1108-4926.
242. Quantum disordered phase near the Mott transition in the staggered-flux Hubbard model on a square lattice, Chia-Chen Chang and R.T. Scalettar, submitted to Phys. Rev. Lett.
243. Phase diagramof the Bose-Hubbard model on a ring-shape lattice with tunable weak links, K. Hettiarachchilage, V. G. Rousseau, Ka-Ming Tam, M. Jarrell, and J. Moreno, submitted to Phys. Rev. A.

244. Dual Fermion Dynamical Cluster study of the Falicov-Kimball model, S. Pathak, S.-X. Yang, K.-M. Tam, H. Fotso, J. Moreno and Mark Jarrell, in preparation.

DMR100022

245. Competing atomic and molecular mechanism of thermal oxidation, X. Shen, B.R. Tuttle, and S. T. Pantelides (in
preparation)
246. Phosphorous passivation of the SiO^2/4H-SiC interface, Y. K. Sharma, A. C. Ahyi, T. Issacs-Smith, X. Shen, S.
T. Pantelides, X. Zhu, L. C. Feldman, J. Rozen, J. R. Williams, Solid-State Electronics 68, 103 (2012).
247. Oxidation-induced substrate carbon di-interstitials as a major cause of endemically poor mobilities in 4H-SiC/SiO2
structures" X. Shen and S. T. Pantelides Materials Sciences Forum, in press.

DMR100029

248. E. S. Penev, S. Bhowmick, A. Sadrzadeh, and B. I. Yakobson, "Polymorphism of Two-Dimensional Boron", *Nano Lett.* (2012) DOI: 10.1021/nl3004754
249. E. S. Penev, V. I. Artyukhov, and B. I. Yakobson, "Energetics of Capped Carbon Nanotubes", *ACS Nano* (2012), submitted
250. S. Bhowmick, A. K Singh, and B. I. Yakobson, "Quantum Dots and Nanoroads of Graphene Embedded in Hexagonal Boron Nitride", *J. Phys. Chem. C* 115, 9889 (2011)

DMR110035

251. B. P. Abolins, R. E. Zillich, K. B. Whaley, "A Ground State Monte Carlo Approach for Studies of Dipolar Systems with Rotational Degrees of Freedom," Journal of Low Temperature Physics, 165(5-6):249-260, 2011.
252. C. M. Herdman, K. B. Whaley, "Quantum Monte Carlo study of loop condensed phase of the triangular lattice quantum dimer model" *in preparation*, 2012.
253. C. M. Herdman, K. B. Whaley, "The ground state phase diagram of the quantum 3-dimer model on the triangular lattice" *in preparation*, 2012.

DMR110061

254. Adsorption of polyvinylpyrrolidone on Ag surfaces: Insight into a structure-directing agent. W. Al-Saidi, H. Feng, and K. A. Fichthorn, Nano Letters 12, 997 (2012).
255. Comparison of the binding of PVP and PEO to Ag surfaces: Elements of a successful structure-directing agent. Y. Zhou, W. Al-Saidi, and K. A.Fichthorn; (Publication in preparation).
256. Comparison of approaches for including van der Waals interactions into density-functional theory calculations for adsorption on metal surfaces: DFT-D2 vs. DFT+vdWsurf . W. Al-Saidi, H. Feng, and K. A. Fichthorn; (Publication in preparation).

DMR110063

257. *Fermi liquid parameters of half--- filled Landau levels,* Junping Shao, Edward Rezayi, and Eun--- Ah Kim, in preparation (2012).

DMR120048

258. C. Zhang, M.C. Gao, F. Zhang, Y. Yang, "*Modeling of Phase Stability in the Mo-OSystem*", Apr. 2012, to be submitted to J. American Ceramic Society. (supported by DMR100099)
259. M.C. Gao, L. Ouyang, Ö.N. Dogan, "*First-Principles Screening B2 Stabilizers in CuPd Ternary Alloys for Hydrogen Separation Membranes*", to be submitted to J MaterialsScience, Apr. 2012. (supported by DMR100099)
260. M. Hutchinson, M. Widom, "VASP on a GPU: Application to exact-exchange calculations of the stability of elemental boron," Computer Physics Communications, 183(2012) 1422-1426. (supported by Teragrid-PSC)
261. Ö.N. Doğan, X. Song, S. Chen, M.C. Gao, "*Microstructural Study of High-Temperature Cr-Ni-Al-Ti Alloys Supported by First-Principles Calculations*", Metallurgical and Materials Transactions A, Feb 2012, accepted with revision. (supported by DMR100099)

DMS110016

262. J.P. Astheimer, A.J. Hesford, J.C. Tillett and R.C. Waag, "Large-scale acoustic scattering by breast tissue models derived from magnetic resonance images", *IEEE Trans. Biomed. Eng*., In preparation.

EAR000002

263. A. G. Kalinichev; E. Iskrenova-Tchoukova; W.-Y. Ahn; M. M. Clark; R. J. Kirkpatrick (2011) Effects of Ca^{2+} on supramolecular aggregation of natural organic matter in aqueous solutions: A comparison of molecular modeling approaches. Geoderma, 169, 27-32.
264. Y. Marcus; A. G. Kalinichev (2012) Supercritical Water. In: Water: From Hydrogen Bonding to Dynamics and Structure, ed. By R. Ludwig, Wiley VCH, in print.
265. A. G. Kalinichev, B. F. Ngouana Wakou, N. Loganathan (2012) Effects of organics on the adsorption and mobility of metal cations in clay systems: Computational molecular modeling approach. Proceedings of the NEA Clay Club Workshop *Clays under Nano- to Microscopic resolution*, Karlsruhe, Germany, September 6-8, 2011, in print.
266. A. G. Kalinichev (2012) Computational molecular modeling of the multi-scale dynamics of water and ions at cement interfaces. Proceedings of the 1st International Symposium on Cement-Based Materials for Nuclear Wastes - NUWCEM, Avignon, France, October 11-14, 2011, Springer-Verlag, in print.
267. J. W. Singer; A. O. Yazaydin; R. J. Kirkpatrick; G. M. Bowers. Structure and Transformation of Amorphous Calcium Carbonate: A Solid-State ^{43}Ca NMR and Computational Molecular Dynamics Investigation. *Chemistry of Materials* **2012**, submitted.
268. C. P. Morrow; A. Ö. Yazaydin; G. M. Bowers; A. G. Kalinichev; R. J. Kirkpatrick. Structure and Dynamics of Smectite Clay Surface and Interlayer Hydration: Molecular Dynamics Investigation of Na-Hectorite. *Journal of Physical Chemistry C* **2012**, in preparation.
269. M. Saharay; A. Ö. Yazaydin; R. J. Kirkpatrick. Effect of Drying on Amorphous Calcium Carbonate Cluster: A Molecular Dynamics Study. **2012**, in preparation.
270. A. G. Kalinichev. The effects of organic molecules on cation-clay interactions in aqueous environment. Laboratoire PECSA (Physicochimie des Electrolytes, Colloïdes et Sciences Analytiques), UMR 7195, Université Pierre et Marie Curie-Paris 6, France, February 2, 2011.
271. P. Kumar, A. G. Kalinichev, R. J. Kirkpatrick.Structure and H-bonding of aqueous carbonate species from ab initio MD simulation. CECAM Workshop *Microscopic-Scale View of CO2 Sequestration,* Centre Européen de Calcul Atomique et Moléculaire, Lausanne, Switzerland. June 22–24, 2011.
272. A. G. Kalinichev, J. Wang, R. J. Kirkpatrick. Hydrogen bonding and molecular ordering of water at mineral-solution interfaces (keynote talk) Symposium 08k: Water Structure and Hydrogen Bonding on Mineral and Nanoparticle Surfaces, 2011 Goldschmidt Conference, Prague, Czech Republic, August 14–19, 2011.
273. A. G. Kalinichev. Fluid-fluid phase separation under metamorphic conditions: MD simulations of a generalized composition H_2O-CO_2-NaCl. Symposium 19e: Simulation of Geofluids from Melts to Aqueous Solutions, 2011 Goldschmidt Conference, Prague, Czech Republic, August 14–19, 2011.
274. A. G. Kalinichev. Molecular models of natural organic matter and its colloidal aggregation in aqueous solutions: Challenges and opportunities for computer simulations (keynote talk). 32nd International Conference on Solution Chemistry, La Grande Motte, France, August 28– September 02, 2011.
275. R. J. Kirkpatrick. NMR Spectroscopy and Molecular Modeling of Aqueous Fluids in Nano- Confinement and at Solid Surfaces. Colloquim Seminar, Department of Chemistry, Michigan State University, East Lansing, MI, USA, September 22, 2011.
276. C. P. Morrow, A. Ö. Yazaydin, R. J. Kirkpatrick. Structure of Na^+ and H_2O in Na-Hectorite via Molecular Dynamics Simulations. 48th Annual Meeting of the Clay Minerals Society, Lake Tahoe, Nevada, USA, September 25 – 30, 2011.

EAR100013

277. Cooper, C.M., L.-N. Moresi, and A. Lenardic, "Three-dimensional simulations of heat flow partitioning between continents and oceans", (in prep), 2012.
278. Cooper, C.M, L.-N., Moresi and A. Lenardic, "The Effects of Size, Configuration and Distribution of Continents on the Efficiency of Heat Transport, American Geophysical Union, Fall Meeting, 2011.

ECS110004

279. F. Wei and A. E. Yılmaz, "A more efficient and scalable parallelization of the adaptive integral method for 3-D BIOEM analysis part I: Algorithm," *IEEE Trans. Antennas Propagat.,* submitted Sep. 2011, under revision.

280. F. Wei and A. E. Yılmaz, “A more efficient and scalable parallelization of the adaptive integral method for 3-D BIOEM analysis part II: Applications,” *IEEE Trans. Antennas Propagat.,* submitted Sep. 2011, under revision.
281. F. Wei and A. E. Yılmaz, “A hybrid message passing/shared memory parallelization of the adaptive integral method for multi-core clusters,” *Parallel Comp.,* vol. 37, no. 6-7, pp. 279-301, June-July 2011. http://www.sciencedirect.com/science/article/pii/S0167819111000354
282. F. Wei and A. E. Yılmaz, “Parallel AIM-accelerated volume-surface integral equation solution for bioelectromagnetics,” submitted to *ICEAA,* Sep. 2012. (invited)
283. C. S. Geyik, F. Wei, J. W. Massey, and A. E. Yılmaz, “FDTD vs. AIM for bioelectromagnetic analysis,” submitted to *IEEE Antennas Propagat. Soc. Int. Symp.,* July 2012, accepted.
284. F. Wei, J. W. Massey, C. S. Geyik, and A. E. Yılmaz, “Error measures for comparing bioelectromagnetic simulators,” submitted to *IEEE Antennas Propagat. Soc. Int. Symp.,* July 2012, accepted.
285. K. Yang, F. Wei, and A. E. Yılmaz, “Multigrid versus pre-corrected FFT/AIM for biolectromagnetics: When is O(N) better than O(NlogN)?” in *Proc. Computational Electromagnetics Int. Workshop,* pp. 153- 158, Oct. 2011. (invited) http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6047351
286. F. Wei and A. E. Yılmaz, “A 2-D decomposition based parallelization of AIM for 3-D BIOEM problems,” in *Proc. IEEE Antennas Propagat. Soc. Int. Symp.,* pp. 3158-3161, July 2011. http://ieeexplore.ieee.org/xpls/abs_all.jsp?tp=&arnumber=5997203
287. T. Malas, F. Wei, J. Massey, C. S. Geyik, and A. E. Yılmaz, “Application of AIM to high-resolution bioelectromagnetics simulations,” in *Proc. Appl. Comp. Electromagnetics Symp.,* pp. 558-563, Mar. 2011.
288. F. Wei and A. E. Yılmaz, “Parallel truncated multigrid preconditioning of AIM for biolectromagnetics,” submitted to *USNC/URSI Rad. Sci. Meet.*, July 2012, accepted.
289. T. Malas and A. E. Yılmaz, “Near-field preconditioners for large-scale AIM-accelerated bioelectromagnetics simulations,” in *Proc. USNC/URSI Rad. Sci. Meet.*, July 2011.
290. T. Malas, F. Wei, J. Massey, C. S. Geyik, and A. E. Yılmaz, “High-fidelity biolectromagnetics analysis with volume integral equations: AIM acceleration, parallelization, and preconditioning,” in *Proc.Workshop on Advanced Techniques in Computational Electromagnetics*, Glasgow, UK, June 2011(invited).

ECS120005

291. W.- S. Cho, M. Luisier, D. Mohata, S. Datta, D. Pawlik, S. L. Rommel, and G. Klimeck, “Full band atomistic modeling of homo--- junction InGaAs band--- to--- band tunneling diodes including band gap narrowing”, App. Phys. Lett. 100, 063504 (2012).
292. M. Luisier, “Performance Comparison of GaSb, strained--- Si, and InGaAs Double--- Gate Ultrathin--- Body n--- FETs”, IEEE Elec. Dev. Lett. 32, 1686 (2011).
293. S. G. Kim, M. Luisier, T. B. Boykin, and G. Klimeck, “Effects of interface roughness scattering on radio frequency performance of silicon nanowire transistors”, App. Phys. Lett. 99, 232107 (2011).
294. A. Paul, M. Luisier, and G. Klimeck, “Shape and orientation effects on the ballistic phonon thermal properties of ultra--- scaled Si nanowires”, J. of Applied Physics 110, 114309 (2011).
295. A. Paul, M. Luisier, and G. Klimeck, “Influence of cross--- section geometry and wire orientations on the phonon shifts in ultra-scaled Si nanowires”, J. of Applied Physics 110, 094308 (2011).
296. M. Luisier, “Investigation of thermal transport degradation in rough Si nanowires”, J. of Applied Physics 110, 074510 (2011).
297. M. Luisier, “Anharmonic phonon decay in Si nanowires”, to be presented at the next International Workshop on Computational Electronics, IWCE 15, Madison, WI, USA, May 2012.
298. T. B. Boykin, M. Luisier, N. Kharche, X. Jiang, S. K. Nayak, A. Martini, and G. Klimeck, “Multiband Tight--- Binding Model for Strained and Bilayer Graphene from DFT Calculations”, to be presented at the next International Workshop on Computational Electronics (IWCE 15), Madison, WI, USA, May 2012.
299. M. Luisier, “Designing and Understanding Nanoelectronic Devices through Petascale Simulations”, SIAM Conference on Parallel Processing for Scientific Computing, PP12, Savannah, GA, USA, February 2012.
300. M. Luisier, “Nanoelectronic Device Simulations: where do we stand?” (invited), Second Tsukuba Nanoelectronics Symposium, University of Tsukuba, Japan, December 2011.
301. M. Luisier, M. Lundstrom, D. Antoniadis, and J. Bokor, “Ultimate Device Scaling: Intrinsic Performance Comparisons of Carbon--- based, InGaAs, and Si Field--- effect Transistors for 5 nm Gate Length” (conference **highlight),** IEDM 2011, Washington DC, USA, December 2011.
302. S. G. Kim, M. Luisier, T. B. Boykin, and G. Klimeck, “Effects of Interface Roughness Scattering on RF Performance of Nanowire Transistors” (best **student poster award),** ISDRS 2011, College Park, Maryland, USA, December 2011.
303. M. Salmani--- Jelodar, Y. Tan, M. Luisier, T. B. Boykin, and G. Klimeck, “Single Layer MoS2 Band Structure and Transport”, ISDRS 2011, College Park, Maryland, USA, December 2011.

304. M. Luisier, T. B. Boykin, G. Klimeck, and W. Fichtner, “Atomistic nanoelectronic device engineering with sustained performances up to 1.44 PFlop/s”, Proceedings of the 2011 ACM/IEEE Conference on Supercomputing, 1 of the 5 Gordon Bell Prize Finalists, (2011).

MCA03S014

305. Abbo, L., Antonucci, E., Lionello, R., Miki´c, Z., & Riley, P. 2011, ArXiv e-prints
306. Amari, T., Aly, J.-J., Luciani, J.-F., Mikic, Z., & Linker, J. 2011, Ap. J. Lett., 742, L27
307. Antiochos, S. K., Linker, J. A., Lionello, R., Miki´c, Z., Titov, V., & Zurbuchen, T. H. 2011a, Space Sci. Rev., 148
308. Antiochos, S. K., Miki´c, Z., Titov, V. S., Lionello, R., & Linker, J. A. 2011b, Ap. J., 731, 112
309. Brandt, P. C., Mitchell, D. G., Keika, K., Dialynas, K., Ohtani, S., Merkin, V. G., Yang, J., & Jia, X. 2011, in EPSC-DPS Joint Meeting 2011, 970
310. Bruntz, R., Lopez, R. E., Wiltberger, M., & Lyon, J. G. 2012, Journal of Geophysical Research (Space Physics), 117, 3214
311. Bruntz, R. J., Lopez, R. E., Wiltberger, M. J., & Lyon, J. 2011, AGU Fall Meeting Abstracts, A2
312. Chandra, R., Schmieder, B., Mandrini, C. H., D´emoulin, P., Pariat, E., To¨r¨ok, T., Aulanier, G., Uddin, W., & Linton, M. G. 2011a, First Asia-Pacific Solar Physics Meeting ASI Conference Series, 2011, Vol. 2, pp 323-329 Edited by Arnab Rai Choudhuri & Dipankar Banerjee, 2, 323
313. Chandra, R., Schmieder, B., Mandrini, C. H., D´emoulin, P., Pariat, E., To¨r¨ok, T., & Uddin, W. 2011b, Sol. Phys., 269, 83
314. Cnossen, I., Richmond, A. D., Wiltberger, M., Wang, W., & Schmitt, P. 2011a, Journal of Geophysical Research (Space Physics), 116, 12304
315. Cnossen, I., Richmond, A. D., Wiltberger, M. J., Wang, W., & Schmitt, P. J. 2011b, AGU Fall Meeting Abstracts, C4
316. Del Zanna, G., Aulanier, G., Klein, K.-L., & To¨r¨ok, T. 2011, Astron. Astrophys., 526, A137
317. Downs, C., Roussev, I. I., van der Holst, B., Lugaz, N., & Sokolov, I. 2011a, AGU Fall Meeting Abstracts, A1939
318. Downs, C., Roussev, I. I., van der Holst, B., Lugaz, N., Sokolov, I. V., & Gombosi, T. I. 2011b, Ap. J., 728, 2
319. Farid, S., Winebarger, A., Lionello, R., & Titov, V. 2011, in AAS/Solar Physics Division Abstracts #42, 1826

MCA04N012

320. Bode, P., Ostriker, J. P., Cen, R., & Trac, H. 2012, ArXiv e-prints, Calibration of Nonthermal Pressure in Global Dark Matter Simulations of Clusters of Galaxies
321. Cen, R. 2011a, ArXiv e-prints, Coincidences between OVI and OVII Lines: Insights from High Resolution Simulations of the Warm-Hot Intergalactic Medium
322. —. 2011b, ApJ, 741, 99, Environmentally Driven Global Evolution of Galaxies —. 2011c, ArXiv e-prints, Inconsequence of Galaxy Major Mergers in Driving Star Formation
323. at z > 1: Insights from Cosmological Simulations
324. —. 2011d, ArXiv e-prints, Physics of Coevolution of Galaxies and Supermassive Black Holes
325. —. 2012, ApJ, 748, 121, The Nature of Damped Lyα Systems and Their Hosts in the Standard Cold Dark Matter Universe
326. Cen, R. & Chisari, N. E. 2011, ApJ, 731, 11, Star Formation Feedback and Metal-enrichment History of the Intergalactic Medium
327. Kreckel, K., Joung, M. R., & Cen, R. 2011, ApJ, 735, 132, Simulated Void Galaxies in the Standard Cold Dark Matter Model
328. Mesinger, A., Furlanetto, S., & Cen, R. 2011, MNRAS, 411, 955, 21CMFAST: a fast, seminu- merical simulation of the high-redshift 21-cm signal
329. Santos, M. G., Silva, M. B., Pritchard, J. R., Cen, R., & Cooray, A. 2011, A&A, 527, A93, Probing the first galaxies with the Square Kilometer Array
330. Tonnesen, S. & Cen, R. 2011, ArXiv e-prints, Effects on Galaxy Evolution: Pair Interactions versus Environment
331. Zahn, O., Mesinger, A., McQuinn, M., Trac, H., Cen, R., & Hernquist, L. E. 2011, MN- RAS, 414, 727, Comparison of reionization models: radiative transfer simulations and approximate, seminumeric models
332. Zheng, Z., Cen, R., Trac, H., & Miralda-Escud´e, J. 2010a, ApJ, 716, 574, Radiative Transfer Modeling of Lyα Emitters. I. Statistics of Spectra and Luminosity
333. Zheng, Z., Cen, R., Trac, H., & Miralda-Escude, J. 2010b, ArXiv e-prints, Radiative Transfer Modeling of Lyman Alpha Emitters. II. New Effects in Galaxy Clustering

334. Zheng, Z., Cen, R., Weinberg, D., Trac, H., & Miralda-Escudé, J. 2011, ApJ, 739, 62, Extended Lyα Emission around Star-forming Galaxies
335. Fritz, H. M., Borrero, J. C., Synolakis, C. E., Okal, E. A., Weiss, R., Titov, V. V., Jaffe, B. E., Foteinis, S., Lynett, P. J., Chan, I.-C., & Liu, P. L.-F. 2011, Earth Science Reviews, 107, 66
336. Garraffo, C., Cohen, O., Drake, J. J., & Downs, C. 2011, in American Astronomical Society Meeting Abstracts #218, #407.01
337. Gibson, S. E., de Toma, G., Emery, B., Riley, P., Zhao, L., Elsworth, Y., Leamon, R. J., Lei, J., McIntosh, S., Mewaldt, R. A., Thompson, B. J., & Webb, D. 2011, Sol. Phys., 274, 5
338. Guo, Y., Schmieder, B., Démoulin, P., Wiegelmann, T., Aulanier, G., Török, T., & Bommier, V. 2011, in IAU Symposium, Vol. 273, IAU Symposium, 328–332
339. Howard, R. A., Thernisien, A. F., Vourlidas, A., Plunkett, S. P., Korendyke, C. M., Sheeley, N. R., Morrill, J. S., Socker, D. G., Linton, M. G., Liewer, P. C., De Jong, E. M., Velli, M. M., Mikic, Z., Bothmer, V., & Lamy, P. L. 2011, AGU Fall Meeting Abstracts, F6
340. Jian, L., Russell, C. T., Luhmann, J. G., Riley, P., Hoeksema, J. T., Odstrcil, D., & Petrie, G. J. 2011a, AGU Fall Meeting Abstracts, A2036
341. Jian, L. K., Russell, C. T., Luhmann, J. G., MacNeice, P. J., Odstrcil, D., Riley, P., Linker, J. A., Skoug, R. M., & Steinberg, J. T. 2011b, Sol. Phys., 273, 179
342. Kleiber, W., Sain, S. R., & Wiltberger, M. J. 2011, AGU Fall Meeting Abstracts, A2091
343. Kliem, B., Török, T., & Thompson, W. T. 2011, ArXiv e-prints
344. Linker, J. A., Lionello, R., Mikić, Z., Titov, V. S., & Antiochos, S. K. 2011a, Ap. J., 731, 110
345. Linker, J. A., Mikic, Z., Lionello, R., Riley, P., & Titov, V. S. 2011b, AGU Fall Meeting Abstracts, D4
346. Linker, J. A., Riley, P., Lionello, R., Mikic, Z., & Stevens, M. 2011c, in AAS/Solar Physics Division Abstracts #42, 1405
347. Lionello, R., Linker, J. A., Mikić, Z., Riley, P., & Titov, S., V. 2011a, MHD Simulations of the Global Solar Corona and the Solar Wind, ed. Miralles, M. P. & Sánchez Almeida, J., 101
348. Lionello, R., Linker, J. A., Mikic, Z., Riley, P., Titov, V. S., & Torok, T. 2011b, AGU Fall Meeting Abstracts, B2
349. Lionello, R., Velli, M., Linker, J. A., & Mikic, Z. 2011c, in AAS/Solar Physics Division Abstracts #42, 1807
350. Lopez, R. E., Bhattarai, S. K., Bruntz, R. J., Pham, K. H., Wiltberger, M. J., & Lyon, J. 2011, AGU Fall Meeting Abstracts, C7
351. Lotko, W., Brambles, O. J., Zhang, B., & Wiltberger, M. J. 2011, AGU Fall Meeting Ab- stracts, B3
352. Lugaz, N., Downs, C., Shibata, K., Roussev, I. I., Asai, A., & Gombosi, T. I. 2011, Ap. J., 738, 127
353. Luhmann, J. G., Ellenburg, M. A., Riley, P., Lee, C. O., Arge, C. N., Jian, L., Russell, C. T., Simunac, K., Galvin, A. B., & Petrie, G. J. 2011, AGU Fall Meeting Abstracts, C1
354. Lynch, B. J., Reinard, A. A., Mulligan, T., Reeves, K. K., Rakowski, C. E., Allred, J. C., Li, Y., Laming, J. M., MacNeice, P. J., & Linker, J. A. 2011, Ap. J., 740, 112
355. Lyon, J., & Merkin, V. G. 2011, AGU Fall Meeting Abstracts, D7
356. Merkin, V. G., Lyon, J., McGregor, S. L., & Pahud, D. M. 2011a, AGU Fall Meeting Abstracts, A1932
357. Merkin, V. G., Lyon, J. G., McGregor, S. L., & Pahud, D. M. 2011b, Geophys. Res. Lett., 38, 14107
358. Mertens, C. J., Kress, B. T., Wiltberger, M. J., Tobiska, W., & Xu, X. 2011, AGU Fall Meeting Abstracts, B3
359. Mikic, Z., Torok, T., Titov, V. S., Linker, J. A., Lionello, R., & Riley, P. 2011, AGU Fall Meeting Abstracts, B4
360. Mok, Y., Lionello, R., Mikic, Z., & Linker, J. 2011, in AAS/Solar Physics Division Abstracts #42, 1817
361. Murphy, J. J., Elkington, S. R., Schmitt, P. J., Wiltberger, M. J., & Baker, D. N. 2011, AGU Fall Meeting Abstracts, A1985
362. Pahud, D. M., Merkin, V. G., Hughes, W., & McGregor, S. L. 2011, AGU Fall Meeting Abstracts, B7
363. Pembroke, A., Toffoletto, F., Sazykin, S., Wiltberger, M., Lyon, J., Merkin, V., & Schmitt, P. 2012, Journal of Geophysical Research (Space Physics), 117, 2211
364. Pham, K. H., Lopez, R. E., Bruntz, R. J., Bhattarai, S. K., Wiltberger, M. J., & Lyon, J. 2011, AGU Fall Meeting Abstracts, A2078
365. Rabinovich, A. B., Woodworth, P. L., & Titov, V. V. 2011, Geophys. Res. Lett., 38, 16604
366. Rastaetter, L., Kuznetsova, M. M., Shim, J., Hesse, M., Knipp, D. J., Weimer, D. R., Fuller- Rowell, T. J., Ridley, A. J., Raeder, J., Maruyama, N., Kilcommons, L., & Wiltberger, M. J. 2011, AGU Fall Meeting Abstracts, A5
367. Riley, P., Linker, J. A., Gonzalez-Esparza, J. A., Jian, L., Russell, C., & Luhmann, J. 2012, J. Atmos. Terr. Phys., in press
368. Riley, P., Linker, J. A., Lionello, R., & Mikic, Z. 2011a, J. Atmos. Terr. Phys., in press
369. Riley, P., & Lionello, R. 2011, Sol. Phys., 270, 575
370. Riley, P., Lionello, R., Linker, J. A., Mikic, Z., Luhmann, J., & Wijaya, J. 2011b, Sol. Phys., 274, 361

371. Riley, P., & Luhmann, J. G. 2012, Sol. Phys., 277, 355
372. Riley, P., Mikic, Z., Linker, J. A., Torok, T., Lionello, R., & Titov, V. S. 2011c, AGU Fall Meeting Abstracts, C5
373. Riley, P., Torok, T., Mikic, Z., Linker, J. A., Lionello, R., & Titov, V. S. 2011d, AGU Fall Meeting Abstracts, A2
374. Roussev, I. I., Sokolov, I., Downs, C., & Gombosi, T. I. 2011, AGU Fall Meeting Abstracts, A1935
375. Schmieder, B., D´emoulin, P., Pariat, E., To¨r¨ok, T., Molodij, G., Mandrini, C. H., Dasso, S., Chandra, R., Uddin, W., Kumar, P., Manoharan, P. K., Venkatakrishnan, P., & Srivastava, N. 2011a, Advances in Space Research, 47, 2081
376. Schmieder, B., Mandrini, C., Chandra, R., D´emoulin, P., T¨oro˙k, T., Pariat, E., & Uddin, W. 2011b, in IAU Symposium, Vol. 273, IAU Symposium, 164–168
377. Sokolov, I., Roussev, I. I., Gombosi, T. I., & Downs, C. 2011, AGU Fall Meeting Abstracts, A1936
378. Subbotin, D. A., Shprits, Y. Y., Gkioulidou, M., Lyons, L. R., Ni, B., Merkin, V. G., Toffo- letto, F. R., Thorne, R. M., Horne, R. B., & Hudson, M. K. 2011, Journal of Geophysical Research (Space Physics), 116, 8211
379. Thompson, B. J., Gibson, S. E., Schroeder, P. C., Webb, D. F., Arge, C. N., Bisi, M. M., de Toma, G., Emery, B. A., Galvin, A. B., Haber, D. A., Jackson, B. V., Jensen, E. A., Leamon, R. J., Lei, J., Manoharan, P. K., Mays, M. L., McIntosh, P. S., Petrie, G. J. D., Plunkett, S. P., Qian, L., Riley, P., Suess, S. T., Tokumaru, M., Welsch, B. T., & Woods, T. N. 2011, Sol. Phys., 274, 29
380. Thompson, W. T., Kliem, B., & To¨r¨ok, T. 2012, Sol. Phys., 276, 241
381. Titov, V., Mikic, Z., Torok, T., & Linker, J. A. 2011a, in AAS/Solar Physics Division Abstracts #42, 2303
382. Titov, V. S., Miki´c, Z., Linker, J. A., Lionello, R., & Antiochos, S. K. 2011b, Ap. J., 731, 111
383. Titov, V. S., Mikic, Z., Torok, T., Linker, J. A., & Panasenco, O. 2011c, AGU Fall Meeting Abstracts, B1949
384. Torok, T., Archontis, V., & Titov, V. S. 2011, AGU Fall Meeting Abstracts, C8
385. To¨r¨ok, T., Chandra, R., Pariat, E., D´emoulin, P., Schmieder, B., Aulanier, G., Linton, M. G., & Mandrini, C. H. 2011, Ap. J., 728, 65
386. Torok, T., Panasenco, O., Titov, V., Mikic, Z., Reeves, K., Velli, M., Linker, J., & de Toma, G. 2011, in AAS/Solar Physics Division Abstracts #42, 908
387. To¨r¨ok, T., Panasenco, O., Titov, V. S., Miki´c, Z., Reeves, K. K., Velli, M., Linker, J. A., & De Toma, G. 2011, Ap. J. Lett., 739, L63
388. Velli, M., Lionello, R., Linker, J. A., & Miki´c, Z. 2011, Ap. J., 736, 32
389. Wang, W., Lei, J., Burns, A. G., Qian, L., Solomon, S. C., Wiltberger, M., & Xu, J. 2011, Sol. Phys., 274, 457
390. Weigel, B., Wilson, E., & Wiltberger, M. 2012, in Astrophysics Source Code Library, record ascl:1202.001, 2001
391. Wiltberger, M. J., Hudson, M. K., Elkington, S. R., Wang, W., Burns, A. G., Qain, L., Lopez, R. E., & Bruntz, R. J. 2011, AGU Fall Meeting Abstracts, B7
392. Zhang, B., Lotko, W., Brambles, O., Wiltberger, M. J., Wang, W., Schmitt, P. J., & Lyon, J. 2011a, AGU Fall Meeting Abstracts, A5
393. Zhang, B., Lotko, W., Wiltberger, M. J., Brambles, O. J., & Damiano, P. A. 2011b, Journal of Atmospheric and Solar-Terrestrial Physics, 73, 686

MCA07S029

394. A 4-Body Model for Charge Transfer Collisions, A.L. Harris, J. Peacher and D.H. Madison, to be published in *Theory of Heavy Particle Collisions,* ed. by Dž. Belkić.
395. U. Chowdhury, M. Schulz, and D. H. Madison, Phys. Rev. A **83,** 032712 (7pp) (17 March 2011).
396. K. L. Nixon, Andrew J. Murray, Hari Chaluvadi, Chuangang Ning, and D. H. Madison, J. Chem. Phys. **134,** 174304 (8pp) (2 may 2011).
397. A.L. Harris, B. Milum, and D.H. Madison, Phys. Rev. A **84,** 052718 (4pp) (18 November 2011).
398. U Chowdhury, A L Harris, J L Peacher and D H Madison, *J. Phys. B.* **45** 035203 (7pp) (11 January 2012).
399. Kate L Nixon, Andrew James Murray, Hari Chaluvadi, Sadek Amami, Don H Madison and Chuangang Ning, J. Chem. Phys. **136,** 094302 (8pp) (1 March 2012)
400. Uttam Chowdhury, Allison Harris, Jerry Peacher, and Don Madison, 42nd Annual Meeting of the APS Division of Atomic, Molecular and Optical Physics, Atlanta Georgia, 13–17 June 2011 [Bull. Am. Phys. Soc. **56,** No. 5, 76 (2011)].
401. Uttam Chowdhury, Michael Schulz, and Don Madison, 42nd Annual Meeting of the APS Division of Atomic, Molecular and Optical Physics, Atlanta Georgia, 13–17 June 2011 [Bull. Am. Phys. Soc. **56,** No. 5, 167 (2011)].

402. Four body charge transfer process in proton helium collision, U. Chowdhury, A.L. Harris, J. L. Peacher, and D. H. Madison, XXVII International Conference on Photonic, Electronic and Atomic Collisions, Belfast, Northern Ireland, 27 July – 2 August 2011.
403. Differential cross sections for single ionization of H2 by 75keV proton impact, U. Chowdhury, M. Schulz, and D. H. Madison, XXVII International Conference on Photonic, Electronic and Atomic Collisions, Belfast, Northern Ireland, 27 July – 2 August 2011.
404. Fully differential cross section for Single and Double Capture in Proton-Helium Collisions, U Chowdhury, A. L Harris, J. L. Peacher, and D. H. Madison, ICPEAC 2011 Correlation and Coherence Satellite, University College Dublin, Dublin, Ireland, 4-6 August 2011.
405. Four body charge transfer with target excitation in proton helium collision, U. Chowdhury, A.L. Harris, J. L. Peacher, and D. H. Madison, ICPEAC 2011 Correlation and Coherence Satellite, University College Dublin, Dublin, Ireland, 4-6 August 2011.
406. Differential cross sections for single ionization of H2 by 75keV proton impact, U. Chowdhury, M. Schulz, and D. H. Madison, ICPEAC 2011 Correlation and Coherence Satellite, University College Dublin, Dublin, Ireland, 4-6 August 2011.
407. Fully differential cross sections for four-body charge transfer processes, U. Chowdhury, Allison Harris, Jerry Peacher, and Don Madison, 64th Gaseous Electronics Conference, Salt Lake City, Utah, 15-18 November 2011, [Bull. Am. Phys. Soc. 56, No. 15, 27 (2011)]. **This paper was the winner of the Student Award contest.**
408. Indistinguishability in excitation-ionization, B. Milum, A.L. Harris, and D.H. Madison, 64th Gaseous Electronics Conference, Salt Lake City, Utah, 15-18 November 2011, [Bull. Am. Phys. Soc. 56, No. 15, 65 (2011)].

MCA08X001

409. A Monte Carlo Simulation Study of Methane Clathrate Hydrates Confined in Slit Pores, S. N. Chakraborty and L. D. Gelb, J. Phys. Chem. B 116 (2012) 2183-2197.
410. Boiling point determination using adiabatic Gibbs ensemble Monte Carlo simulations: Appli- cation to metals described by embedded-atom potentials, L. D. Gelb and S. N. Chakraborty, J. Chem. Phys. 135 (2011) 224113.
411. Extension of the Steele 10-4-3 potential for adsorption calculations in cylindrical, spherical, and other pore geometries, D. W. Siderius and L. D. Gelb, J. Chem. Phys. 135 (2011) 084703.

MCA08X018

412. Herman, J., Reed, P. M., and Wagener, T., “Time-varying Sensitivity Analysis Reveals Impacts of Watershed Model Choice on the Inference of Dominant Processes.” Water Resources Research, In-Review.
413. Reed, P.M., and Kollat, J.B., "Visual Analytics Clarify the Promise and Challenges for Massively Parallel Many-Objective Optimization: A Groundwater Monitoring Design Example.", Advances in Water Resources, In-Preparation.
414. Reed, P.M., and Kollat, J.B., "Save Now, Pay Later? Multi-Period Many-Objective Groundwater Monitoring Design Given Systematic Model Errors and Uncertainty.", Advances in Water Resources, v35, n1, 55-68, 2012.
415. Kollat, J.B., Reed, P.M., Wagener, T., "When are Multi-Objective Calibration Tradeoffs in Hydrologic Models Meaningful?", Water Resources Research, 48, W03520, doi:10.1029/2011WR011534,2012.
416. Kollat, J.B., Reed, P.M., Wagener, T., "When are Multi-Objective Calibration Tradeoffs in Hydrologic Models Meaningful?", Water Resources Research, 48, W03520, doi:10.1029/2011WR011534,2012.
417. Kollat, J.B., Reed, P.M., and Maxwell, R., “Many-objective Groundwater Monitoring Network Design using Bias-Aware Ensemble Kalman Filtering, Evolutionary Optimization, and Visual Analytics.” Water Resources Research, v47, W02529, doi:10.1029/2010WR009194, 2011.

MCA08X028

418. Gómez, F. A., Minchev, I., O’Shea, B. W., et al., “Signatures of minor mergers in the Milky Way disc I: The SEGUE G-Dwarf sample,” 2012, ApJ, submitted; arXiv:1201.5898
419. B. Smith, E. Hallman, J.M. Shull & B.W. O’Shea, “The Nature of the Warm/Hot Intergalactic Medium. I. Numerical Methods, Convergence, and O VI Absorption,” 2011, ApJ, 731, 6
420. B. O’Shea et al. “Probing the formation of the MIlky Way’s Satellite Population,” 2012, in “Pro- ceedings of the 3rd Subaru International Conference,” Astronomical Society of the Pacific Conference Proceedings, Proceedings number TBD.
421. M. Chatzikos, C.L. Sarazin, & B.W. O’Shea, “The SLAM Project: I. Descriptions of Simulations,” in preparation for submission to the Astrophysical Journal.

422. M. Chatzikos, C.L. Sarazin, & B.W. O'Shea, "The SLAM Project: II. Self-Similarity of Galaxy Cluster Mergers," in preparation for submission to the Astrophysical Journal.
423. B. Crosby, B.W. O'Shea & M.J. Turk, "Large-volume simulations of Population III Star Formation", in preparation for submission to the Astrophysical Journal.
424. B.W. O'Shea, S. Skillman, E.J. Hallman, T. Jeltema, J.O. Burns & J. Jones, "Simulated gamma-ray emission from a population of galaxy clusters," in preparation for submission to the Astrophysical Journal.
425. C. Peruta, B.W. O'Shea, J. Tumlinson, & T. Beers, "The impact of different IMFs on simulations of stellar populations in the Milky Way," in preparation for submission to the Astrophysical Journal.
426. M. Trenti, B.D. Smith & B.W. O'Shea, "Formation of primordial stars in underdense regions at $z < 10$," in preparation for submission to the Astrophysical Journal.
427. J. Wise, B.D. Smith & B.W. O'Shea, "High-resolution simulations of the transition to metal-enriched star formation," in preparatoin for submission to the Astrophysical Journal.

MCA93S001

428. *"Free Energy Considerations for Nucleic Acids with Dangling Ends near a Surface: A Coarse Grained Approach,"* J. Ambia-Garrido, Arnold Vainrub, and B. Montgomery Pettitt, Journal of Physics: Condensed Matter, 23(32), 325101, (2011)
429. *"The Binding Process of a Nonspecific Enzyme with DNA,"* Chuanying Chen B. Montgomery Pettitt, Biophysical Journal, 101(5), 1139, (2011)
430. *"Peptide Conformational Preferences in Osmolyte Solutions: Transfer Free Energies of Decaalanine,"* Hironori Kokubo, Char Y. Hu, and B. Montgomery Pettitt, Journal of the American Chemical Society, 133(6), 1849, (2011)
431. *"Electrostatic Solvation Free Energy of Amino Acid Side Chain Analogs: Implications for the Validity of Electrostatic Linear Response in Water,"* Bin Lin and B. Montgomery Pettitt, Journal of Computational Chemistry, 32(5), 878, (2011)
432. *"Note: On the Universality of Proximal Radial Distribution Functions of Proteins,"* Journal of Chemical Physics, 134(10), 106101, (2011)
433. *"Fast Calculations of Electrostatic Solvation Free Energy from Reconstructed Solvent Density Using Proximal Radial Distribution Functions,"* Bin Lin, Ka-Yiu Wong, and B. Montgomery Pettitt,The Journal of Physical Chemistry Letters, 2(13), 1626, (2011)
434. *"Note: Packing Contributions to DNA Structure in Phages",* Christopher Myers and B. Montgomery Pettitt, manuscript in preparation.

MCA93S028

435. Xabier Agirrezabala, Hstau Liao, Eduard Schreiner, Jie Fu, Rodrigo Ortiz-Meoz, Klaus Schulten, Rachel Green, and Joachim Frank. Structural characterization of mRNA-tRNA translocation intermediates. Proceedings of the National Academy of Sciences, USA, 2012. In press.
436. James Gumbart, Eduard Schreiner, Daniel N. Wilson, Roland Beckmann, and Klaus Schulten. Mechanisms of SecM-mediated stalling in the ribosome. Submitted.
437. David E. Tanner, Kwok-Yan Chan, James Phillips, and Klaus Schulten. Parallel generalized Born implicit solvent calculations with NAMD. Journal of Chemical Theory and Computation, 7:3635- 3642, 2011.
438. Philip M.D. Severin, Xueqing Zou, Hermann E. Gaub, and Klaus Schulten. Cytosine methylation alters DNA mechanical properties. Nucleic Acids Research, 39:8740-8751, 2011.
439. Xueqing Zou, Wen Ma, Ilia Solov'yov, Christophe Chipot, and Klaus Schulten. Recognition of methylated DNA through methyl-CpG binding domain proteins. Nucleic Acids Research, 40:2747- 2758, 2012.
440. Philip M.D. Severin, Xueqing Zou, Klaus Schulten and Hermann E. Gaub. Effects of cytosine hy- droxymethylation on DNA strand separation. In preparation: preliminary version available.
441. Hang Yu, Anton Arkhipov, Ying Yin, and Klaus Schulten. Membrane Sculpting by F-BAR Do- mains studied by Molecular Dynamics Simulations. In preparation: preliminary version available.
442. Eleonore von Castelmur, Johan Strumpfer, Barbara Franke, Julijus Bogomolovas, Sonia Barbieri, Hiroshi Qadota, Petr V. Konarev, Dmitri I. Svergun, Siegfried Labeit, Guy, M. Benian, Klaus Schulten, and Olga Mayans. Identification of an N-terminal inhibitory region as the primary mechanosensory regulator of twitchin kinase. Submitted.
443. Chaitanya Sathe, Xueqing Zou, Jean-Pierre Leburton, and Klaus Schulten. Computational inves- tigation of DNA detection using graphene nanopores. ACS Nano, 5:8842-8851, 2011.
444. Yanxin Liu, Jen Hsin, HyeongJun Kim, Paul R Selvin, and Klaus Schulten. Extension of a three- helix bundle domain of myosin VI and key role of calmodulins. Biophysical Journal, 100:2964-2973, 2011.
445. Eduard Schreiner, Leonardo G. Trabuco, Peter L. Freddolino, and Klaus Schulten. Stereochemical errors and their implications for molecular dynamics simulations. BMC Bioinformatics, 12:190, 2011.

446. Ramya Gamini, Marcos Sotomayor, Christophe Chipot, and Klaus Schulten. Cytoplasmic domain filter function in the mechanosensitive channel of small conductance. Biophysical Journal, 101:80- 89, 2011.
447. Jen Hsin, Loren LaPointe, Alla Kazy, Christophe Chipot, Alessandro Senes, and Klaus Schulten. Oligomerization state of photosynthetic core complexes is correlated with the dimerization affinity of a transmembrane helix. Journal of the American Chemical Society, 133:14071-14081, 2011.
448. Kwok-Yan Chan, James Gumbart, Ryan McGreevy, Jean M. Watermeyer, B. Trevor Sewell, and Klaus Schulten. Symmetry-restrained flexible fitting for symmetric EM maps. Structure, 19:1211- 1218, 2011.
449. Fatemeh Khalili-Araghi, Emad Tajkhorshid, Benoit Roux, and Klaus Schulten. Molecular dynamics investigation of the w current in the Kv1.2 voltage sensor domains. Biophysical Journal, 102:258-267, 2012.
450. Kwok-Yan Chan, Leonardo G. Trabuco, Eduard Schreiner, and Klaus Schulten. Cryo-EM modeling by the molecular dynamics flexible fitting method. Biopolymers, 2012. In press.
451. Whasil Lee, Johan Strumpfer, Vann Bennett, Klaus Schulten, and Piotr E. Marszalek. Mutation of conserved histidines alters the tertiary structure and nanomechanics of consensus ankyrin repeats. Submitted.
452. Martin van Eijk, Michael J. Rynkiewicz, Mitchell R. White, Kevan L. Hartshorn, Xueqing Zou, Klaus Schulten, Dong Luo, Erika C. Crouch, Tanya M. Cafarella, James F. Head, Henk P. Haags- man, and Barbara A. Seaton. Unique sugar-binding site mediates the distinct anti-influenza activity of pig surfactant protein D. Submitted.
453. Yanxin Liu, Johan Strmpfer, Peter L. Freddolino, Martin Gruebele, and Klaus Schulten. Structural Characterization of λ-Repressor folding from All-Atom Molecular Dynamics Simulations. In press.
454. Christopher Harrison and Klaus Schulten. Quantum and classical dynamics of ATP hydrolysis in solvent. Submitted.

MCA94P017

455. Fan, J.. Saunders, M.G., & Voth, G.A. (2012). Coarse-graining reveals key impacts of heterogeneity on actin filaments. *Biophys.J.* in preparation.
456. Knight, C., and G. A. Voth. (2012). Coarse Graining away Electronic Structure: a rigorous route to accurate condensed phase interaction potentials. *Mol. Phys.* in press.
457. Knight, C. & Voth, G. A. (2012). The Curious Case of the Hydrated Proton. *Acc. Chem. Res.* 45, 101-109.
458. Lai C.-L., Jao C. C., Lyman E., McMahon H. T., Langen R & Voth G. A. (2012). Membrane Binding and Self-Association of the Epsin N-Terminal Homology Domain. in preparation.
459. Li, H., Voth, G. A., (2012), The Influenza A M2 Proton Channel: Conformational Ensemble, Water Dynamics, and Proton Permeation, in preparation
460. Li, J. & Voth, G.A. (2012). Molecular dynamics insights into human adenosine A2A receptor activation and deactivation, *PNAS*. in preparation.
461. Lyman, E., Cui, H. S. & Voth, G. A. (2011). Reconstructing protein remodeled membranes in molecular detail from mesoscopic models. *Phys. Chem. Chem. Phys*. 13, 10430-10436.
462. Mim, C., Cui, H., Gawronski-Salerno, J. A., Lyman, E., Voth, G. A. & Unger, V. M. (2012). Structural Basis of Membrane Bending by the N-BAR Protein Endophilin. *Cell* 149, 137-145.
463. Pfaendtner, J., Volkmann, N., Hanein, D., Dalhaimer, P., Pollard, T. D. & Voth, G. A. (2012). Key Structural Features of the Actin Filament Arp2/3 Complex Branch Junction Revealed by Molecular Simulation. *J. Mol. Biol.* 416, 148–161.
464. Saunders, M. G. & Voth, G. A. (2011). Water molecules in the nucleotide binding cleft of actin: effects on subunit conformation and implications for ATP hydrolysis. *J. Mol. Bio.* 413, 279-291.
465. Saunders, M. & Voth, G. A. (2012). Comparison between actin filament models: coarse-graining reveals essential differences. *Structure* 20, 641-653.
466. Saunders, M & Voth, G.A. . (2012). Nucleotide regulation of G-actin’s propeller twist angle. in preparation.
467. Sinitskiy, A., Saunders, M. & Voth, G. A. (2012). Optimal Number of Coarse-Grained Sites in Different Components of Large Biomolecular Complexes. *The Journal of Physical Chemistry B.* Just Accepted Manuscript DOI: 10.1021/jp2108895.
468. Srivastava, A. & Voth, G. A. (2012) A Hybrid Approach for Highly Coarse-grained Lipid Bilayer Models. *The Journal of Physical Chemistry B.* in preparation.
469. Yamashita, T. & Voth, G. A. (2012). Insights into the Mechanism of Proton Transport in Cytochrome c Oxidase. *Journal of the American Chemical Society* 134, 1147-1152.

MCA95C006

470. Berenguer, M., M. Surcel, I. Zawadzki, M. Xue, and F. Kong, 2011: The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part II: Intercomparison between numerical models and with nowcasting. Mon. Wea. Rev., in press.

471. Clark, A. J., John S. Kain, Patrick T. Marsh, James Correia, Jr., Ming Xue, and Fanyou Kong, 2012a: Application of a 3-dimensional object algorithm to convection-allowing forecasts. *Wea. Forecasting*, conditional accepted.
472. Clark, A. J., J. S. Kain, D. J. Stensrud, M. Xue, F. Kong, M. C. Coniglio, K. W. Thomas, Y. Wang, K. Brewster, J. Gao, X. Wang, S. J. Weiss, and J. Du, 2011, Probabilistic Precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. Mon. Wea. Rev., 139, 1410-1418.
473. Clark, A. J., John S. Kain, Patrick T. Marsh, James Correia, Jr., Ming Xue, and Fanyou Kong, 2012a: Application of a 3-dimensional object algorithm to convection-allowing forecasts. Wea. Forecasting, conditional accepted.
474. Clark, A. J., Steven J. Weiss, John S. Kain, Israel L. Jirak, Michael Coniglio, Christopher J. Melick, Christopher Siewert, Ryan A. Sobash, Patrick T. Marsh, Andrew R. Dean, Ming Xue, Fanyou Kong, Kevin W. Thomas, Yunheng Wang, Keith Brewster, Jidong Gao, Xuguang Wang, Jun Du, David R. Novak, Faye E. Barthold, Michael J. Bodner, Jason J. Levit, C. Bruce Entwistle, Tara L. Jensen, James Correia, Jr., 2012b: An Overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. Bull. Amer. Meteor. Soc., 93, 55-74.
475. Dawson, D. T., II, M. Xue, A. Shapiro, and J. A. Milbrandt, 2012: Impact of multi-moment microphysics on real-data simulations of the 3 May 1999 Oklahoma City tornadic supercell and associated tornadoes. Mon. Wea. Rev., To be submitted.
476. Dong, J., M. Xue, and K. K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phy*., **112**, 41-61.
477. Dong, J. and M. Xue, 2012: Coastal WSR-88D Radar Data Assimilation with Ensemble Kalman Filter for Analysis and Forecast of Hurricane Ike (2008). *Quart. J. Roy. Met. Soc.* Conditionally accepted. 2
478. Du, N., M. Xue, K. Zhao, and J. Min, 2012: Impact of assimilating airborne Doppler radar velocity data using the ARPS 3DVAR on the analysis and prediction of hurricane Ike (2008). *J. Geophy. Res.*, Submitted.
479. Gasperoni, N. A., M. Xue, R. D. Palmer, and J. Gao, 2012: Sensitivity of convective initiation prediction to near-surface moisture when assimilating radar refractivity: Impact tests using OSSEs. *J. Atmos. Ocean. Tech*., Submitted.
480. Ge, G., J. Gao, M. Xue, and K. K. Droegemeier, 2012: Diagnostic pressure equation as a weak constraint in a storm-scale three dimensional variational radar data assimilation system. *J. Atmos. Ocean. Tech*., Accepted.
481. Govindankutty, K., and X. Wang, 2012: Assess observation impacts in NCEP hybrid GSI-EnKF data assimilation system for GFS using observation system experiments and ensemble based observation impact metrics. *Mon. Wea. Rev.*, in preparation.
482. Hall, J. D. and M. Xue, 2012: Structure and evolution of vortex Rossby waves and asymmetries in numerically simulated typhoon Morakot (2009) with and without terrain. J. Atmos. Sci., To be submitted.
483. Hall, J. D., M. Xue, L. Ran and Lance M. Leslie, 2012: High-resolution modeling of typhoon Morakot (2009): Vortex Rossby waves and their role in extreme precipitation over Taiwan. *J. Atmos. Sci*., Conditionally accepted.
484. Hu, X.-M., P. Klein, and M. Xue, 2012: A modeling study of nocturnal warming events associated with cold-frontal passages in Oklahoma, To be submitted.
485. Hu, X.-M., P. Klein, M. Xue, F. Zhang, D. C. Doughty, and J. D. Fuentes, 2012: Impact of LLJs-induced vertical mixing on boundary layer O3, *Geophys. Res. Let*., Submitted.
486. Johnson, A., X. Wang, F. Kong, and M. Xue, 2011a: Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous Weather Testbed 2009 Spring Experiment. Part I: Development of object-oriented cluster analysis method for precipitation fields. *Mon. Wea. Rev*., **139**, 3673-3693.
487. Johnson, A., X. Wang, M. Xue, and F. Kong, 2011b: Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous Weather Testbed 2009 Spring Experiment. Part II: season-long ensemble clustering and implication for optimal ensemble design. *Mon. Wea. Rev.*, **139**, 3694-3710.
488. Kain, J. S., M. Xue, M. C. Coniglio, S. J. Weiss, F. Kong, T. L. Jensen, B. G. Brown, J. Gao, K. Brewster, K. W. Thomas, Y. Wang, C. S. Schwartz, and J. J. levit, 2010: Assessing Advances in the Assimilation of Radar Data within a Collaborative Forecasting-Research Environment. *Wea. Forecasting*, **25**, 1510-1521.
489. Kain, J. S., M. C. Coniglio, J. Correia, A. C. Clark, P. T. Marsh, C. Zeigler, V. Lakshmanan, S. D. Miller, S. R. Dembek, S. J. Weiss, F. Kong, and M. Xue, 2012: Convective Initiation: A feasibility study for probabilistic forecasts based on explicit numerical guidance. To be submitted to the *Bull. Amer. Meteor. Soc*.
490. Li, Y., X. Wang and M. Xue, 2012: Assimilation of radar radial velocity data with the WRF ensemble-3DVAR hybrid system for the prediction of hurricane Ike (2008). *Mon. Wea. Rev.*, conditionally accepted. 3
491. Putnam, B. J., M. Xue, Y. Jung, N. A. Snook, and G. Zhang, 2012: The analysis and prediction of microphysical states and polarimetric variables of a mesoscale convective system using double-moment microphysics, multi-network radar data, and the Ensemble Kalman Filter. *Mon. Wea. Rev*., In preparation.

492. Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, M. C. Coniglio, 2009: Next-day convection-allowing WRF model quidance: A second look at 2 vs. 4 km grid spacing. *Mon. Wea. Rev*., 137, 3351-3372.
493. Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, M. C. Coniglio, M. S. Wandishim, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263-280.
494. Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev*., **139**, 3446-3468.
495. Snook, N., M. Xue, and Y. Jung, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev*., In press.
496. Tanamachi, R. L., M. Xue, Y. Jung, K. Brewster, H. B. Bluestein, and M. I. Biggerstaff, 2012: Assimilation via EnKF of high-resolution mobile Doppler radar data collected in a weakly tornadic supercell during Project VORTEX2. *Mon. Wea. Rev*., in preparation.
497. Wang, S., M. Xue, and J. Min, 2012: A four-dimensional asynchronous ensemble square-root filter (4DEnSRF) and tests with simulated radar data. *Quart. J. Roy. Meteor. Soc*., submitted.
498. Wang, X., J. Whitaker, D. Parrish, D. Kleist, R. Treadon, J. Derber, 2012a: GSI-based hybrid ensemble-variational data assimilation system for NCEP Global Forecast System: test with single resolution. *Mon. Wea. Rev.* , in preparation.
499. Wang, X., T. Lei and co-authors, 2012b: Ensemble 4DVAR for NCEP GSI-EnKF hybrid data assimilation system: methods and test results. *Mon. Wea. Rev.* , in preparation.
500. Wang, Y., Y. Jung, T. Supinie, and M. Xue, 2012c: A hybrid MPI/OpenMP parallel algorithm and performance analysis for an ensemble square root filter designed for dense observations. *J. Atmos. Ocean Tech*., In preparation.
501. Xue, M., and J. Dong, 2012: Impact of assimilating best track minimum sea level pressure data together with coastal Doppler radar data on hurricane analysis and prediction at a cloud-resolving resolution, *Quart. J. Roy. Meteor. Soc*., To be submitted.
502. Xue, M., J. Schleif, F. Kong, K. K. Thomas, Y. Wang, K. Zhu, and J. S. Whitaker, 2012: Track and intensity forecasting of Hurricanes: Impact of cloud-resolving resolution and ensemble Kalman filter data assimilation on 2010 Atlantic season forecasts. *Wea. Forecasting*, Being submitted.

MCA99S024

503. Time variability in simulated ultracompact and hypercompact H II region," Galv´an- Madrid, R., Peters, T., Keto, E. R., Mac Low, M.-M., Banerjee, R., Klessen, R. S. 2011, Monthly Not. Roy. Astron. Soc., 416, 1033.
504. Approximations for modelling CO chemistry in giant molecular clouds: a comparison of approaches," Glover, S. C. O., Clark, P. C. 2012 Monthly Not. Roy. Astron. Soc., 421, 116-131.
505. On the Relationship Between Molecular Hydrogen and Carbon Monoxide Abundances in Molecular Clouds," Glover, S. C. O., & Mac Low, M.-M. 2011, Monthly Not. Roy. Astron. Soc., 412, 337.
506. The Abundance of Molecular Hydrogen and its Correlation with Midplane Pressure in Galaxies: Non-Equilibrium, Turbulent, Chemical Models ," 2012, Mac Low, M.-M. & Glover, S. C. O. Astrophys. J., 746, 135
507. Magnetorotational Turbulence Transports Angular Momentum in Strati_ed Disks with Low Magnetic Prandtl Number but Magnetic Reynolds Number above a Critical Value ," Oishi, J. S., & Mac Low, M.-M. 2011, Astrophys. J., 740, 18.
508. The Interplay of Magnetic Fields, Fragmentation and Ionization Feedback in High-Mass Star Formation," Peters, T., Banerjee, R., Klessen, R. S., & Mac Low, M.-M. 2011, Astrophys. J., 729, 72.
509. Modelling CO emission - II. The physical characteristics that determine the X factor in Galactic molecular clouds," Shetty, R., Glover, S. C., Dullemond, C. P., Ostriker, E. C., Harris, A. I., Klessen, R. S. 2011, Monthly Not. Roy. Astron. Soc., 415, 3253-3274.
510. Vertical structure of a supernova-driven turbulent magnetized ISM," Hill, A. S., Joung, M. R., Mac Low, M.-M., Benjamin, R. A., Ha_ner, L. M., Klingenberg, C., Waagan, K. 2012, Astrophys. J., in press (arXiv:1202.0552).
511. Phurbas: An Adaptive, Lagrangian, Meshless, Magnetohydrodynamics Code. I. Algorithm," Maron, J. L., McNally, C. P., Mac Low, M.-M. 2012, Astrophys. J. Suppl., in press (arXiv:1110.0835).
512. Phurbas: An Adaptive, Lagrangian, Meshless, Magnetohydrodynamics Code. II. Implementation and Tests," McNally, C. P., Maron, J. L., Mac Low, M.-M. 2012, Astrophys. J. Suppl., in press (arXiv:1110.0836).

513. The Density Variance{Mach Number Relation in Supersonic Turbulence: I. Isothermal, magnetised gas," Molina, F. Z., Glover, S. C. O., Federrath, C., Klessen, R. S. 2012, Monthly Not. Roy. Astron. Soc., submitted (arXiv:1203.2117).
514. The stellar IMF at very low metallicities," Dopcke, G., Glover, S. C. O., Clark, P., & Klessen, R. S. 2012, Astrophys. J., submitted (arXiv:1203.6842)
515. Star formation in metal-poor gas clouds," Glover, S. C. O., Clark, P. C. 2012, Monthly Not. Roy. Astron. Soc., submitted (arXiv:1203.4251).
516. A Well-Posed Kelvin-Helmholtz Instability Test and Comparison," McNally, C. P., Lyra, W., & Passy, J.-C. 2012, Astrophys. J., submitted.
517. Are molecular outows around high-mass stars driven by ionization feedback?" Peters, T., Klaassen, P. D., Mac Low, M.-M., Klessen, R. S., & Banerjee, R. Astrophys. J., submitted.
518. A resolution study for baroclinic disk instability," Raettig, N., Lyra, W., & Klahr H. 2012, Astrophys. J., submitted
519. How long does it take to form a molecular cloud?" Clark, P. C., Glover, S. C. O., Klessen, R. S., Bonnell, I. A., 2012, Monthly Not. Roy. Astron. Soc., in prep.
520. Rossby wave instability at dead zone boundaries in 3D resistive magnetohydrodynamical global models of protoplanetary disks," Lyra, W. & Mac Low, M.-M. 2012, Astrophys. J., in prep.
521. Sharp eccentric rings in planetless hydrodynamical models of debris disks, " Lyra, W. & Kuchner, M. J. 2012, Astrophys. J., in prep.
522. Planet formation through baroclinic vortices in local models of circumstellar disks," Raettig, N., Lyra, W., Klahr H., & Mac Low, M.-M., 2012, Astrophys. J., in prep.

MCB050014

523. P. W. Thomas, M. Zheng, S. Wu, H. Guo, D. Liu, D. Xu and W. Fast, Biochem. **50**, 10102 (2011).
524. S. Wu, C. Zhang, R. Cao, D. Xu and H. Guo, J. Phys. Chem. B **115**, 10360 (2011).
525. X. Wang, S. Wu, D. Xu, D. Xie and H. Guo, J. Chem. Inf. Comput. Sci. **51**, 1074 (2011).
526. Z. H. Ke, G. K. Smith, Y. K. Zhang and H. Guo, J. Am. Chem. Soc. **133**, 11103 (2011).
527. Z. H. Ke, H. Guo, D. Xie, S. L. Wang and Y. K. Zhang, J. Phys. Chem. B **115**, 3725 (2011).
528. C. M. Johnson, A. F. Monzingo, Z. Ke, D.-W. Yoon, T. W. Linsky, H. Guo, J. D. Robertus and W. Fast, J. Am. Chem. Soc. **133**, 10951 (2011).
529. G. K. Smith, Z. Ke and H. Guo, J. Phys. Chem. B **115**, 13713 (2011).

MCB060037

530. Y. Cai, N. E. Geacintov, and S. Broyde (2012) Nucleotide excision repair efficiencies of bulky carcinogen – DNA adducts are governed by a balance between stabilizing and destabilizing interactions. Biochemistry, 51(7):1486-99.
531. H. Mu, L. Wang, K. Kropachev, L. Zhang, A. Kolbanovskiy, M. Kolbanovskiy, N. E. Geacintov, and S. Broyde. (2012) Molecular dynamics simulations elucidate factors explaining different nucleotide excision repair efficiencies of 2-acetylaminofluorene- and 2- aminofluorene-(C8)guanine lesions in the three NarI sequence contexts. Manuscript in preparation.
532. L. Lior, L. Wang, S. Wang, S. Broyde and Y. Zhang (2012), The nucleotidyl transfer reaction in human DNA polymerase κ utilizes a WMSA mechanism. Manuscript in preparation.
533. J. Yang, S. Broyde and Y. Zhang (2012), Mechanism of the methyl transfer reaction in M.HhaI, Manuscript in preparation.
534. S. Ding, K. Kropachev, Y. Cai, M. Kolbanovskiy, S. A. Durandina, Z. Liu, Shafirovich V, Broyde S, Geacintov NE. (2011) Structural, energetic and dynamic properties of guanine(C8)-thymine(N3) cross-links in DNA provide insights on susceptibility to nucleotide excision repair. *Nucleic Acids Res.* 2011 Dec 1. [Epub ahead of print]
535. Y. Cai, L. Wang, S. Ding, A. Schwaid, N. E. Geacintov and S. Broyde (2011). Intercalative conformations of the 14R (+)- and 14S (−)-trans-anti-DB[a,l]P-N6-dA adducts: molecular modeling, MD simulations and spectroscopic studies. *Chem. Res. Toxicol.* 24(4):522-531
536. D. Reeves, M. Hong, M. Kolbanovskiy, Y. Cai, S. Ding, J. Krzeminsky, A. Kolbanovskiy, J. L. Bolton, S. Amin, K. El-Bayoumy, S. Broyde and N. E. Geacintov (2011). Impact of conformations of PhIP-derived DNA adducts on efficiencies of human nucleotide excision repair. *Nucleic Acids Res.* 39(20):8752-64. *Featured Article.*
537. Y. Liu, D. Reeves, K. Kropachev, Y. Cai, M. Kolbanovskiy, A. Kolbanovskiy, J. L. Bolton, S. Broyde, B. Van Houten, and N. E. Geacintov (2011) Probing for DNA damage with β- hairpins: Similarities in incision efficiencies of bulky DNA adducts by prokaryotic and human nucleotide excision repair systems *in vitro. DNA Repair.* 10(7):684-96

538. H. Zheng, Y. Cai, S. Broyde and Y. Zhang (2010). Energetic studies on base-flipping of dibenzo[a,l]pyrene-damaged DNA duplexes. *Chem. Res. Toxicol.* 23(12):1868–1870
539. Y. Cai, L.Wang, S. Ding, A. Schwaid, N.E. Geacintov and S. Broyde (2010). A bulky DNA lesion derived from a highly potent polycyclic aromatic tumorigen stabilizes nucleosome core particle structure. *Biochemistry,* 49(46):9943-9945 *Highlighted Article.*
540. Y. Cai, K. Kropachev, M. Kolbanovskii, R. Xu, A. Kolbanovskii, S. Amin, D.J. Patel, S. Broyde, and N. E. Geacintov (2010). Distant neighbor base sequence context effects in human nucleotide excision repair of a benzo[a]pyrene-derived lesion. *J. Mol. Biol.* 399(3):397-409.

MCB070059

541. Vashisth, H. and C.F. Abrams, Docking of insulin to a structurally equilibrated insulin receptor ectodomain. Proteins, 2010;78(6):1531-43.
542. Vashisth, H. and C.F. Abrams, All-atom structural models for complexes of insulin-like growth factors IGF1 and IGF2 with their cognate receptor. J Mol Biol, 2010;400:645-658.
543. Nyce, H., S. Stober, C.F. Abrams, and M.M. White, Mapping Spatial Relationships between Residues in the Ligand-Binding Domain of the 5-HT3 Receptor Using a Molecular Ruler. Biophys J, 2010;98(9):1847-55.
544. Abrams, C.F. and E. Vanden-Eijnden, Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. Proc Natl Acad Sci USA, 2010;107(11):4961-6.
545. Emileh, A. and C.F. Abrams, A mechanism by which binding of the broadly neutralizing antibody b12 unfolds the inner domain alpha1 helix in an engineered HIV-1 gp120. Proteins, 2011;79(2):537-46.
546. Gangupomu, V.K. and C.F. Abrams, All-atom models of the membrane-spanning domain of HIV-1 gp41 from metadynamics. Biophys J, 2010;99(10):3438-44.
547. Stober, S. and C.F.Abrams, Enhanced meta-analysis of acetylcholine binding protein structures reveals conformational signatures of agonism in nicotinic receptors, Protein Sci, 2011;21:307-317.
548. Vashisth, H. and C.F. Abrams, DFG-flip in the insulin-receptor kinase is facilitated by a helical intermediate in the activation loop, (accepted, Biophys J, 2012).
549. C. F. Abrams and E. Vanden-Eijnden, "On-the-fly free energy parameterization," (submitted, Chem. Phys. Lett.)
550. Emileh, E. and C.F. Abrams, *All-atom models of unliganded HIV-1 gp120 from temperature-accelerated molecular dynamics,* (to be submitted, Proteins)
551. Baker, M. and C. F. Abrams, *Effects of the Arginine-to-Lysine Mutation at position 694 on the Conformational Ensemble of the HIV-1 gp41 Membrane-Spanning Domain,* (to be submitted, J. Phys. Chem. B)
552. Stober, S., and C. F. Abrams, "Mechanism of the transition from the normal to the amyloidogenic form of β2 microglobulin from on-the-fly string method calculations", (to be submitted, JACS)
553. Emileh, E., Tuzer, F., Chaiken, I., and C. F. Abrams, "All-atom models of peptide-triazole dual-site entry inhibitors in complex with HIV-1 gp120," (to be submitted).

MCB080077

554. Provasi, D., Camacho-Artacho, M., Negri, A., Mobarec, J.C., Filizola, M. "Ligand-Induced Modulation of the Free-Energy Landscape of G Protein-Coupled Receptors Explored by Adaptive Biasing Techniques." *PLOS Computational Biology* (2011) 7(10):e1002193. [PMID: 22022248]
555. Fribourg, M., Moreno, J.L., Holloway, T., Provasi, D., Baki, L., Mahajan, R., Park, G., Adney, S.K., Hatcher, C., Eltit, J.M., Ruta, J.D., Albizu, L., Li, Z., Umali, A., Shim, J., Fabiato, A., MacKerell, A.D. Jr., Brezina, V., Sealfon, S.C., Filizola, M., Gonzalez-Maeso, J., Logothetis, D.E. "Decoding the Signaling of a GPCR Heteromeric Complex Reveals a Unifying Mechanism of Action of Antipsychotic Drugs" *Cell* (2011) 147 (5):1011-1023 [PMID: 22118459]
556. Zhu, J., Choi, W.-S., McCoy, J. G., Negri, A., Zhu, J., Naini, S., Li, J., Shen, M., Huang, W., Bougie, D., Rasmussen, M., Aster, R., Thomas, C. J., Filizola, M., Springer, T. A., Coller, B. S. "Structure-Guided Design of a High-Affinity Platelet Integrin $\alpha_{IIb}\beta_3$ Receptor Antagonist That Disrupts Mg^{2+} Binding to the MIDAS." *Science Translational Medicine* (2012) 4: 125ra32 [PMID: 22422993]

MCB090160

557. Deighan, M., M. Bonomi, and J. Pfaendtner. 2012. Efficient Simulation of Explicitly Solvated Proteins In The Well-Tempered Ensemble. *under review*.
558. Pfaendtner, J., N. Volkmann, D. Hanein, P. Dalhaimer, T. D. Pollard, and G. A. Voth. 2011. Key Structural Features of the Actin Filament Arp2/3 Complex Branch Junction Revealed by Molecular Simulation. J. Mol. Biol. *in press.*

MCB100052

559. V.K. Gupta, C. D. Eggleton, `` Effect of cell and microvillus mechanics on the transmission of applied loads to single bonds in dynamic force spectroscopy." Physical Review E **84**, 011912 (2011).
560. V.K. Gupta, C. D. Eggleton, `` **A** theoretical method to determine unstressed off-rate from multiple bond force spectroscopy." Published online on 25th Feb 2012 at Colloids and Surface B: Biointerfaces.
561. V.K. Gupta, C. D. Eggleton, ``Effect of viscoelasticity on the analysis of single molecule force spectroscopy on live cells." submitted to Biophysical Journal (2012).
562. I. Sraj, A. C. Szatmary, S. A. Desai, D. W.M. Marr, and C. D. Eggleton, "Erythrocyte deformation in high-throughput optical stretchers", accepted for publication in Physical Review E on April 10, 2012.
563. Alex C. Szatmary and Charles D. Eggleton. Elastic capsule deformation in general irrotational linear flows. Submitted to Fluid Dynamic Research 2011.

MCB100057

564. Lee, S., M. Warthaka, C. Yan, T.S. Kaoud, A. Piserchio, R. Ghose, P. Ren*, and K.N. Dalby*, *A Model of a MAPK*Substrate Complex in an Active Conformation: A Computational and Experimental Approach.* **PLoS ONE** (2011) 6, e18594. PMCID: PMC3073974.
565. Piserchio, A., M. Warthaka, A.K. Devkota, T.S. Kaoud, S. Lee, O. Abramczyk, P. Ren, K.N. Dalby, and R. Ghose*, *Solution NMR Insights into Docking Interactions Involving Inactive ERK2.* **Biochemistry** (2011) 50, 3660-3672. PMCID: PMC3103835.
566. Ren, P. *, C. Wu, and J.W. Ponder*, *Polarizable Atomic Multipole-based Molecular Mechanics for Organic Molecules.* **Journal of Chemical Theory and Computation** (2011) 7, 3143-3161 PMCID: PMC3196664.
567. Lee, S., M. Warthaka, C. Yan, T.S. Kaoud, P. Ren*, and K.N. Dalby*, *Examining docking interactions on ERK2 with modular peptide substrates.* **Biochemistry** (2011) 50, 9500-10 (PMCID in progress).
568. Gardner, D.P., P. Ren, S. Ozer, and R.R. Gutell*, *Statistical potentials for hairpin and internal loops improve the accuracy of the predicted RNA structure.* **Journal of Molecular Biology** (2011) 413, 473-483 PMCID: PMC3242814
569. Wu, J., X. Zhen, H. Shen, G. Li, and P. Ren*, *Gay-Berne and electrostatic multipole based coarse-grain potential in implicit solvent.* **The Journal of Chemical Physics** (2011) 135, 155104 (PMCID in progress).
570. Kryscio, D.R., Y. Shi, P. Ren, and N.A. Peppas*, *Molecular docking simulations for macromolecularly imprinted polymers.* **Industrial & engineering chemistry research** (2011) 50, 13877-13884 PMCID: PMC3266373.
571. Shi, Y., C.Z. Zhu, S.F. Martin, and P. Ren*, *Probing the Effect of Conformational Constraint on Phosphorylated Ligand Binding to an SH2 Domain Using Polarizable Force Field Simulations.* **The journal of physical chemistry. B** (2012) 116, 1716-27 (PMCID in progress).
572. Zhang, Jiajing, Wei Yang, Jean-Philip Piquemal, and Pengyu Ren*, *Modeling Structural Coordination and Ligand Binding in Zinc Proteins with a Polarizable Potential.* **Journal of Chemical Theory and Computation** (2012) in press, DOI: 10.1021/ct200812y.
573. 10.Wu, J.C., G. Chattre, and P. Ren*, *Automation of AMOEBA polarizable force field parameterization for small molecules.* **Theoretical Chemistry Accounts** (2012) 131, 1138- 1148 (PMCID in progress)
574. Marjolin, A., C. Gourlaouen, C. Clavaguéra, P. Ren, J. Wu, N. Gresh, J.-P. Dognon, and J.-P. Piquemal, *Toward accurate solvation dynamics of lanthanides and actinides in water using polarizable force fields: from gas-phase energetics to hydration free energies.* **Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)** (2012) 131, 1-14.
575. J. Schnieders, M., T. S. Kaoud, C. Yan, K. N. Dalby*, and P. Ren*, *Computational Insights for the Discovery of Non-ATP Competitive Inhibitors of MAP Kinases.* **Current Pharmaceutical Design (Current Drug Metabolism)** (2012) 18, 1173-1185.

MCB100062

576. Feng Zhao and Jinbo Xu*. A position-specific distance-dependent statistical potential for protein structure and functional study. Submitted to STRUCTURE.
577. Morten Kallerg, Haipeng Wang, Jian Peng, Sheng Wang, Zhiyong Wang, Hui Lu and Jinbo Xu*. RaptorX: a web-server for template-based protein modeling. Submitted to Nature Protocols.
578. Jianzhu Ma, Jian Peng, Sheng Wang and Jinbo Xu*. A Conditional Neural Fields model for protein threading. ISMB 2012. (In Press)
579. Jian Peng and Jinbo Xu*. RaptorX: exploiting structure information for protein alignment by statistical inference. PROTEINS, 2011. (invited to CASP9 special issue).
580. Zhiyong Wang and Jinbo Xu*. A Conditional Random Fields Method for RNA Sequence-Structure Relationship Modeling and Conformation Sampling. Bioinformatics (Proceedings of ISMB/ECCB 2011).
581. Jian Peng and Jinbo Xu*. A multiple-template approach to protein threading. PROTEINS, 2011.

582. Zhiyong Wang, Jian Peng, Feng Zhao and Jinbo Xu*. Protein 8-class secondary structure prediction using conditional neural fields. IEEE BIBM 2010.
583. Raghavendra Hosur, Jinbo Xu, Jadwiga Bienkowska and Bonnie Berger. iWRAP: An interface threading approach with application to prediction of cancer related protein- protein interactions. Journal of Molecular Biology, December 2010.
584. Jian Peng and Jinbo Xu*. Low-homology protein threading. Bioinformatics (Proceedings of ISMB2010), vol 16, issue 12, 2010.
585. Feng Zhao, Jian Peng and Jinbo Xu*. Fragment-free approach to protein folding using conditional neural fields. Bioinformatics (Proceedings of ISMB2010), vol 16, issue 12, 2010.

MCB110016

586. **Xiao, C.,** Fischer, M.G., Bolotaulo, M.D., Ulloa-Rondeau, N., Avila, G.A., Suttle, C.A. (2011) Cryo-EM study of the Cafeteria roenbergensis virus. Talk presented at: ***American Society for Virology 28th Annual Meeting*** (Minneapolis, Minnesota)
587. Enriquez, A.S, Bolotauolo, D.M., Fischer, M.G., Suttle, C.A., **Xiao, C.** (2012)Structural Studies of the capsid of a Giant Marine Virus. Poster presented at: **Rio Grande Branch of the American Society for Microbiology Annual Meeting** (Las Cruces, New Mexico)
588. Xiao, C., Fischer, M.G., Bolotaulo, M.D., Ulloa-Rondeau, N., Avila, G.A., Suttle, C.A. Dissect the Capsomer Arrangement of the Giant Marine Virus *Cafeteria roenbergensis* Virus, in preparation.

MCB110017

589. Natesan S., Wang T., Lukacova V., Bartus V., Khandelwal A., and Balaz S. Rigorous treatment of multispecies multimode ligand-receptor interactions in 3D-QSAR: CoMFA analysis of thyroxine analogs binding to transthyretin. *Journal of Chemical Information and Modeling* **2011,** *51,* 1132-1150
590. Natesan S., Subramaniam R., Bergeron, C., and Balaz S. Binding affinity prediction for ligands and receptors forming tautomers and ionization species: Inhibition of MAPKactivated protein kinase 2. *Journal of Medicinal Chemistry* **2012**, *55*, 2035–2047
591. Natesan S., Wang T., Lukacova V., Bartus V., Khandelwal A., Subramaniam R., and Balaz S.: Cellular quantitative structure-activity relationship (cell-QSAR): Conceptual dissection of receptor binding and intracellular disposition in filaricidal activities of Selwood antimycins. *Journal of Medicinal Chemistry* **DOI:** 10.1021/jm201371y, Publication Date (Web): April 2, 2012

MCB110022

592. The iPlant Collaborative: Cyberinfrastructure for Plant Biology. Goff, SA. et al (2011). Frontiers in Plant Science 2. doi:10.3389/fpls.2011.00034
593. A Scalable Parallel Implementation of the Neighbor Joining Algorithm for Phylogenetic Trees. Robert McLay, Dan Stanzione, Sheldon McKay, Travis Wheeler. 11th Annual IEEE International Workshop on High Performance Computational Biology (submitted).
594. Understanding angiosperm diversification using small and large phylogenetic trees. Smith SA, Beaulieu JM, Stamatakis A, Donoghue MJ. Am J Bot. 2011 Mar;98(3):404--- 14.
595. Quantitative Trait Locus Analysis Using a Partitioned Linear Model on a GPU Cluster. Peter E. Bailey, Tapasya Patki, Gregory M. Striemer, Ali Akoglu, David K. Lowenthal, Peter Bradbury, Matthew Vaughn§, Liya Wang, and Stephen Goff. 11th 11th Annual IEEE International Workshop on High Performance Computational Biology. May 2012.
596. An API To Feed the World. Dooley, R (July 19, 2011). TeraGrid 2011: Extreme Digital Discovery (Salt Lake City, Utah).
597. Building an Environment to Facilitate Discoveries for Plant Sciences. Lenards A., Merchant N. and Stanzione D (2011) GCE11: Proceedings of the 2011 ACM Workshop on Gateway Computing Environments (Seattle, Washington)

In addition to the enumerated publications directly authored by iPlant Collaboative Staff, several of the iPlant Direct Access Program members have indicated that manuscripts resulting from work facilitated by this project are under preparation.

MCB110062

598. Bolia A. Gerek ZN, Keskin O, Dev K and Ozkan SB. “The Binding affinities of proteins interactin with the PDZ domain of PICK1” (2012) Proteins 80: 80:1393-1408.
599. Gerek ZN and Ozkan SB “ (2011) “ Identifying the hidden allostery in PDZ by pertubation residue scanning” Plos Comp Biol (10):e1002154

600. Glembo TJ, Farrell DW, Gerek ZN, Thorpe MF, Ozkan SB. .(2012) Collective dynamics differentiates functional divergence in protein evolutionPLos Comp. Biol (3):e1002428.

MCB120102

601. A. Narayanan, L. L. LeClaire, D. L. Barber, and M. P. Jacobson. "Phosphorylation of the Arp2 Subunit Relieves Auto-inhibitory Interactions for Arp2/3 Complex Activation". *PLoS Computational Biology*, 7 (2011) e1002226. PMC3220268
602. J. D. Sadowsky, M. A. Burlingame, D. W. Wolan, C. L. McClendon, M. P. Jacobson, J. A. Wells. Turning a protein kinase (PDK1) on or off from a single allosteric site via disulfide trapping. *PNAS*. 2011, 108(15), pp 6051-6061. PubMed Central PMC3076885.
603. McClendon, C.L., Hua L., Barreiro G., and Jacobson, M.P. "Comparing Conformational Ensembles Using the Kullback-Leibler Divergence Expansion". Journal of Chemical Theory and Computation, accepted.
604. Datta D., McClendon, C.L., Jacobson, M.P., and Wells, J.A. "Substrate and inhibitor-induced dimerization and cooperativity in caspase-1 but not caspase-3." Submitted to Journal of Biological Chemistry.

OCE030007

605. Boe, J., Hall, A., Colas, F., McWilliams, J.C., Qu, X., Kurian, J., Kapnick, S., 2011. What shapes mesoscale wind anomalies in coastal upwelling zones? Clim. Dyn. 36, 2037–2049. Doi:10.1007/s00382-011-1058-5.
606. Buijsman, M.C., Kanarska, Y., McWilliams, J.C., 2010a. On the generation and evolution of nonlinear internal waves in the South China Sea. J. Geophys. Res. 115. C02012, doi:10.1029/2009JC005275.
607. Buijsman, M.C., McWilliams, J.C., Jackson, C.R., 2010b. East-west asymmetry in nonlinear internal waves from Luzon Strait. J. Geophys. Res. 115. C10057, doi:10.1029/2009JC006004.
608. Buijsman, M.C., Uchiyama, Y., McWilliams, J.C., Hill-Lindsay, C.R., 2012. Modeling semidiurnal internal tide variability in the Southern California Bight. J. Phys. Oceanogr. 42, 62–77.
609. Colas, F., Capet, X., McWilliams, J.C., 2012a. Mesoscale eddy buoyancy flux and eddy-induced circulation in eastern-boundary upwelling systems. J. Phys. Oceanogr. Submitted.
610. Colas, F., McWilliams, J.C., Capet, X., Kurian, J., 2011. Heat balance and eddies in the Peru-Chile current system. Clim. Dyn. Doi:10.1007/s00382-011-1170-6.
611. Colas, F., Molemaker, J., Capet, X., McWilliams, J.C., 2012b. Seasonal and geographical variability of surface-layer submesoscale circulations In presparation.
612. Colas, F., Wang, X., Capet, X., Chao, Y., McWilliams, J.C., 2012c. Untangling the roles of wind, run-off and tides in Prince William Sound. Cont. Shelf Res. Accepted.
613. Deutsch, C., Brix, H., Ito, T., Frenzel, H., Thompson, L., 2011. Climate-forced variability of ocean hypoxia. Science , 336–339.
614. Dong, C., McWilliams, J.C., Hall, A., Hughes, M., 2011. Numerical simulation of a synoptic event in the Southern California Bight. J. Geophys. Res. 116. C05018, doi:10.1029/2010JC006578.
615. Gruber, N., Lachkar, Z., Frenzel, H., Marchesiello, P., Mu¨nnich, M., McWilliams, J.C., Nagai, T., Plattner, G.K., 2011. Eddy-induced reduction of biological production in eastern boundary upwelling systems. Nature Geoscience 4, 787–792. Doi:10.1038/ngeo1273.
616. Gula, J., Molemaker, M.J., McWilliams, J.C., 2012. Shingle eddies and submesoscale filaments in the Gulf Stream In preparation.
617. Kurian, J., Colas, F., Capet, X., McWilliams, J.C., Chelton, D.B., 2011. Eddy properties in the California Current System. J. Geophys. Res. 116, C08027.
618. Lemari´e, F., Debreu, L., Shchepetkin, A.F., McWilliams, J.C., 2012a. On the stability and accuracy of the harmonic and biharmonic adiabatic mixing operators in ocean models. Ocean Modell. Accepted.
619. Lemari´e, F., Jolivet, P., McWilliams, J., 2012b. Mathematical representation of turbulent air/sea fluxes in the atmospheric surface layer: Assessement and application to force a mesoscale ocean model In preparation.
620. Lemari´e, F., Kurian, J., Shchepetkin, A.F., Molemaker, M.J., Colas, F., McWilliams, J.C., 2012c. Are there inescapable issues prohibiting the use of terrain-following coordinates in climate models? Ocean Modell. 42, 57–79. Doi:10.1016/j.ocemod.2011.11.007.
621. Li, Z., Chao, Y., Farrara, J., McWilliams, J., 2012. Impacts of distinct observations during the 2009 Prince William Sound field experiment: A data assimilation study. Cont. Shelf Res. Accepted.
622. Liang, J.H., McWilliams, J., Deutsch, C., Sullivan, P., D'Asaro, E., McNiel, C., Romero, L., Melville, K., 2012a. Significance of gas bubbles in air-sea gas transfer under hurricanes: A comparison between observations and Large Eddy Simulations In preparation.
623. Liang, J.H., McWilliams, J.C., Kurian, J., Wang, P., Colas, F., 2012b. Mesoscale variablility in the northeastern tropical Pacific: Forcing mechanisms and eddy properties. J. Geophys. Res. Submitted.
624. Liang, J.H., McWilliams, J.C., Sullivan, P.P., Baschek, B., 2011. Modeling bubbles and dissolved gases in the ocean. J. Geophys. Res. 116. C03015, doi:10.1029/2010JC006579.

625. Liang, J.H., McWilliams, J.C., Sullivan, P.P., Baschek, B., 2012c. Large Eddy Simulation of the bubbly ocean: New insights on subsurface bubble distribution and bubble-mediated gas transfer. J. Geophys. Res. 117. C04002.
626. Mason, E., Colas, F., Molemaker, J., Shchepetkin, A.F., Troupin, C., McWilliams, J.C., Sangra`, P., 2011. Seasonal variability in the Canary Basin: A numerical study. J. Geophys. Res. 116. C06001, doi:10.1029/2010JC006665.
627. Mason, E., Molemaker, M.J., Shchepetkin, A., Colas, F., McWilliams, J.C., Sangr`a, P., 2010. Procedures for offline grid nesting in regional ocean models. Ocean Modell. 35, 1–15. Doi:10.1016/j.ocemod.2010.05.007.
628. McWilliams, J.C., 2010. A perspective on submesoscale geophysical turbulence, in: IUTAM Symposium on Turbulence in the Atmo- sphere and Oceans. Proceedings of Newton Institute Conference on The Nature of High Reynolds Number Turbulence, Oct. 30 – Dec. 12, 2008. volume 28, pp. 131–141. Doi:10.1007/978-94-007-0360-5 11.
629. McWilliams, J.C., Colas, F., Molemaker, J., 2012a. Filamentary spiral vortices In preparation.
630. McWilliams, J.C., Huckle, E., Liang, J.H., Sullivan, P.P., 2012b. The wavy Ekman layer: Langmuir circulations, breaking waves, and Reynolds stress. J. Phys. Oceanogr. Submitted.
631. Molemaker, M.J., Gula, J., McWilliams, J.C., 2012a. Sustained submesoscale instabilty in a Gulf Stream Ring In preparation.
632. Molemaker, M.J., McWilliams, J.C., 2010. Local balance and cross-scale flux of available potential energy. Journal of Fluid Mechanics 645, 295–314. Doi:10.1017/S0022112009992643.
633. Molemaker, M.J., McWilliams, J.C., Capet, X., 2010. Balanced and unbalanced routes to dissipation in an equilibrated Eady flow. Journal of Fluid Mechanics 654, 33–63.
634. Molemaker, M.J., McWilliams, J.C., Dewar, W.K., 2012b. Submesoscale generation of mesoscale anticyclones in the California Undercurrent In preparation.
635. Nagai, T., Gruber, N., Frenzel, H., McWilliams, J.C., Plattner, G., 2012. Eddy nutrient transport in the California Current System In preparation.
636. Nencioli, F., Dong, C., Dickey, T., Washburn, L., McWilliams, J.C., 2010. A vector geometry based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the Southern California Bight. J. Atmos. Oceanic Technology 27, 564–579.
637. Romero, L., McWilliams, J.C., Siegel, D., Ohlmann, J.C., Y. Uchiyama, Y., 2012. Submesoscale Lagrangian oceanic dispersion in the Southern California Bight In preparation.
638. Shchepetkin, A.F., McWilliams, J.C., 2011. Accurate Boussinesq oceanic modeling with a practical, “stiffened” equation of state. Ocean Modell. 38, 41–70. Doi:10.1016/j.ocemod.2011.01.010.
639. Sullivan, P.P., McWilliams, J.C., 2010. Dynamics of winds and currents coupled to surface waves. Annu. Rev. Fluid Mech. 42, 19–42.
640. Uchiyama, Y., McWilliams, J.C., Idica, E., Stolzenbach, K., 2012. Buoyant sewage outfall dispersal in the Southern California Bight. J. Geophys. Res. In preparation.
641. Uchiyama, Y., McWilliams, J.C., Shchepetkin, A.F., 2010. Wave-current interaction in a three-dimensional circulation model with a vortex force formalism: Application to the surf zone. Ocean Modell. 34, 16–35. Doi:10.1016/j.ocemod.2010.04.002.
642. Wang, X., Chao, Y., Zhang, H., Farrara, J., Li, Z., Jin, X., Park, K., Colas, F., McWilliams, J., Paternostro, C., Shum, C., Yi, Y., Schoch, C., Olsson, P., 2012. Modeling tides in Prince William Sound, Alaska and their influence on circulation. Cont. Shelf Res. Accepted.
643. Watson, J.R., Hays, C.G., Raimondi, P.T., Mitarai, S., Siegel, D.A., Dong, C., McWilliams, J.C., Blanchette, C.A., 2011. Currents connecting communities: The decay of nearshore community similarity with ocean circulation. Ecology 92, 1193–1200.
644. Watson, J.R., Mitarai, S., Siegel, D.A., Caselle, J., Dong, C., McWilliams, J.C., 2010. Realized and potential larval connectivity in the Southern California Bight. Mar. Ecol. Progr. Ser. 401, 31–48.

OCE110011

645. Flanagan, J.D., Radko, T., Shaw, W. and Stanton, T. 2012: The interaction of shear with double diffusion with application to the Weddell Sea. *In preparation*.
646. Radko, T. and Flanagan, J.D. 2012: Constraint on diapycnal transport in a double-diffusive thermocline. *In preparation*.
647. Flanagan, J.D., Hebert, M. and Radko, T., 2012: Numerical simulation, mean-field theory and multiscale analysis of thermohaline intrusions. *In preparation*.

OCE110013

648. McClean, J. L. and E. Yulaeva, Mesoscale air-sea interactions in a fine-resolution Earth System Model. *Ocean Sciences 2012*, Salt Lake City, Utah, 2012.

649. McClean, J.L, M. Maltrud, D. Bader, and L. Evans, Fully-coupled fine-resolution CCSM simulations: A prototype and advancements. *WCRP Open Science Conference,* Denver CO, October 2011.
650. McClean, J. L., E. Yulaeva, and Y.Y. Kim, Mesoscale air-sea interaction in a fine-resolution Earth System Model: Agulhas eddies. *OSTST Science Team Meeting,* San Diego CA, October 2011

PHY060019

651. Measurement of the Positive Muon Lifetime and Determination of the Fermi Constant to Part--- per--- Million Precision, D.M. Webber et al., Phys. Rev. Lett. 106, 041803 (2011)
652. The 10 bar hydrogen time projection chamber of the MuCap experiment, J. Egger et al., Nucl. Instr. Meth. A 628, 199 (2011)

PHY080005

653. "Axial couplings in heavy hadron chiral perturbation theory at the next- to-leading order" W. Detmold, C. -J. D. Lin and S. Meinel. Phys. Rev. D 84, 094502 (2011)
654. "Axial couplings and strong decay widths of heavy hadrons" W. Detmold, C. -J. D. Lin and S. Meinel. arXiv:1109.2480 [hep-lat], to appear in Physical Review Letters.
655. "Excited-state spectroscopy of triply-bottom baryons from lattice QCD" S. Meinel. arXiv:1202.1312 [hep-lat], submitted to Physical Review
656. "Calculation of the heavy-hadron axial couplings g1, g2, and g3 using lattice QCD" W. Detmold, C. -J. D. Lin and S. Meinel. arXiv:1203.3378 [hep-lat], submitted to Physical Review.
657. "Axial couplings of heavy hadrons from domain-wall lattice QCD" W. Detmold, C. -J. D. Lin and S. Meinel. arXiv:1203.3600 [hep-lat] PoS LATTICE 2011, 166 (2011)
658. "Masses and axial couplings of heavy hadrons from lattice QCD" S. Meinel, Poster, The 19th Particles and Nuclei International Conference, MIT, Cam- bridge, USA.
659. "Axial couplings of heavy hadrons" W. Detmold, talk presented at workshop on "Beautiful Mesons and Baryons on the Lattice", Trento, Italy, April 2012.
660. "Excited state spectroscopy of triply botom baryons" S. Meinel, talk presented at workshop on "Beautiful Mesons and Baryons on the Lattice", Trento, Italy, April 2012.

PHY080042

661. T. DeGrand, Y. Shamir, and B. Svetitsky Running coupling and mass anomalous dimension of SU(3) gauge theory with two flavors of symmetric-representation fermions Phys. Rev. D 82, 054503 (2010) [16 pp.]
662. T. DeGrand, Y. Shamir, and B. Svetitsky Mass anomalous dimension in sextet QCD arXiv:1201.0935 [hep-lat], January 2012 [9 pp.] Submitted to Phys. Rev. D
663. T. DeGrand, Y. Shamir, and B. Svetitsky SU(4) lattice gauge theory with decuplet fermions: Schr¨odinger functional analysis arXiv:1202.2675 [hep-lat], February 2012 [15 pp.] Phys. Rev. D, in press

PHY080043

664. A. P. Hickman, D. O. Kashinski, R. F. Malenda, F. Gatti, and D. Talbi, "Calculation of dissociating autoionizing states using the block diagonalization method: Application to N2H", J. Phys. Conf. Ser. 300, 012016 (8 pages) (2011).
665. D. O. Kashinski, D. Talbi, and A. P. Hickman, "*Ab initio* calculations of autoionizing states using block diagonalization: Collinear diabatic states for dissociative recombination of electrons with N2H", Chem. Phys. Lett. **529**, 10–15 (2012).

PHY090083

666. Single photoionization of highly charged atomic ions including the full electromagnetic-field potential M. S. Pindzola, Sh. A. Abdel-Naby, F. Robicheaux, and J. Colgan Phys. Rev. A 85, 032701 (2012) Published March 1, 2012
667. Electron-impact ionization of Al^{2+} Di Wu, S. D. Loch, M. S. Pindzola, and C. P. Balance, Phys. Rev. A 85, 012711 (2012) Published January 25, 2012
668. Electron-impact ionization of Li_2 using a time-dependent close-coupling method M. S. Pindzola, Sh. A. Abdel-Naby, J. A. Ludlow, F. Robicheaux, and J. Colgan Phys. Rev. A 85, 012704 (2012) Published January 17, 2012.
669. Proton-impact excitation of lithium using a time-dependent close-coupling method Teck-Ghee Lee and M. S. Pindzola Phys. Rev. A 84, 052712 (2011) Published November 15, 2011
670. Level-resolved R-matrix calculations for the electron-impact excitation of Ne3+ and Ne6+ J. A. Ludlow, T. G. Lee, C. P. Ballance, S. D. Loch, and M. S. Pindzola Phys. Rev. A 84, 022701 (2011) Published August 1, 2011

671. K-Shell X-Ray Spectroscopy of Atomic Nitrogen M. M. SantAnna, A. S. Schlachter, G. O¨hrwall, W. C. Stolte, D.W. Lindle and B. M. McLaughlin Phys. Rev. Letts. 107, 033001 (2011), Published July 11, 2011
672. Valence-shell photoionization of chlorinelike Ar+ ions A. M. Covington, A. Aguilar, I. R. Covington, G. Hinojosa, C. A. Shirley, R. A. Phaneuf, I. A´ lvarez, C. Cisneros, I. Dominguez-Lopez, M. M. SantAnna, A. S. Schlachter, C. P. Ballance and B. M. McLaughlin Phys. Rev. A 84, 013413 (2011), Published July 18, 2011
673. Electron-impact ionization of Al S D Loch, C P Ballance, D Wu, Sh A Abdel-Naby and M S Pindzola (2012) J. Phys. B: At. Mol. Opt. Phys. 45 065201
674. Photoionization cross section calculations for the halogen-like ions Kr+ and Xe+ B M McLaughlin and C P Ballance (2012) J. Phys. B: At. Mol. Opt. Phys. 45 085701
675. Dielectronic recombination of Au20+: a theoretical description of the res- onances at low electron energies C P Ballance, D C Griffin, S D Loch and N R Badnell (2012) J. Phys. B: At. Mol. Opt. Phys. 45 045001
676. Fine-structure photoionization cross sections of Fe II V Fivet, M A Bautista and C P Ballance (2012) J. Phys. B: At. Mol. Opt. Phys. 45 035201
677. Photoionization Cross-Sections for the trans-iron element Se+ from 18 eV to 31 eV Brendan M McLaughlin and Connor P Ballance J. Phys. B: At. Mol. Opt. Phys. 45 095202 (2012)
678. A large-scale R-matrix calculation for electron-impact excitation of the Ne2+, O-like ion B M McLaughlin, Teck-Ghee Lee, J A Ludlow, E Landi, S D Loch, M S Pindzola and C P Ballance (2011) J. Phys. B: At. Mol. Opt. Phys. 44 175206, Published August 18

PHY090084

679. Evidence of Critical Balance in Kinetic Alfven Wave Turbulence Simulations TenBarge, J. M. and Howes, G. G. Phys. Plasmas, **19**, 055901 (2012).
680. *A Weakened Cascade Model for Turbulence in Astrophysical Plasmas* Howes, G. G., TenBarge, J. M., and Dorland, W. Phys. Plasmas **18,** 102305 (2011).
681. Prediction of the Proton-to-Total Turbulent Heating in the Solar Wind Howes, G. G. Astrophys. J. **738,** 40 (2011).
682. *Gyrokinetic Simulations of Solar Wind Turbulence from Ion to Electron Scales* Howes, G. G., TenBarge, J. M., Dorland, W., Quataert, E., Schekochihin, A. A., Numata, R., and Tatsuno, T. Phys. Rev. Lett. **107,035004** (2011).
683. *Collisionless Dissipation at Electron Scales in the Solar Wind* TenBarge, J. M. and Howes, G. G. Phys. Rev. Lett., submitted (2012).
684. *Interpreting Magnetic Variance Anisotropy Measurements in the Solar Wind* TenBarge, J. M., Podesta, J. J., Klein, K. G., and Howes, G. G. Astrophys. J., submitted (2012).
685. *Measuring the Fundamental Building Block of Astrophysical Turbulence in the Laboratory* Howes, G. G., Drake, D. J., Nielson, K. D., Skiff, F., Kletzing, C. A., and Carter, T. A. Nature, in preparation (2012).

PHY100030

686. B. Urbanc, M. Betnel, L. Cruz, H. Li, E. Fradinger, B.H. Monien, and G. Bitan, “Structural Basis for Aβ1−42 Toxicity Inhibition by Aβ C-terminal Fragments: Discrete Molecular Dynamics Study,” *J. Mol. Biol.* **410,** 316-328 (2011).
687. B. Barz and B. Urbanc, “Dimer Formation Enhances Structural Differences between Amyloid _-Protein (1-40) and (1-42): An Explicit-Solvent Molecular Dynamics Study,” *PLoS ONE*, in press (2012).
688. D. Meral, B. Barz, M. Betnel, and B. Urbanc, “The effect of A_ C-terminal toxicity inhibitors on amyloid-protein folding,” in preparation.
689. I. Eaves, B. Barz, J.-E. Shea, and D. Urbanc, “Oligomer formation of tau R2+ fragment: A computational study,” in preparation.

PHY100033

690. L. Dessart, E. O’Connor, and C. D. Ott, The Arduous Journey to Black-Hole Formation in Potential Gamma-Ray Burst Progenitors, Submitted to the Astrophys. J.; ArXiv:1203.1926 (2012).
691. B. Dasgupta, E. P. O’Connor, and C. D. Ott, Role of collective neutrino flavor oscillations in core-collapse supernova shock revival, Phys. Rev. D. 85, 065008 (2012).
692. F. Löffler, J. Faber, E. Bentivegna, T. Bode, P. Diener, R. Haas, I. Hinder, B. C. Mundim, C. D. Ott, E. Schnetter, et al., The Einstein Toolkit, To appear in Class. Quantum Grav.; arXiv:1111.3344 (2011).
693. C. D. Ott, E. Abdikamalov, E. O’Connor, C. Reisswig, R. Haas, P. Kalmus, S. Drasco, A. Burrows, and E. Schnetter, Correlated Gravitational Wave and Neutrino Signals from General-Relativistic Rapidly Rotating Iron Core Collapse, Submitted to Phys. Rev. D.; arXiv:1204.0512 (2012).

694. E. Abdikamalov, A. Burrows, C. D. Ott, F. Löffler, E. O'Connor, J. C. Dolence, and E. Schnet- ter, A New Monte Carlo Method for Time-Dependent Neutrino Radiation Transport, Submit- ted to the Astrophys. J.; arXiv:1203.2915 (2012).
695. M. Blazewicz, S. R. Brandt, P. Diener, D. M. Koppelman, K. Kurowski, F. Löffler, E. Schnet- ter, and J. Tao, A massive data parallel computational framework on petascale/exascale hybrid computer systems, in PARALLEL COMPUTING 2011 (ParCo2011) (2011), URL http://cacuda.googlecode.com/files/ParCo2011.pdf.
696. C. D. Ott, E. P. O'Connor, and B. Dasgupta, New Aspects and Boundary Conditions of Core- Collapse Supernova Theory, Proceedings of the HAvSE 2011 conference. ArXiv:1111.6282 (2011), 1111.6282.
697. J. Logue, C. D. Ott, I. S. Heng, P. Kalmus, and J. Scargill, Inferring Core-Collapse Super- nova Physics with Gravitational Waves, Submitted to Phys. Rev. D.; arXiv:1202.3256 (2012), 1202.3256.
698. T. Damour, A. Nagar, D. Pollney, and C. Reisswig, Energy Versus Angular Momentum in Black Hole Binaries, Phys. Rev. Lett. 108, 131101 (2012).
699. N. Bishop, D. Pollney, and C. Reisswig, Initial data transients in binary black hole evolutions, Class. Quantum Grav. 28, 155019 (2011).
700. E. O'Connor and C. D. Ott, Black Hole Formation in Failing Core-Collapse Supernovae, Astrophys. J. 730, 70 (2011).

PHY110009

701. S. X. Hu, **"Attosecond** timing the ultrafast charge-transfer process in atomic collisions", Phys. Rev. A **83,** 041401(R) (2011).
702. S. X. Hu, B. Militzer, V. N. Goncharov, S. Skupsky, "First-principles *equation- of-state table of deuterium for inertial confinement fusion applications",* Phys. Rev. B **84,** 224109 (2011).
703. S. X. Hu, B. Militzer, V. N. Goncharov, S. Skupsky, "Strong *Coupling and Degeneracy Effects in Inertial Confinement Fusion Implosions",* Phys. Rev. Lett. **104,** 235003 (2010).
704. S. X. Hu, "Optimizing the FEDVR-TDCC code for exploring the quantum dynamics of two-electron systems in intense laser pulses", Phys. Rev. E **81,** 056705 (2010).
705. M. Chini, B. Zhao, H. Wang, Y. Cheng, S. X. Hu, Z. Chang, "*Sub-cycle AC Stark shift*", Phys. Rev. Lett. (submitted).
706. S. X. Hu, "*Attosecond controlling double-ionization of He*", Phys. Rev. Lett. (to be submitted).
707. J.M. Ngoko Djiokap, S. X. Hu, A. F. Starace, "*Asymmetries in Producing He+(n=2) with an Intense Few-Cycle Attosecond Pulse*", Phys. Rev. Lett. (in preparation).
708. S. X. Hu, L. A. Collins, and B. I. Schneider, "*Threshold ionization of H2 + with an attosecond XUV pulse*", Phys. Rev. A (in preparation).

G XSEDE Citation Report

Content for this appendix is pending the publication and subsequent citation of a research paper on the XSEDE project. XSEDE allocated users will be requested to cite this paper in their publications, thus enabling a more comprehensive analysis of the science impact of the XSEDE project.

H XSEDE EOT Event Details

The following are events conducted during the quarter by one or more XSEDE teams. Collectively, more than 1,500 people have been engaged through 33 workshops, 15 tutorials, 14 presentations, 7 education meetings, and campus visits.

Type	Title	Location	Date(s)	Hours	Number of Participants	Number of Under-represented people	Method	Funding Sources
Workshop	Computational Chemistry for Chemistry Educators	University of Pittsburgh; Pittsburgh, PA	June 11-15		22		Live	XSEDE
Workshop	Computational Thinking from a Parallel Perspective	Southern University; Baton Rouge, LA	June 11-14		21		Live	XSEDE
Workshop	Computational Biology for Biology Educators	Montgomery College; Georgetown, MD	July 16-18		22		Live	XSEDE
Workshop	Intro to Parallel Programming and Cluster Computing	University of Oklahoma; Norman, OK	Aug 5-11		41		Live	XSEDE
Workshop	Kinder, Gentler Supercomputing: Using HPC Resources and Visualization Tools	Kean University; Union, NJ	July 23-25		18		Live	XSEDE
Workshop	Computational physics developers workshop		June 27 – July 2		10		Live	XSEDE
Panel	Campus Bridging in XSEDE, OSG, and Beyond	Internet2 Spring Members' Meeting	24 Apr 2012	1315 - 1430 EDT	32	10	Live	XSEDE
Webinar	"What is campus bridging and why should XSEDE Campus Champions	Michigan City, IN	8 Jun 2012	1300 - 1400 EDT	15	7	Webcast	XSEDE

	Care?”							
Campus Visit	Math Teachers: modeling skills in the sciences	Dept of Defense Dependent Schools (DoDDS), Stuttgart, Germany	April 17		9		Live	
Workshop	System Dynamics and Modeling in Biology	University of Freiburg, Germany	April 18-19		15		Live	
Campus Visit	Math Teachers: modeling skills in the sciences	Dept of Defense Dependent Schools (DoDDS), Wiesbaden, Germany	April 20		7		Live	
Symposium	STEMposium, week of computational modeling and simulation tutoring, workshops and mentoring	DoDDS-Europe, Oberwesel, Germany	April 21-28		107		Live	
Workshop	Computational Science	East Carolina University	May 10		5		Live	
Workshop	Introduction to Computational Thinking	Pamlico Schools, Bayboro NC	May 11		19		Live	
Workshop	Intro to Computational Thinking	Universidad de Sagrado Corazon, San Juan, PR	May 22-25		30		Live	
Workshop	Computational Thinking across the Curriculum	Union College, Schenectady, NY	May 30-31		66		Live	
Workshop	AgentSheets teacher Institute, Agent Modeling in Science	University of Colorado, Boulder	June 5-9		39		Live	
Workshop	Modeling your World in partnership	Harkers Island, NC	June 18-20		25	7	Live	

	with The Bridge Downeast							
Workshop	Modeling your World in partnership with The Bridge Downeast		June 25-29		17		Live	
Workshop	Modeling in Marine Science	West Carteret High School, Morehead City, NC	June 21		47		Live	
Education Meeting		Clark Atlanta			33			
Education Meeting		North Carolina			19			
Education Meeting		College of Southern Nevada			12			
Education Meeting		Great Basin College			7			
Education Meeting		Stockton College			20			
Education Meeting		Norfolk University			40			
Education Meeting		Nashville			67			
Workshop		East Carolina University			6			
Workshop		Richmond Public Schools			18			
Workshop		Virginia Military Institute			66			
Workshop		North Carolina School of Math and Science			16			
Workshop		Monash University, Australia			35			
Workshop		John Monash School of Science and Math			440			

Workshop		NC Academy of Science			35			
Workshop		West Florida			18			
Workshop		Elon University			16			
Workshop		Drake			25			
Workshop		NC Science Teachers			40			
Workshop		Clemson University			3			
Workshop		Arapaho Charter School, NC			9			
Workshop		St. Egbert School, NC			9			
Workshop	XSEDE workshop Nashville, TN	Vanderbilt U.	May 7-8	14	60	25	Live	XSEDE
Workshop	XSEDE workshop Raleigh, NC	North Carolina Central U.	May 14-15	14	36	18	Live	XSEDE
Workshop	NICS/XSEDE Spring workshop	U. of Tennessee	March 19-23	30	40	10	Live	XSEDE, NICS, RDAV
Workshop	NICS/OLCF Fall training	Oak Ridge N.L.	October 18-20, 2011	20	35	10	Live	XSEDE, NICS, RDAV
Presentation	XSEDE Campus Bridging	CASC Fall Meeting	September, 2011	1	45	10	Live	XSEDE
Presentation	Regional CI Support for Bridging and Research Engagement	Internet2 Fall Member’s Meeting	October, 2011	1	50	8	Live	XSEDE
BOF	Campus Bridging GFFS Pilot Program	SC11	November, 2011	1	28	5	Live	XSEDE
Presentation	Campus Bridging	ACTI-CCI/ACTI-CSG	January, 2012	1	25	3	Live	XSEDE
Presentation	Penguin at IU and Cloud Services	CASC Spring Meeting	March, 2012	2	50	10	Live	XSEDE
Presentatio	Campus	Internet2	April, 2012	1	42	8	Live	XSEDE

n	Bridging in XSEDE, OSG, and beyond	Spring Members meeting						
Presentation	What is Campus Bridging and Why Should I Care	XSEDE12	July, 2012	1	35	5	Live	XSEDE
Presentation	Campus Bridging Pilot Projects Panel	XSEDE12	July, 2012	1	35	5	Live	XSEDE
Tutorial	"XSEDE System Administrator Tutorial"	TeraGrid '11 Salt Lake City	July 2011.				Live	
Tutorial	"XSEDE System Administrator Tutorial"	TeraGrid '11 Salt Lake City	July 2011.				Live	
Tutorial	"XSEDE End User Tutorial"	TeraGrid '11 Salt Lake City	July 2011.				Live	
Tutorial	"Globus Online Introduction"	TeraGrid'11 Salt Lake City	July 18 2011				Live	
Presentation	"Using Globus Online for Reliable, Secure File Transfer"	ORNL live webcast	August 4 2011.				Webcast	
Presentation	"Using Globus Online for Reliable, Secure File Transfer"	Network for Earthquake Engineering Simulation (NEES) live webcast	August 5 2011.				Webcast	
Presentation	"Data Movement with Globus Online"	ESnet live webcast	September 8 2011.				Webcast	
Tutorial	"Globus Online: Reliable File Transfer. No IT Required"	IEEE Cluster Conference tutorial	September 26 2011.				Live	
Tutorial	"Globus Online File Transfer for University of California Grid Users"	UCLA Live webcast	September 28 2011.				Webcast	
Tutorial	"Globus Online: Reliable File Transfer. No IT	ALCF Getting Started Workshop	October 5 2011.				Live	

	Required"	tutorial						
Presentation	"Powering Collaboration with Cloud-Hosted Data Movement"	Great Plains Network live webinar	October 7 2011.				Webcast	
Tutorial	"Reliable high-performance file transfer... as a service"	National Center for Supercomputing Applications seminar	October 24 2011.				Live	
Presentation	"Reliable high-performance file transfer made easy"	XSEDE Forum live webinar	December 13 2011.				Webcast	
Tutorial	"Reliable high-performance file transfer... as a service"	Cornell University seminar	January 12 2012.				Live	
Tutorial	"Reliable high-performance file transfer... as a service"	ALCF workshop presentation	January 25 2012.				Live	
Presentation	"Globus Online for New XSEDE Users"	XSEDE webcast	January 27 2012.				Webcast	
Tutorial	"Reliable high-performance file transfer... as a service"	University of Michigan live webcast	January 31 2012.				Webcast	
Tutorial	"Reliable high-performance file transfer... as a service"	NERSC live webcast	February 1 2012.				Webcast	
Tutorial	"Reliable high-performance file transfer... as a service"	European Globus Community Forum tutorial	March 26 2012.				Live	
Tutorial	"Reliable high-performance file transfer... as a service"	University of Pittsburgh live webcast	May 7 2012.				Webcast	
Tutorial	"Reliable Data Movement with Globus Online"	Data Intensive Science and Cloud Computing Tutorial University of	May 31 2012.				Live	

		Chicago						