

Comet Virtual Cluster 'User' Experience

XSEDE ECSS Symposium – December 20, 2016

Trevor Cooper (SDSC), tcooper@sdsc.edu

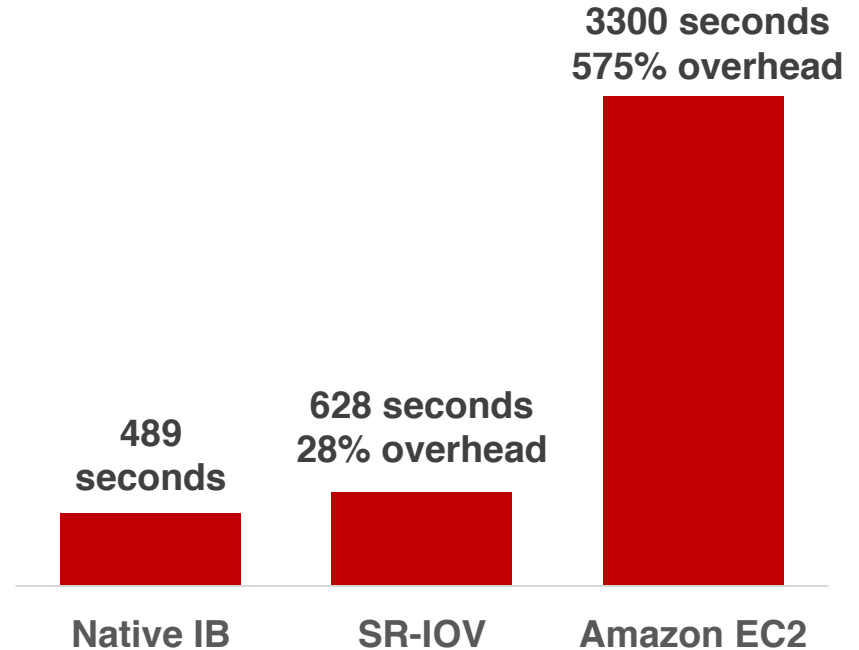


Comet Virtual Clusters
provide a **near bare metal HPC**
performance and management
experience

Quantum ESPRESSO

(Density Functional Theory)

- 48-core (3 node) calculation
- CG matrix inversion
(**irregular communication**)
- 3D FFT matrix transposes
(All-to-all communication)
- SR-IOV incurs modest performance hit...
- ...but **5 times faster** than Amazon



http://glennklockwood.blogspot.com/2013/12/high-performance-virtualization-sr-iov_14.html

Comet System Characteristics

- ~1.8 - 2.0 Pflop/s Peak
- **Compute nodes (1944 nodes)**
 - 2 x Intel® Xeon® E5-2680 v3 @ 2500 MHz
 - 128 GB DDR4 DRAM (4 x 16 GB, 2 NUMA)
 - 2 x Intel® SSD DC S3500 Series (160GB, 2.5in SATA 6Gb/s, 20nm, MLC)
- **GPU nodes (36 nodes)**
 - Add 2 x nVIDIA K80 (12 GB GDDR5) / node to Compute
- **Large-memory nodes (4 nodes)**
 - 4 x Intel® Xeon® E7-8860 v3 @ 2200 MHz
 - 1.5 TB DDR4 DRAM (48 x 32 GB, 4 NUMA)
- **Hybrid fat-tree topology**
 - FDR (56 Gbps) InfiniBand
 - Rack-level (72 nodes) full bisection bandwidth
 - 4:1 oversubscription cross-rack
- **Performance Storage**
 - 7 PB @ 200 GB/s
 - Scratch & Persistent Storage
- **Durable Storage (reliability)**
 - 6 PB @ 100 GB/s
- **Gateway and Frontend hosting nodes (8 nodes)**
 - 2 x Intel® Xeon® E5-2640 v3 @ 2600 MHz
 - 256 GB DDR4 DRAM (16 x 16GB, 2 NUMA)
 - 2 x Intel® SSD DC S3500 Series (240GB, 2.5in SATA 6Gb/s, 20nm, MLC)
- **Image NAS (2 servers)**
 - 2 x Intel® Xeon® E5-2640 v2 @ 2000 MHz
 - 128 GB DDR3 DRAM (8 x 16GB, 2 NUMA)
 - 50 TB (raw) ZFS (zvol and NFS)

Comet Virtual Compute Resources

Comet VC Compute

- 24 vCPU, 120 GiB, 100 GB SSD
- 56 Gpbs MPI & IPoIB
- 10 Gbps Shared NAT
- 576 SUs / day

Amazon C3 or C4 Instances

- c4.4xlarge (16 vCPU, 30 GiB, EBS-only)
- c3.4xlarge (16 vCPU, 30 GiB, 2 x 160 GB SSD)
- 10 Gbps single flow inside
- 5 Gbps outside cluster
- \$ 20 / day

Comet Virtual Clusters

Frontend hosting nodes are **single purpose**

- Comet VC KVM Host
- Shared nodes / modest resources
- Virtual frontends can be up 24/7/365
- Does not debit allocation

Comet compute nodes are **dual purpose**

- Standard HPC batch computing
- Comet VC KVM Host
- One virtual compute node per physical compute node
- Allocated on-demand and charged to allocation automatically

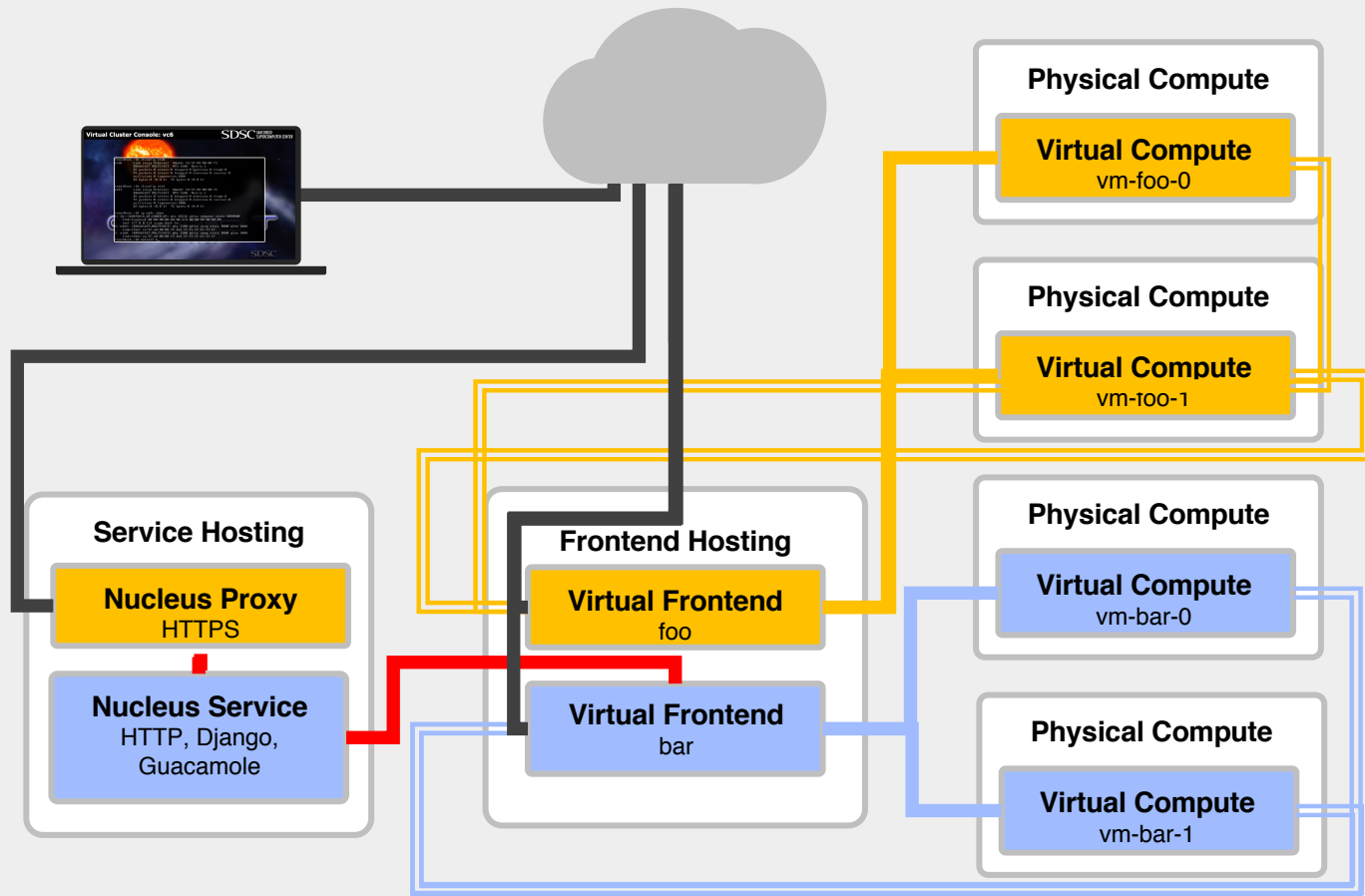
Comet Virtual Clusters

Beowulf-style cluster

- Frontend node + Compute nodes
- Private cluster management network
- Optional additional high-speed network(s)

Comet Virtual Clusters (VC) can run any OS

- Cluster deployment and management is the responsibility of the VC admin



DEMO

Comet VC Team

- **SDSC – Systems**
 - Trevor Cooper
 - Dmitry Mishin
 - Christopher Irving
 - Eva Hocks
 - Haisong Cai
- **IU – User Support**
 - Gregor von Laszewski
 - Fugang Wang

help@xsede.org
(reference Comet Virtual Clusters)