Toward Flood Inundation Mapping at the Continental Scale: the Cyberinfrastructure Approach

Yan Liu

CyberGIS Center
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign

XSEDE ECSS Symposium Webinar
January 17, 2017

Outline

- Background
- Computational challenges and strategies
- High-performance hydrological information analysis software and services - TauDEM
- National coverage of Height Above Nearest Drainage (HAND)
- National inundation mapping
- Conclusion and discussion

- NFIE: National Flood Interoperability Experiment. PI: David Maidment, UT Austin
    - Part 1 was accomplished by Si Liu and John Cazes@TACC
- TauDEM: Terrain analysis using digital elevation models. PI: David Tarboton, Utah State Univ.

Team

- UT Austin
  - David Maidment – methodology development; PI
  - Xing Zheng – HAND methodology; hydro model coupling
  - Justine Yu – Hydro model integration
- Utah State University
  - David Tarboton - methodology development; TauDEM development
  - Nazmus Sazib – TauDEM integration test; HAND methodology
  - Pabitra Dash – HydroShare web
  - Tony Castronova - JupyterHub
- University of Illinois at Urbana-Champaign (UIUC)
  - Yan Liu – coordination; computation; workflow integration
  - Ahmet Yildirim (left) – Inlet identification; block-wise data decomposition in TauDEM; future d8 and dinf algorithm integration
  - Kornelijus Survila – breakthrough of the new flow direction algorithm in TauDEM
  - Nathan Casler, Xingchen Hong, Kiumars Soltani, Dandong Yin, Hao Hu – scalable data storage, query, visualization, JupyterHub environment
- USGS
  - Larry Stanislawski – NHD and methodology consulting
  - NHDPlus team
  - 3DEP national elevation program
- ESRI
  - Steve Kopp, Dean Djokic, Daniel Siegel – TX inundation mapping; commercial solutions
- RENCI
  - Ray Idaszak – iRODS sharing with HydroShare
  - Yi Hong – HydroShare web
- NCAR
  - Channel geometry calculation
- NOAA NWS
  - Nation Water Center (NWC) - summer institute

Real-Time Flood Inundation Mapping

Existing

130 Stream Gage Reaches

Proposed

All of Medium Resolution NHD

	Existing	Proposed	Ratio
Number of Mapped Reaches	130	2,691,344	
Total Mapped Length (Km)	2256	5,192,824	2302
Average Reach Length (Km)	17.4	1.93	0.11

A tremendous application of USGS elevation and hydrography data!

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

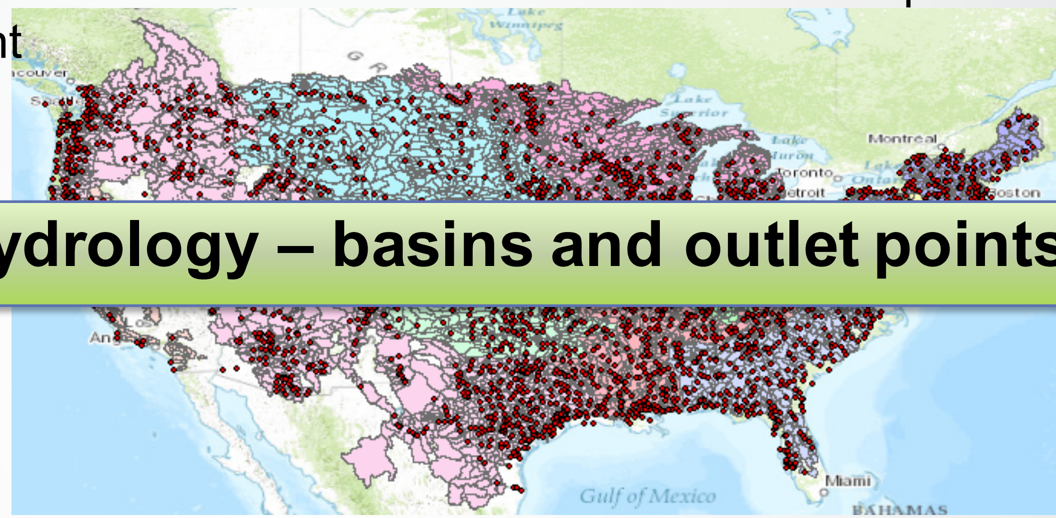
Flow Continuum Model – a national stream network, atmosphere to oceans, coast to coast

Blanco River at Wimberley
Two basins and one forecast point

becomes ⇩

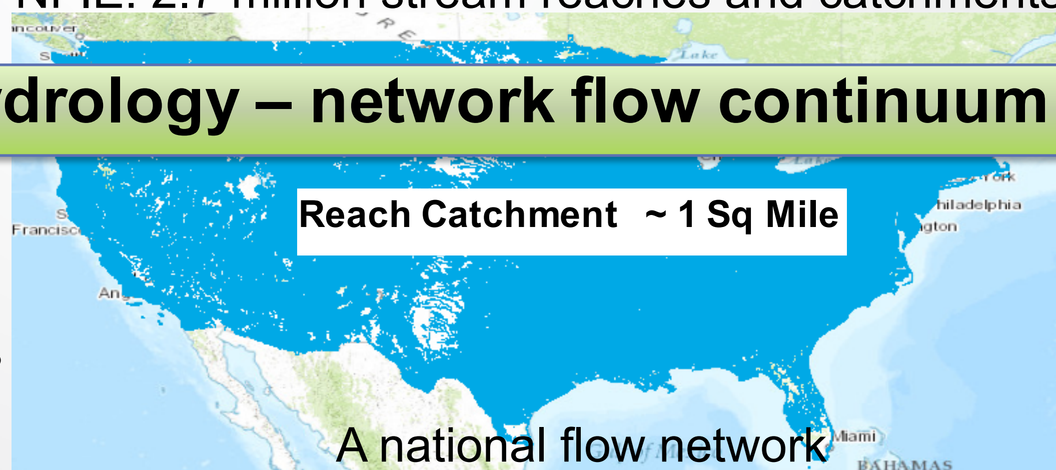
130 Catchments and Flowlines uniquely labelled

Current: 6600 basins and 3600 forecast points



Watershed Hydrology – basins and outlet points

NFIE: 2.7 million stream reaches and catchments

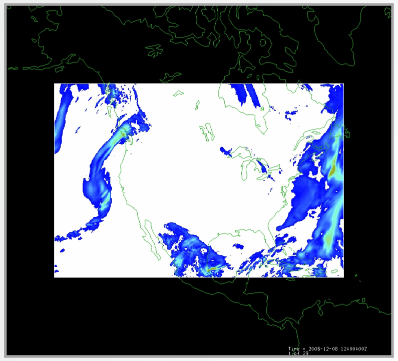


Continental Hydrology – network flow continuum

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

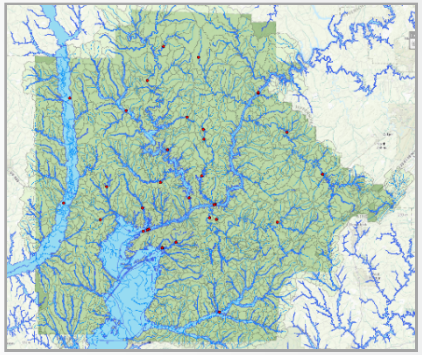
National Water Model

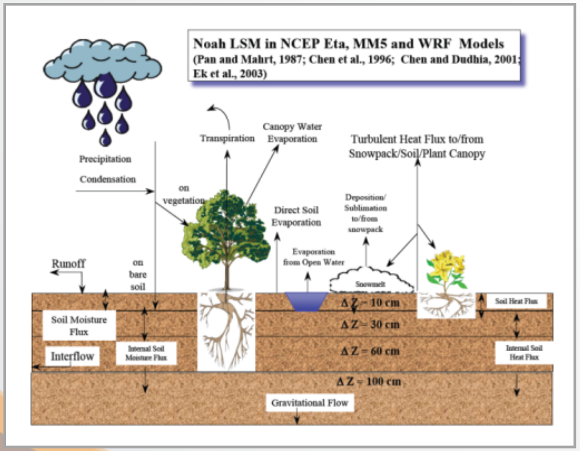
Meteorology



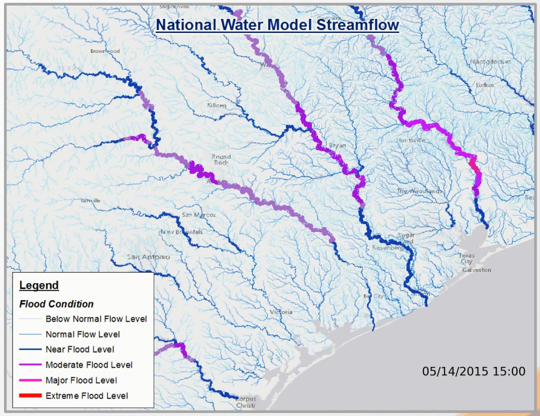
2.7 million catchments

Mapping and Impacts





Hydrology

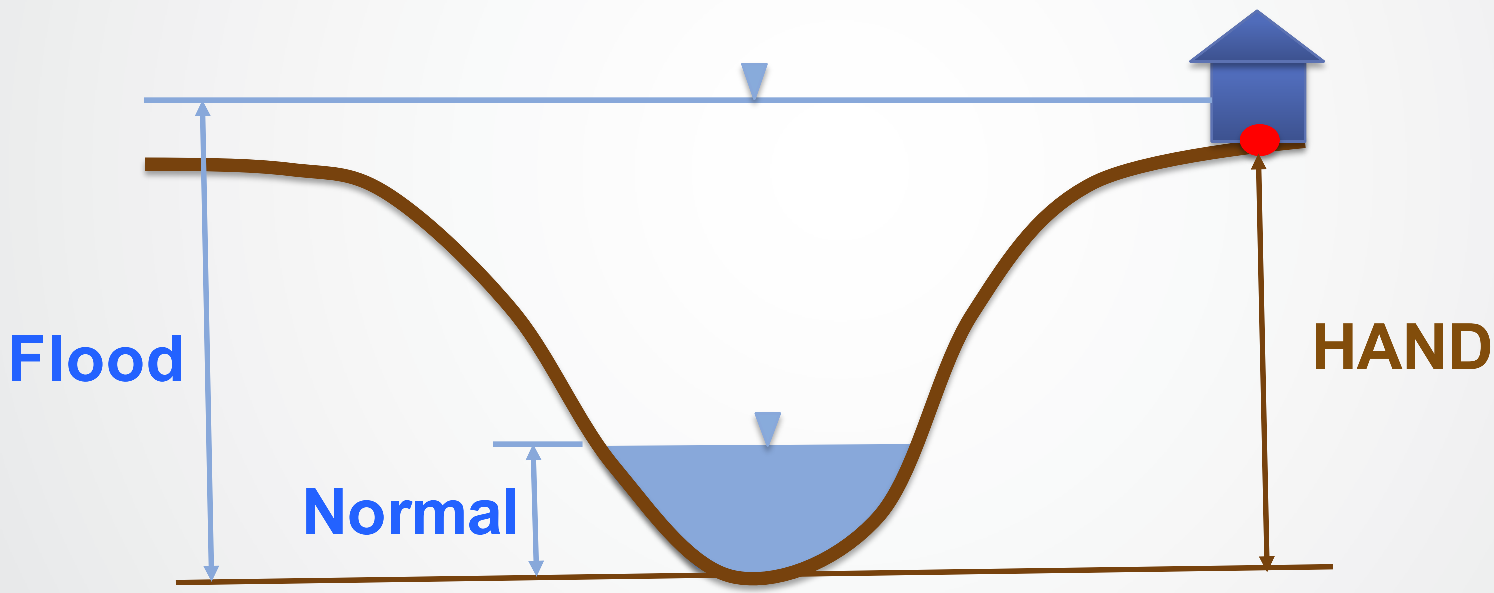


Hydraulics

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

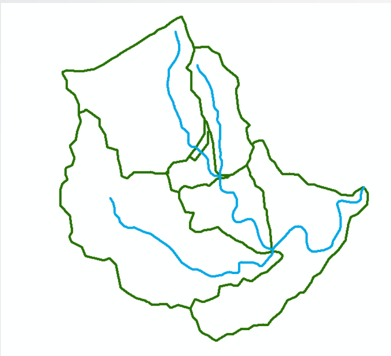
Method for Determining Flood Risk: Height Above Nearest Drainage (HAND)

Flooding occurs when Water Depth is greater than HAND



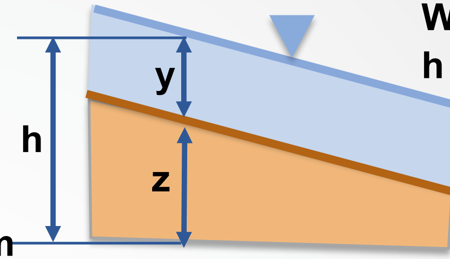
Source: Maidment, NOAA NWS leadership team briefing. 01/09/2017

Flood Inundation Mapping – NHDPlus-HAND Method

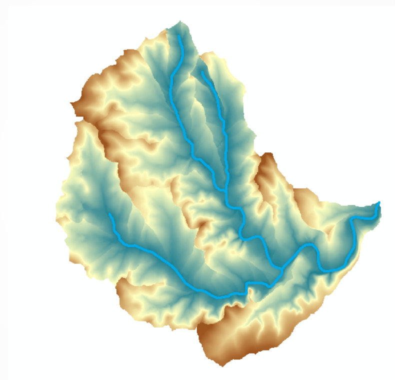


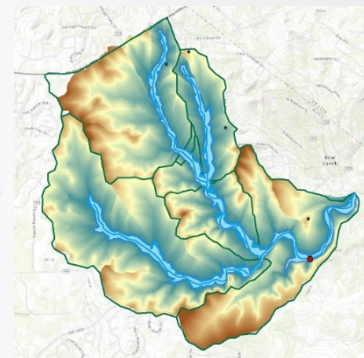
NHDPlus

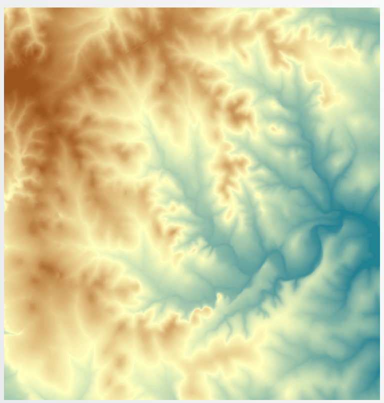
Catchments and Flowlines



Water surface elevation
h = z + y







Digital Elevation Model

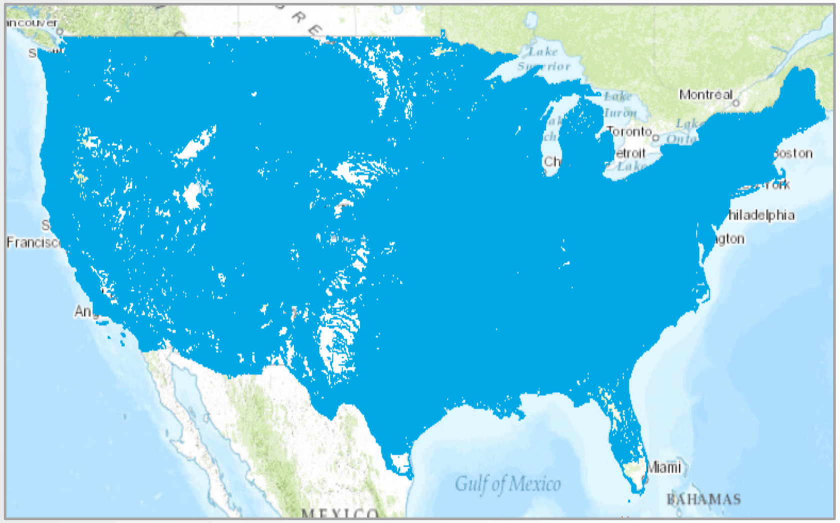
Height Above Nearest Drainage (HAND)
(relative elevation of land surface cell above cell in NHDPlus stream to which it flows)

Inundation map
Height above drainage < 15 ft)

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

Experiment for 2016:

Combine hydrography and elevation to define river channel geometry and flood inundation extent for 5 million km of stream reaches over continental US



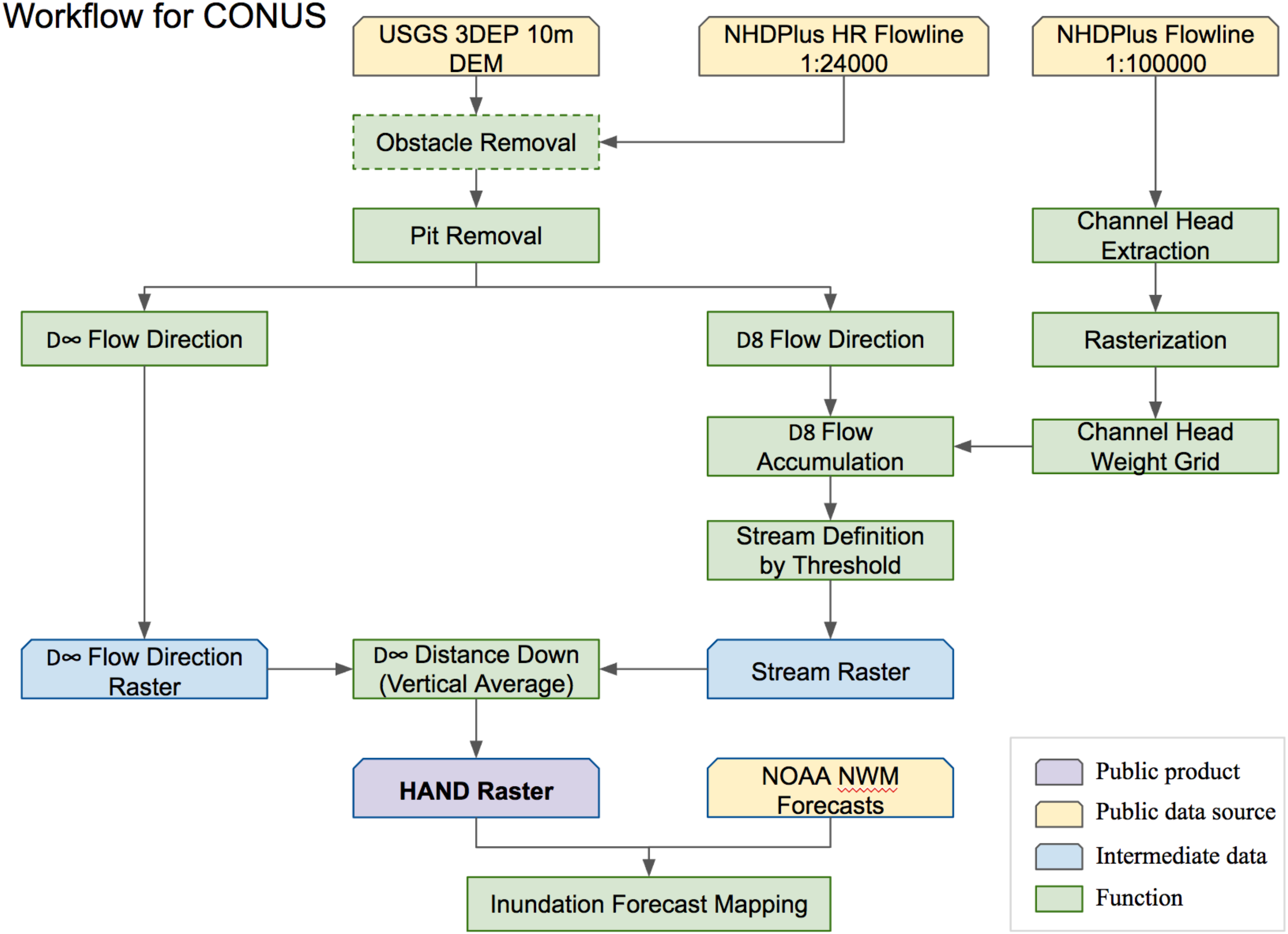
National Hydrography Dataset

National Elevation Dataset

Use the National CyberGIS Facility (ROGER supercomputer) operated by the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

HAND Workflow for CONUS



May 29, 2016. Derived from the original version by Xing Zheng

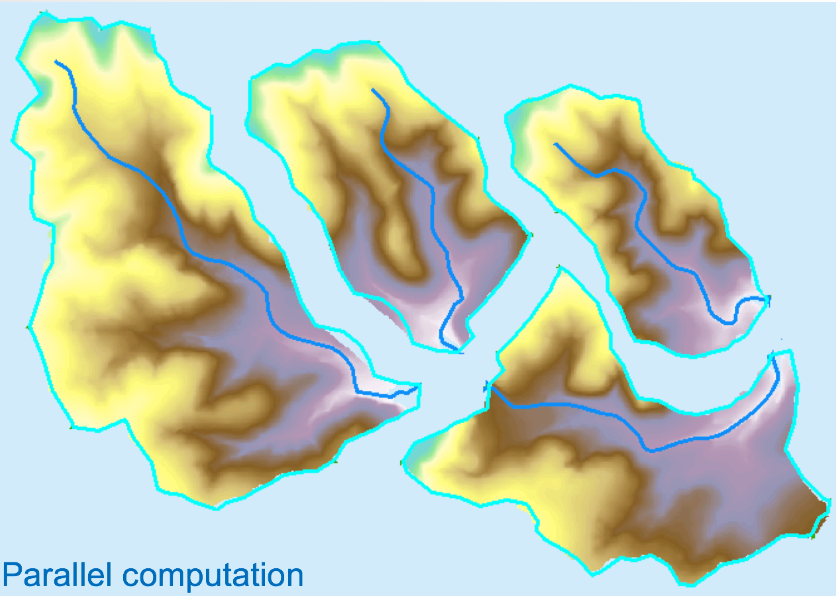
Computational Challenges

- Data
  - Input
    - DEM: USGS 3DEP elevation data: 10m (667GB GeoTIFF tiles; 180 billion cells)
    - NHDPlus MR: 1km (2.67 million vectors in FileGDB format)
    - NHDPlus HR: 250m (30 million vectors in FileGDB format)
  - Runtime and output
    - Multiple intermediate and final raster output of similar DEM input size
    - Multiple intermediate and final vector output, size proportional to NHDPlus resolution
- Computing
  - Parallel computing code – TauDEM
    - Scalability enhancement in the XSEDE ECSS TauDEM project
  - IO
    - Each TauDEM function reads multiple rasters and writes multiple rasters
  - Memory
    - Up to 6x input DEM size

Strategies

- High-throughput computing on HUC6 level
- High-performance computing for each HUC6 unit

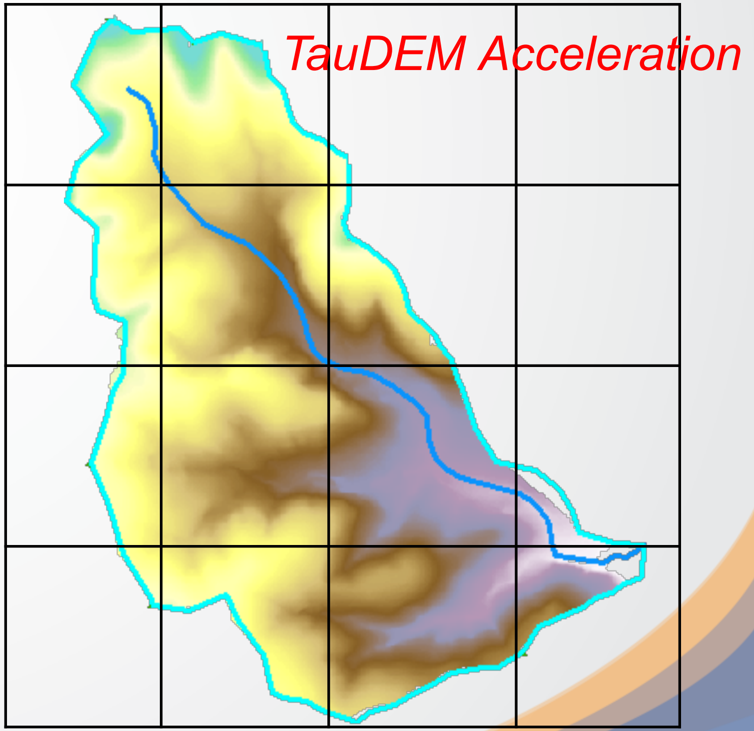
Hydrological unit decomposition



- Parallel computation
- High resolution (e.g. LIDAR) in some areas
- Incremental improvement

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

Parallel computing on individual units

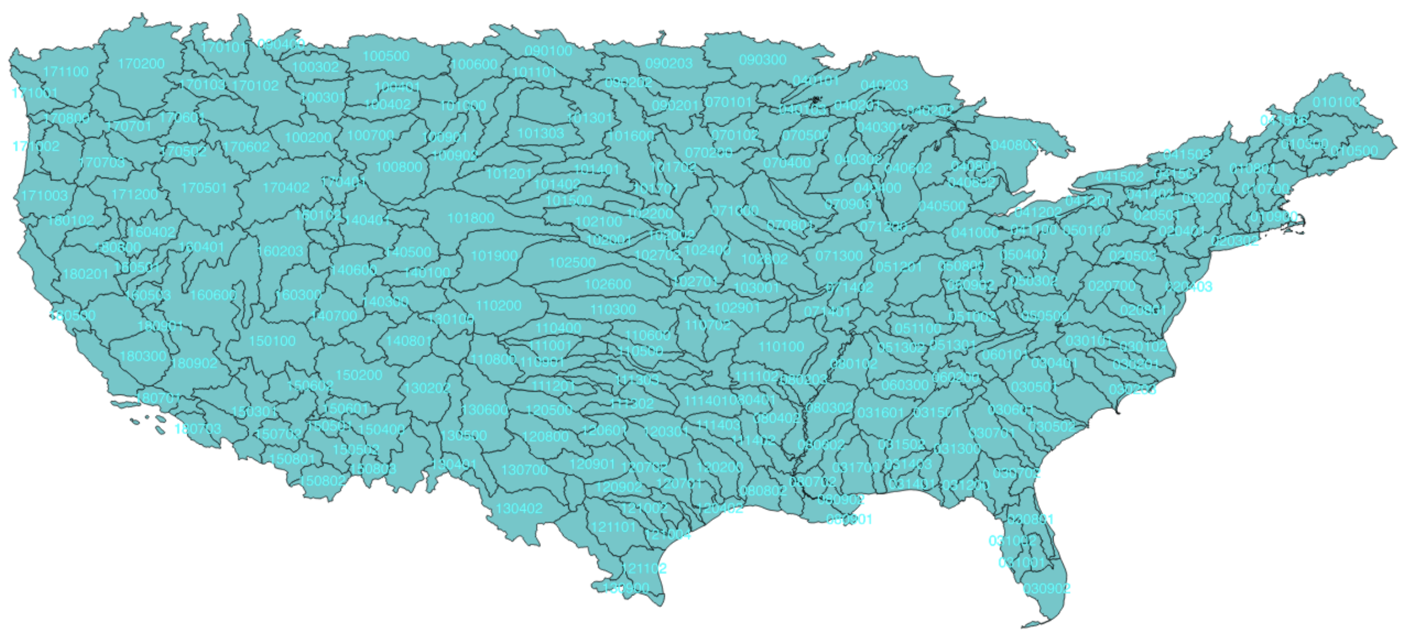


HUC6 Decomposition for CONUS

HUC system:
https://en.wikipedia.org/wiki/Hydrological_code

Name	Level	Digits	Average size (square miles)	Number of HUs (approximate)	Example name	Example code (HUC)
Region	1	2	177,560	21	Pacific Northwest	17
Subregion	2	4	16,800	222	Lower Snake	1706
Basin	3	6	10,596	370	Lower Snake	170601
Subbasin	4	8	700	2,200	Imnaha River	17060102
Watershed	5	10	227 (40,000–250,000 acres)	22,000	Upper Imnaha River	1706010201
Subwatershed	6	12	40 (10,000–40,000 acres)	160,000	North Fork Imnaha River	170601020101

HUC6 map:
331 units for CONUS



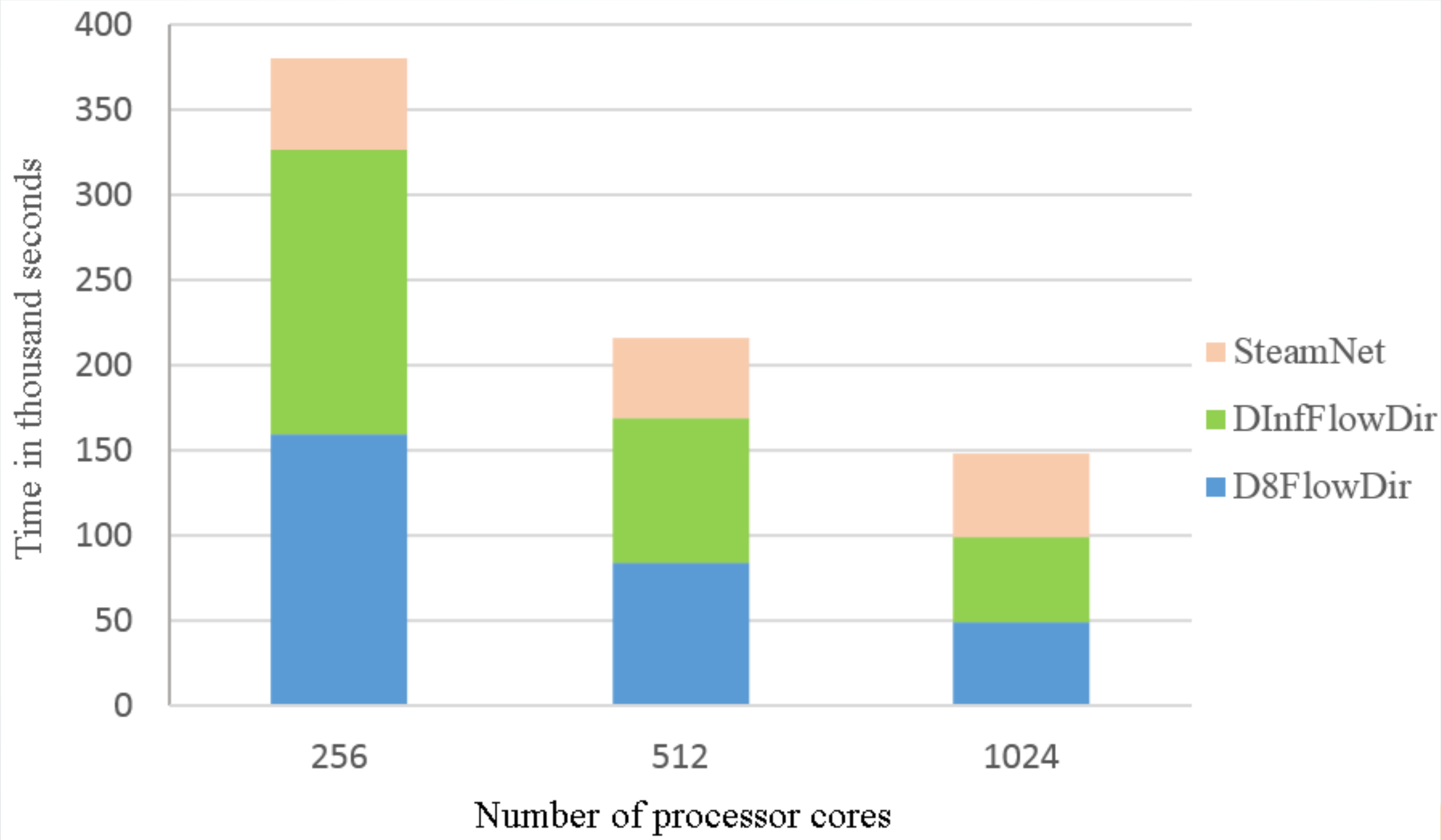
TauDEM Acceleration

- High-performance enhancements
  - Work update on the new flow direction algorithms
- TauDEM as science gateway services

TauDEM

- TauDEM - a Parallel Computing Solution to DEM-based Hydrological Information Analysis
  - Open source software
  - A suite of DEM tools for the extraction and analysis of hydrologic information from topographic data
  - A growing user community
- Parallel Computing in TauDEM
  - Parallel programming model: Message Passing Interface (MPI)
  - Spatial data decomposition
    - Each process reads a sub-region for processing
    - MPI communication for exchanging runtime hydrological information at border cells
    - Each process writes a sub-region defined by output data decomposition
    - Row-wise and block-wise decomposition
  - Input/output (IO)
    - MPI IO for DEM read and write – old version
    - GDAL IO

Three Costly TauDEM Functions

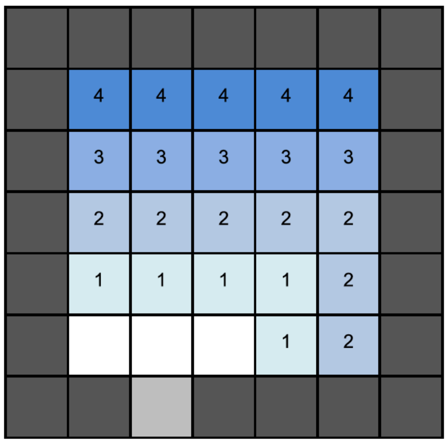


Execution time of the three most costly TauDEM functions on a 36GB DEM dataset. (*Fan et al. 2014*)

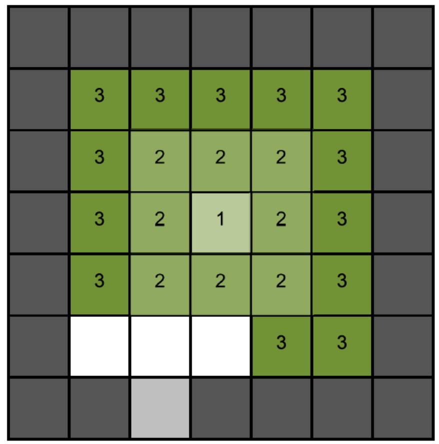
TauDEM Flow Direction Algorithm Acceleration

- Performance issues
  - 10m NED tile (1° x 1°) – 400MB
    - 3 hours with 1 processor
    - 1 hour with 4 processors
  - 5 NED tiles combined – 2.2GB
    - 1 hour with 100 (faster) processors
  - Most of the time is spent resolving flat regions
- Algorithm issues (Garbrecht & Martz 1997)
  - Pixel by pixel approach
  - No knowledge about the flat regions
    - Yet flat regions are independent
  - Every cell re-checked every iteration
    - ...even if the cell shouldn't change
    - ...even if the whole region is complete
    - Communication after every iteration

Flat Resolving Algorithm

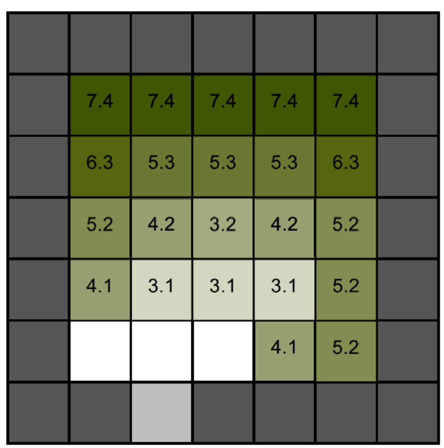


(a) Step 1



(b) Step 2

(c) Step 3



(d) Step 4

Step1. Gradient away from higher;
Step 2. Gradient toward lower;
Step 3. Combination;
Step 4. Recursive resolving if new flats appear

- Dark gray cells have higher elevation
- Light gray cells have lower elevation

New Parallel Algorithm

Algorithm 1 Parallel D8 flow algorithm

Require: Input elevation DEM D_p of the processor p
Ensure: Flow direction grid F_p of the processor p

```
1: F_p ← CalculateSlope(D_p)
2: if there exist cells with FLAT value in F_p then
3:     IS ← findIslands(F_p)
4:     while there exists shared flat regions do
5:         F_p ← resolveSharedFlats(IS)
6:     end while
7:     F_p ← resolveLocalFlats(IS)
8: end if
```

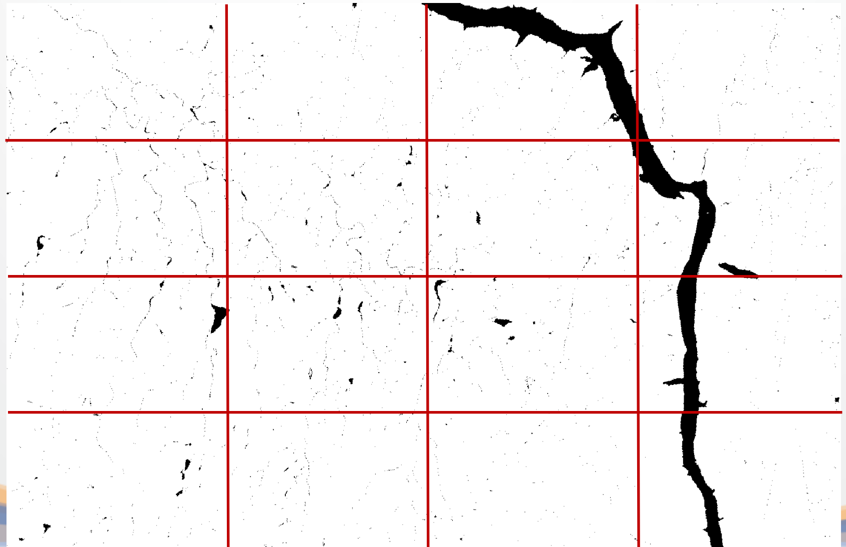
Survila, K., Yildirim, A.A., Li, T., Liu, Y., Tarboton, and D., Wang, S. 2016. "A Scalable High-performance Topographic Flow Direction Algorithm for Hydrological Information Analysis". *Proceedings of the 2016 Annual Conference on Extreme Science and Engineering Discovery Environment (XSEDE'16)*. July 17-21. Miami, Florida, USA

Sequential Algorithm Improvements

- Rewrite of the sequential algorithm
  - Identification of flat regions
  - Every flat cell is visited only once in flat resolving
- Worst case complexity reduced from $O(n^2)$ to $O(n)$
- Results
  - 10m NED tile (1° x 1°) – 400MB
  - Before:
    - 3 hours with 1 processor
    - 1 hour with 4 processors
  - After: 15 seconds with 1 processor

New Parallelization Strategies

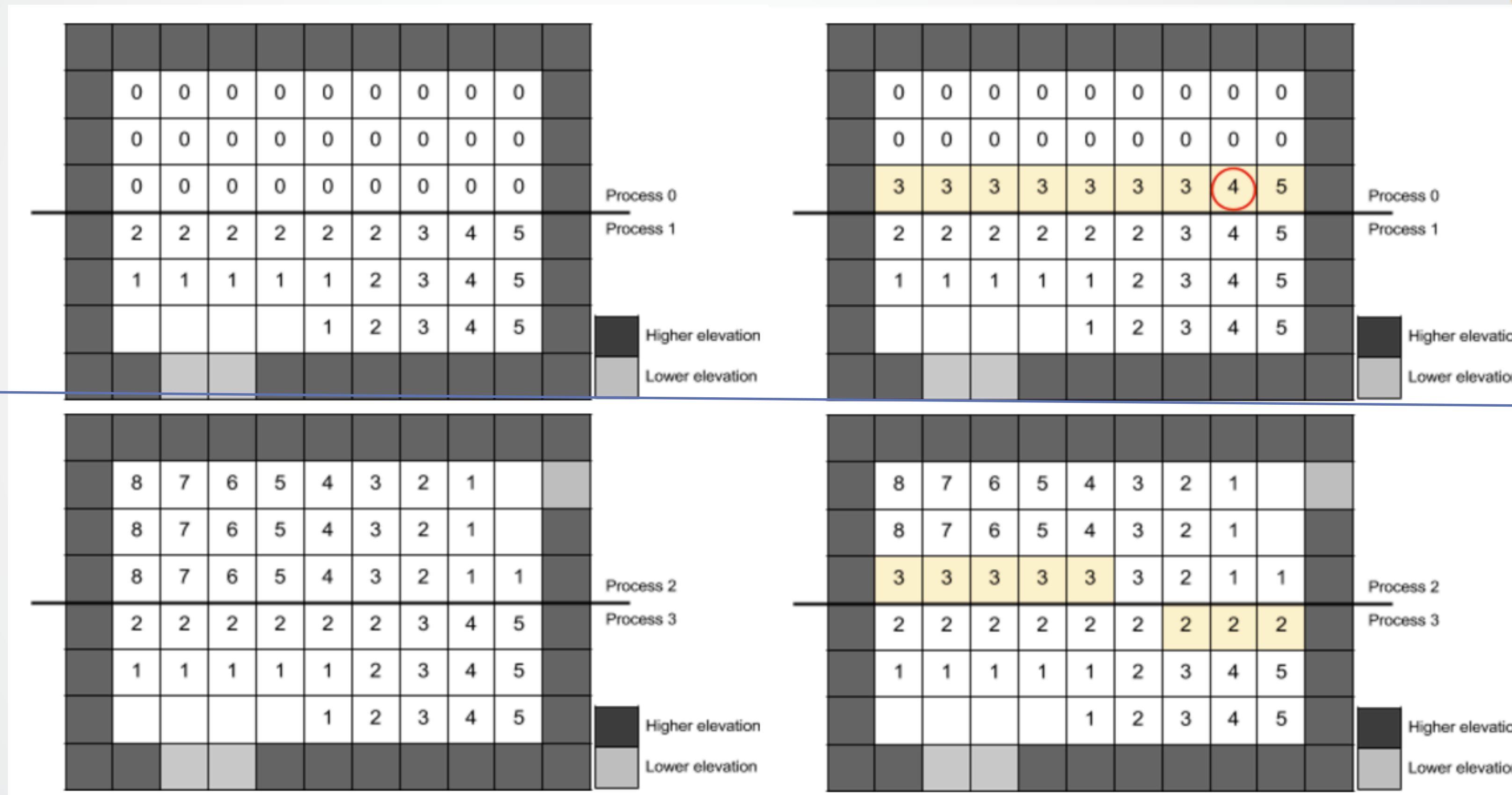
- Flat regions are independent
  - Figure out which regions cross the partition boundary and need to be synchronized
    - connected-component labeling
  - Rest of regions need no communication!
- Shared area of flat regions can vary from 10% to 80% based on region sizes and number of processors
  - Whole US 10m NED - 70% of flat areas are shared (with 200 processors)
  - Heavily influenced by data decomposition
  - Sparse partitioning
  - New communication layer implementation using non-blocking MPI calls



Resolving Shared Flats

Case 1

Case 2



Parallelization Results on ROGER

- Chesapeake DEM (~16 GB)
  - 100 cores original D8 - **8000s**
  - 20 cores new D8 - **177s**
  - 40 cores new D8 - 135s
- Contiguous US DEM (~700GB)
  - 750 seconds with 200 cores
    - 368 seconds spent in flat resolution
  - Infeasible with old algorithm

Performance

New D8 (100 million cells)

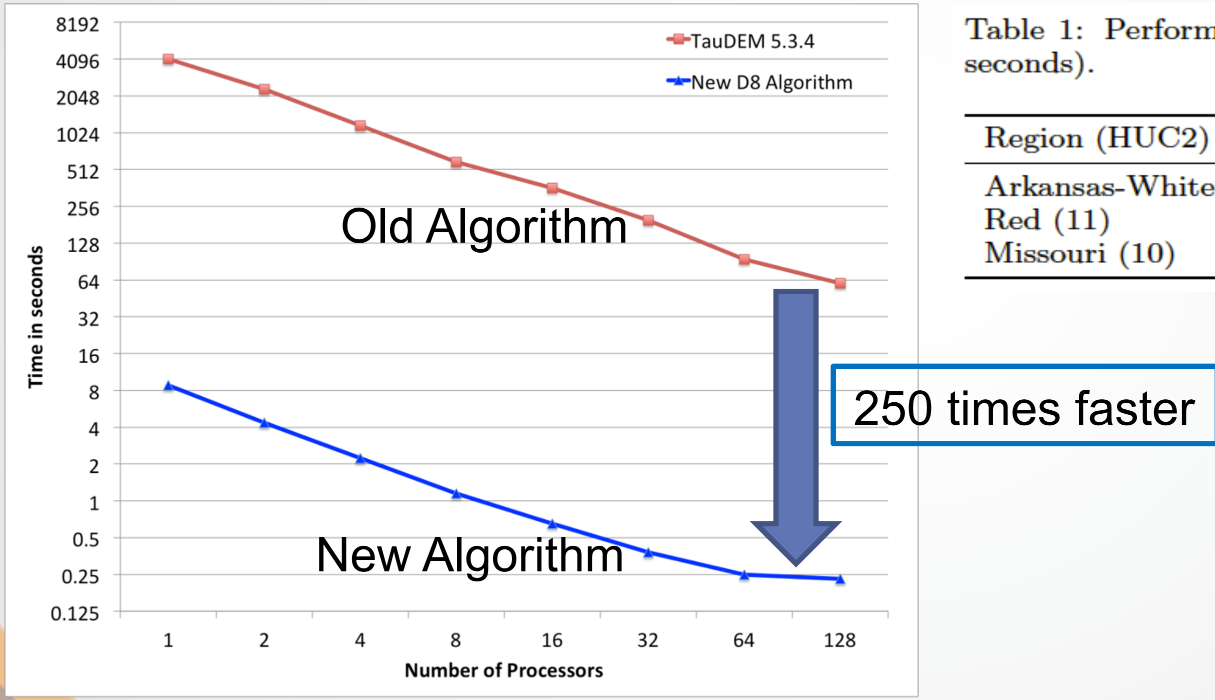


Table 1: Performance on large watershed DEMs (Time in seconds).

Region (HUC2)	Size	Num Flats	Time
Arkansas-White-Red (11)	170919x114734	433,807,730	39.57
Missouri (10)	227191x177787	1,225,082,919	91.43

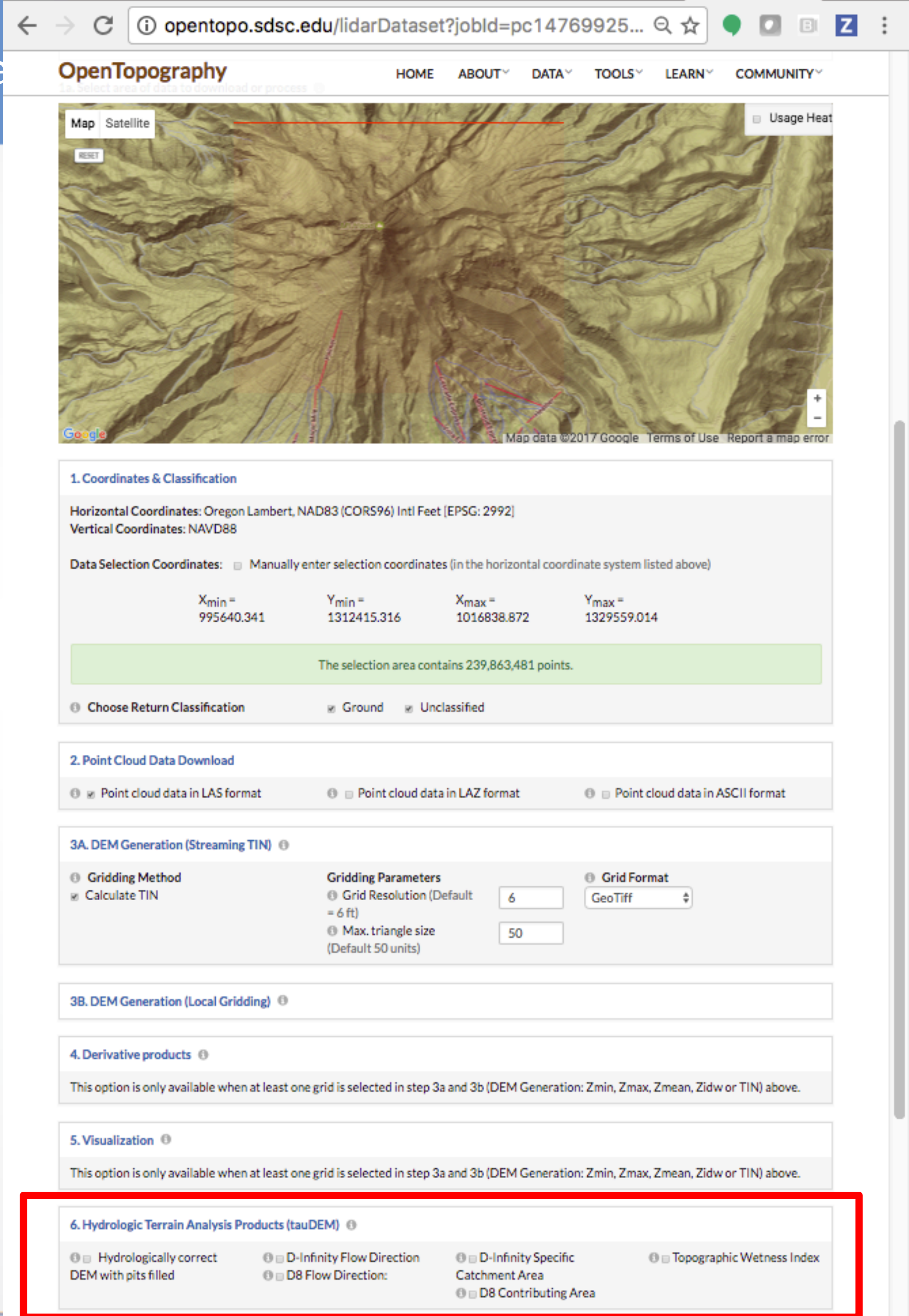
Github repository: https://github.com/cybergis/cybergis-toolkit/tree/master/taudem

TauDEM as Science Gateway Services

- LiDAR-based TauDEM analysis via OpenTopography
- Data and computing coupling through HydroShare-CyberGIS interoperability

TauDEM in OpenTopography

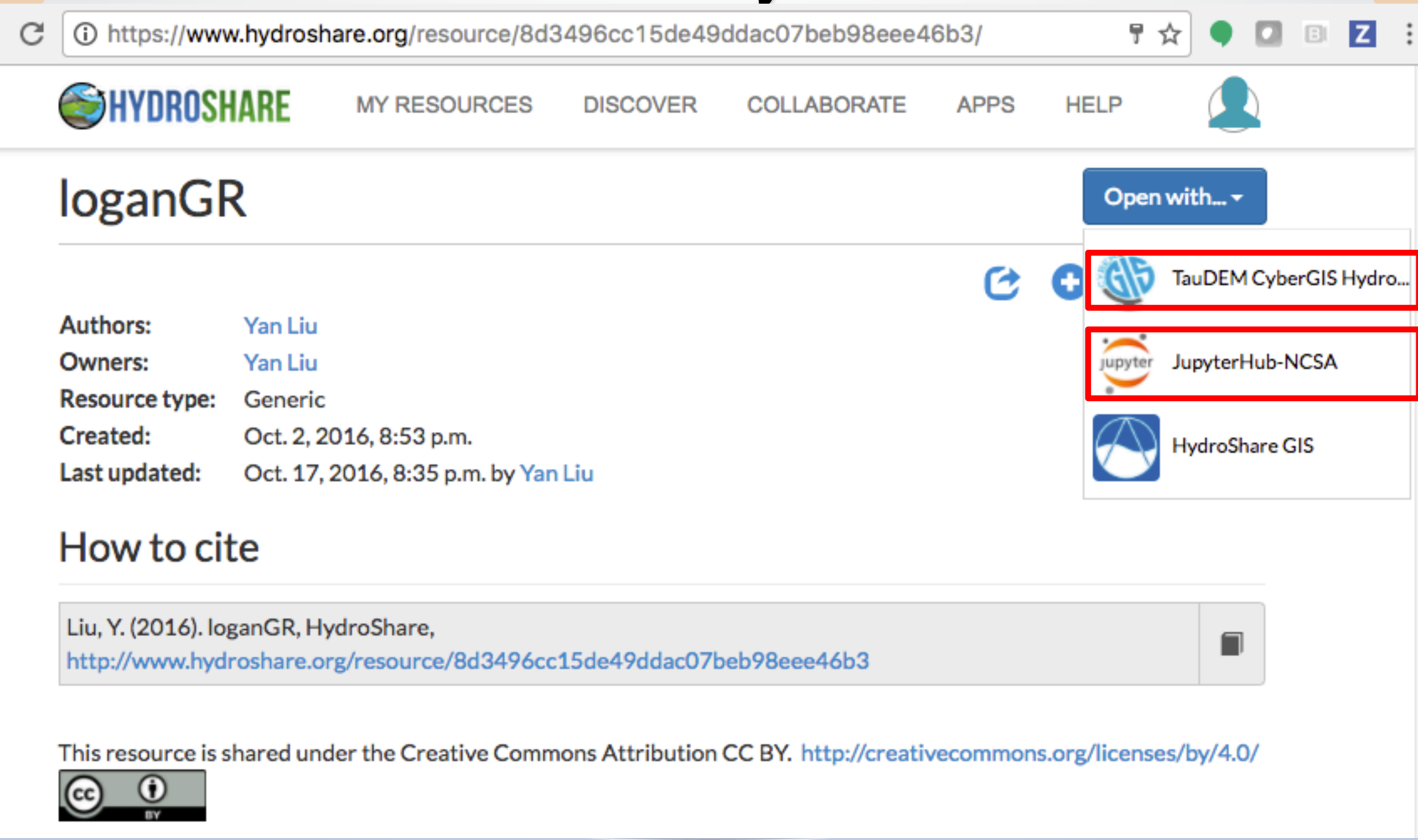
- Opal2 service integration
- On Gordon IO node
- Collaboration with OpenTopo team@SDSC

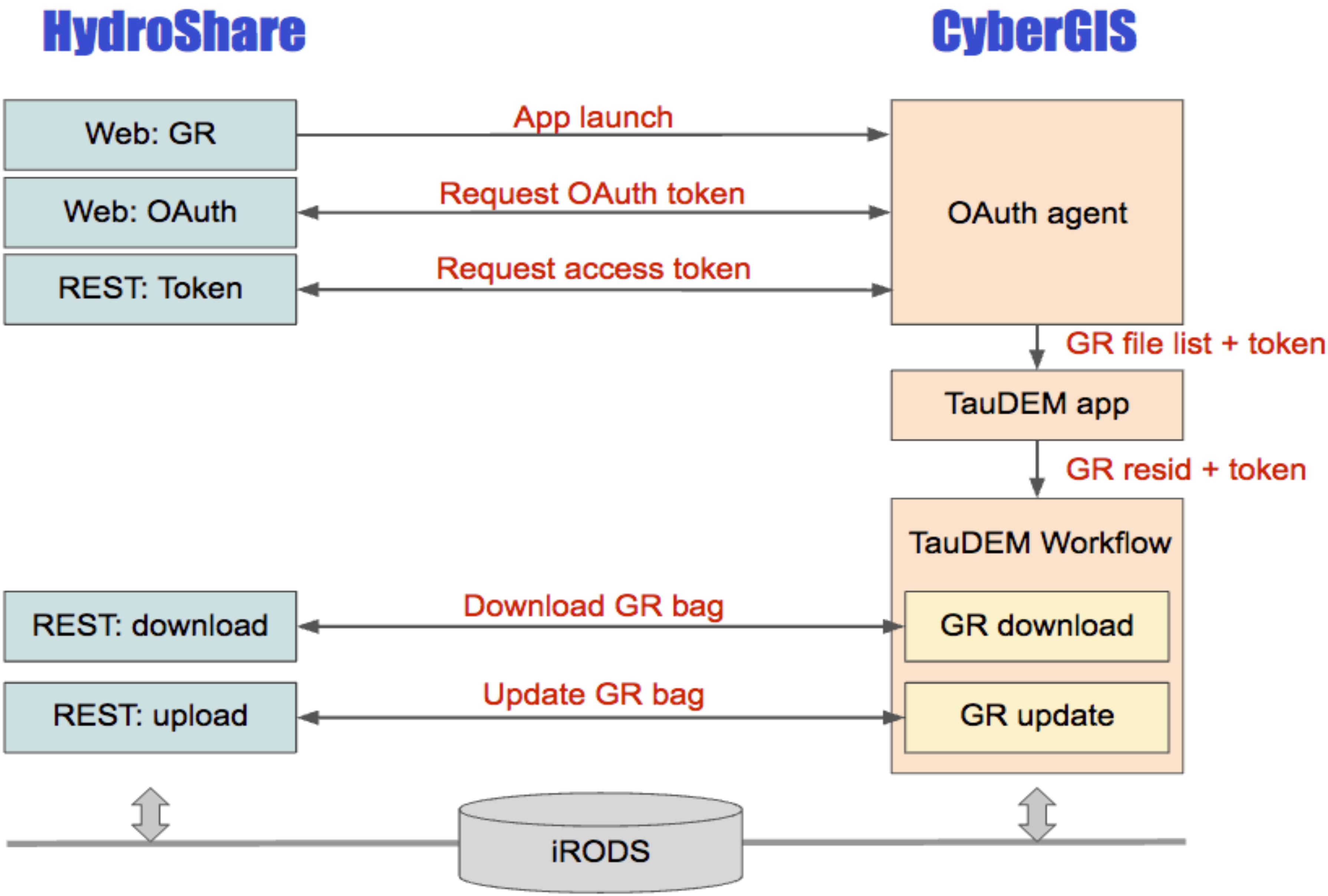


CyberGIS-HydroShare Interoperability

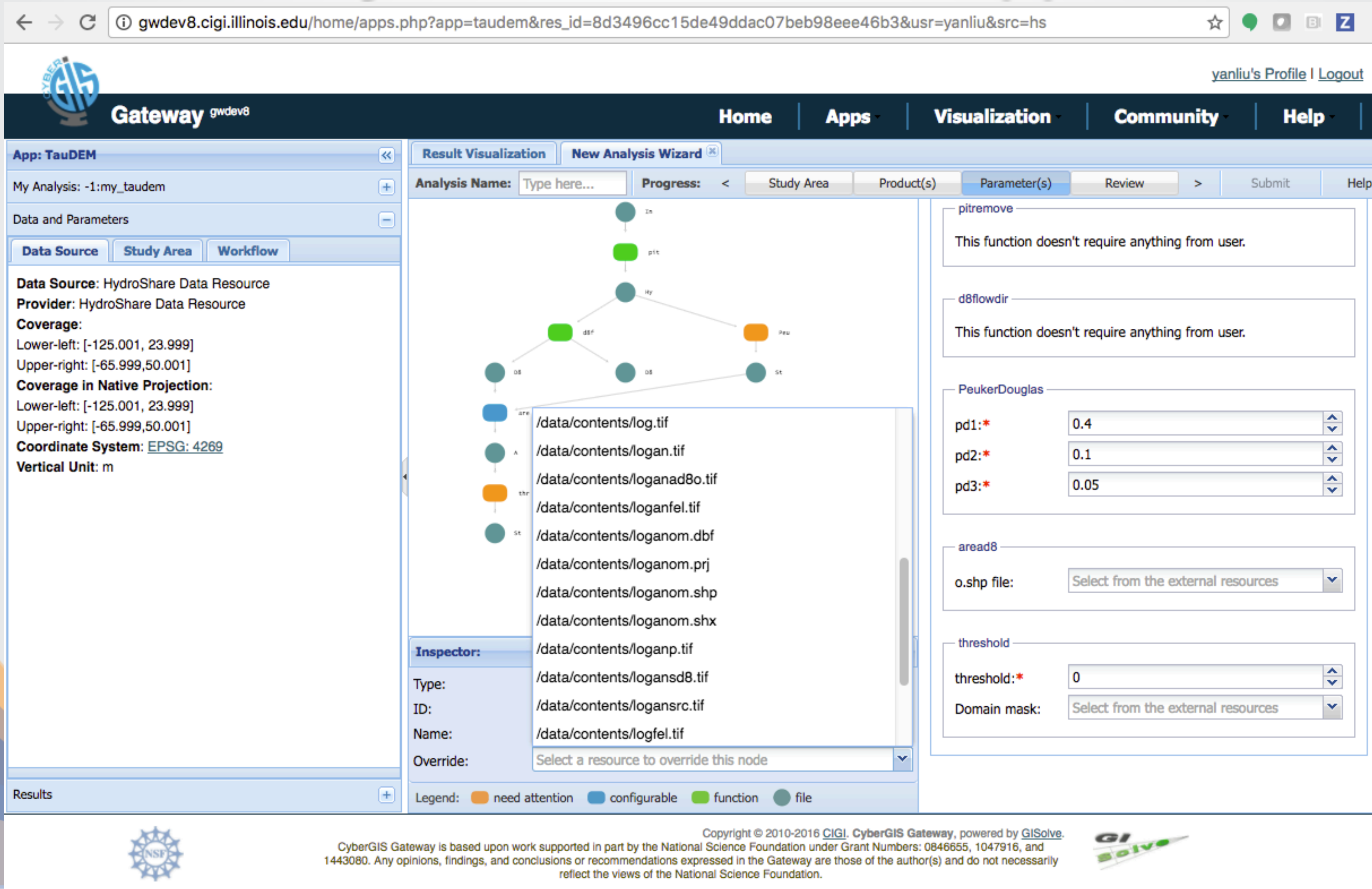
- CyberGIS Gateway as an online problem-solving environment
  - Provider of analysis apps, computation on cyberinfrastructure, and interactive analysis environment
  - TauDEM gateway app
- HydroShare as a hydrological data portal
  - Data store for hydro-related data, including GIS data
  - REST API based on Django REST framework
  - App framework as a simple wrapper of external services
- Interoperability model
  - Portal-portal interoperability
    - Security: Oauth
    - Service interaction: REST
    - Web: HTML5 and web mapping
  - Distributed data management via iRODS
    - User data space model (one generic resource (1GR))
  - Interactive analysis using JupyterHub

TauDEM in HydroShare




HydroShare
CyberGIS
Web: GR
App launch
Web: OAuth
Request OAuth token
REST: Token
Request access token
OAuth agent
GR file list + token
TauDEM app
GR resid + token
TauDEM Workflow
REST: download
Download GR bag
GR download
REST: upload
Update GR bag
GR update
iRODS

CyberGIS TauDEM app



hydroroger.ncsa.illinois.edu/user/yanliu/notebooks/notebooks/help/taudem.ipynb

jupyter taudem Last Checkpoint: Yesterday at 4:15 PM (autosaved)

Welcome yanliu | Control Panel | Logout

File Edit View Insert Cell Kernel Widgets Help | Python 2

Markdown | CellToolbar

TauDEM Raster Processing Notebook

This notebook demonstrates basic watershed processing using the TauDEM GIS library. For more information about TauDEM, see http://hydrology.usu.edu/taudem/taudem5/. This tutorial is divided into the following three categories:

- Setup and Preparation
- GIS Processing using TauDEM
- New Resource Creation using HydroShare

....

2. Remove DEM Pits

In [ ]:

```
# Fill the DEM Pits

# set the output paths
fill = os.path.join(data_directory, 'loganfel.tif')
cmd = 'pitremove -z %s -fel %s' % (raw_dem_path, fill)
taudem.run_cmd(cmd, processors=4)
```

3. Calculate Flow Directions

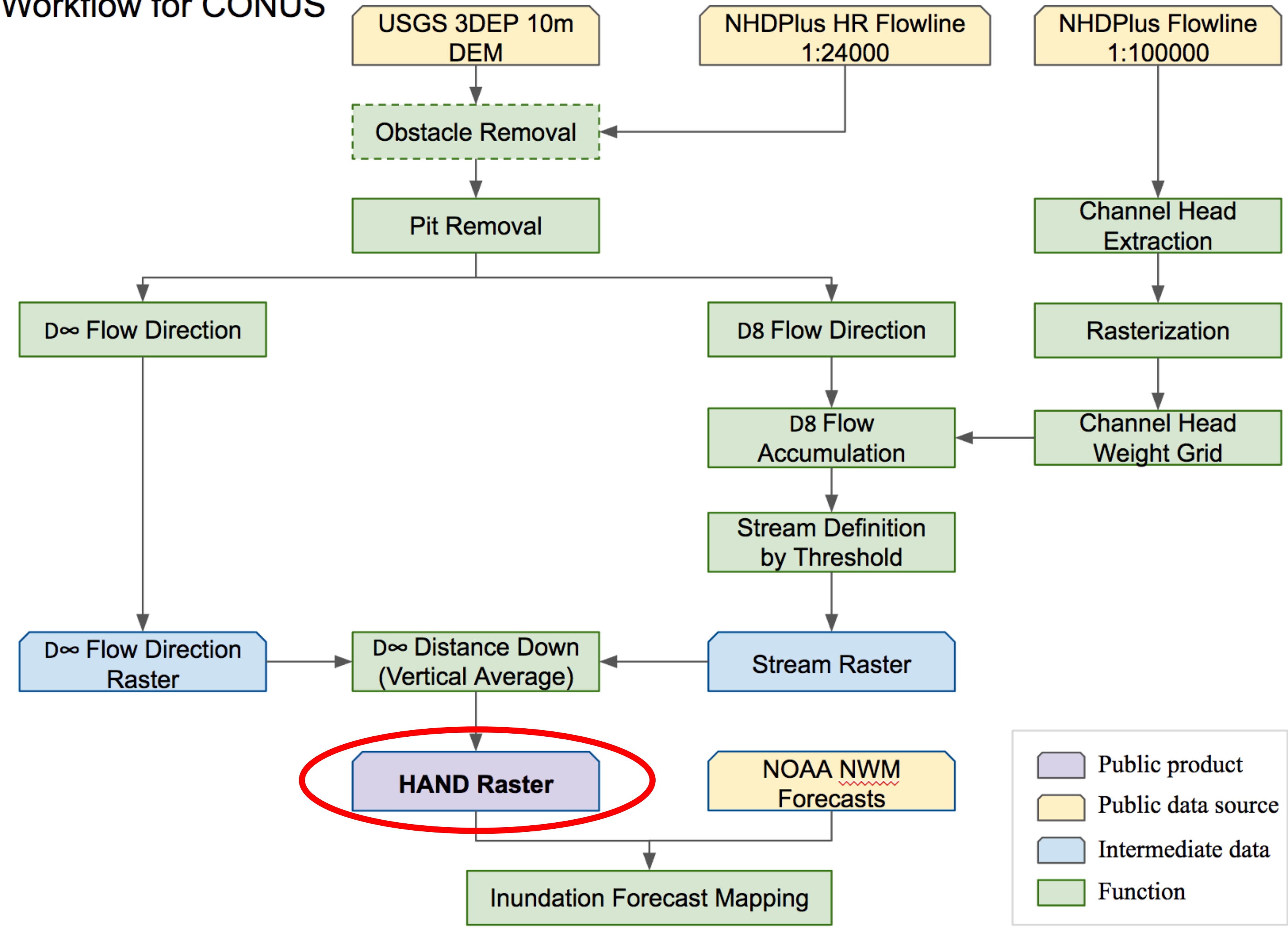
In [ ]:

```
#####################
# D8 Flow Direction #
#####################

# set the output paths
fdr = os.path.join(data_directory, 'fdr.tif')  # flowdir
sd8 = os.path.join(data_directory, 'sd8.tif')  # slope

cmd = 'd8flowdir -fel %s -sd8 %s -p %s' % (fill, sd8, fdr)
taudem.run_cmd(cmd, processors=4)
```

HAND Workflow for CONUS



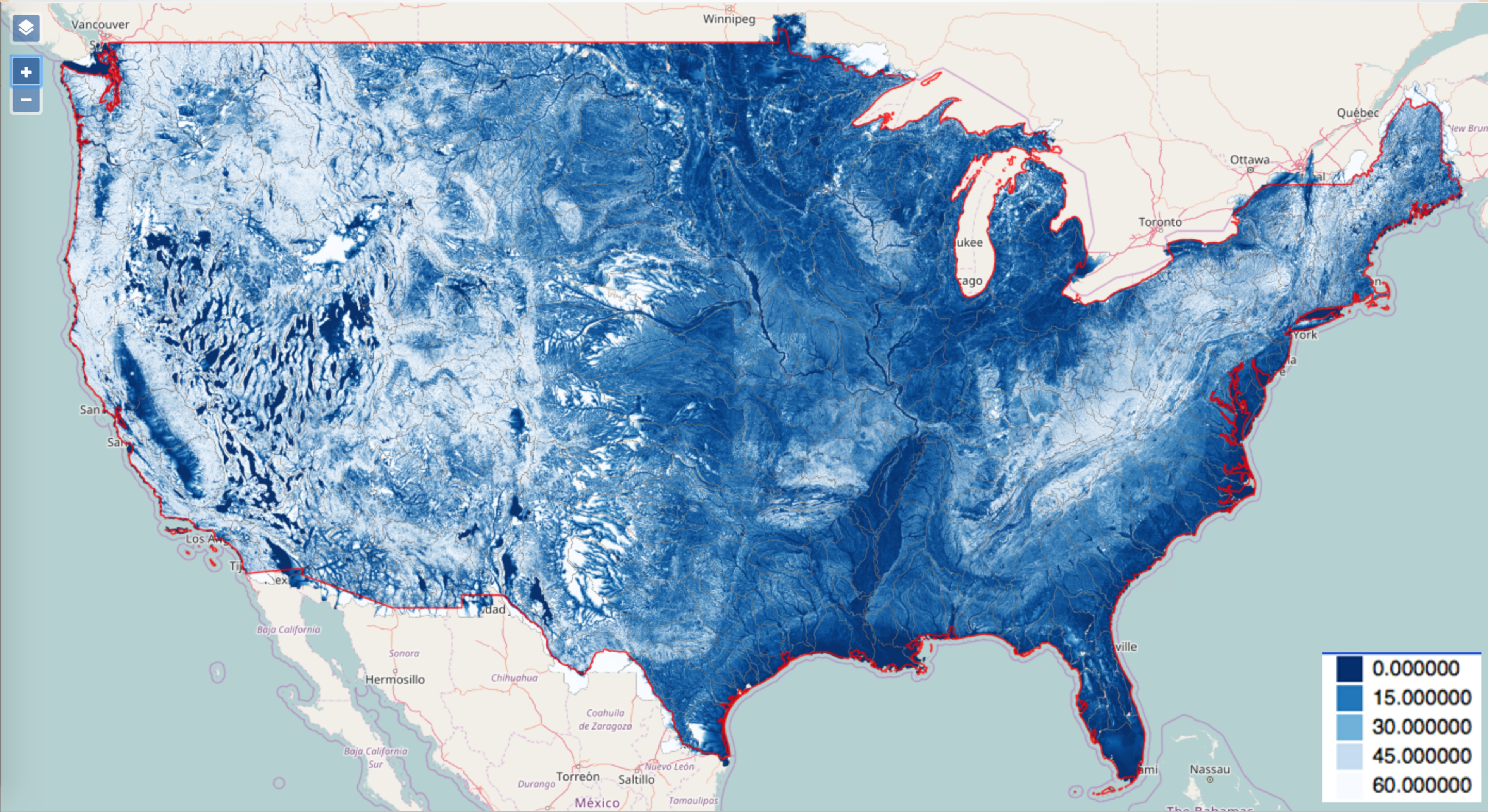
May 29, 2016. Derived from the original version by Xing Zheng

Workflow Development

- Collaboration
  - UT Austin, Utah State Univ., UIUC, RENCI, USGS
  - NSF project collaboration: CyberGIS, HydroShare, NFIE, ROGER MRI
- Code
  - https://github.com/cybergis/nfie-floodmap
- Cyberinfrastructure
  - ROGER@UIUC
  - Job management: PBS (Torque and Maui)
  - MPI: mpich 1.3
  - TauDEM: 5.3.5 with GDAL IO support and new flow direction algorithms
  - GDAL: 1.11.4 and 2.0.2 (with PROJ4, GEOS, HDF5/NETCDF4, Esri FileGDB, spatialite/sqlite, Java, Python support)
  - Data sharing: iRODS and web server (on ROGER cloud)
  - Ensemble computation: GNU Parallel
  - Visualization
    - Tiled Map Service (TMS) (http://nfie.roger.ncsa.illinois.edu/nfie/TMS/)

CyberGIS Center for Advanced Digital and Spatial Studies

HAND 10m for CONUS



HAND Output

- Intermediate + final results: 4.9TB
- HAND raster
  - 363GB compressed GeoTIFF as VRT (virtual raster format)
  - Online map (zoom level 5-10) through Tile Map Service (TMS)
    - http://nfie.roger.ncsa.illinois.edu/nfiedata/maps/maps.html#source=..%2Fyanliu%2Fviz%2Fhuc6-mercator-new-singlelayer.json&extent=-14392000_2436200_-7279500_6594375

Metadata for each HUC6 HAND output

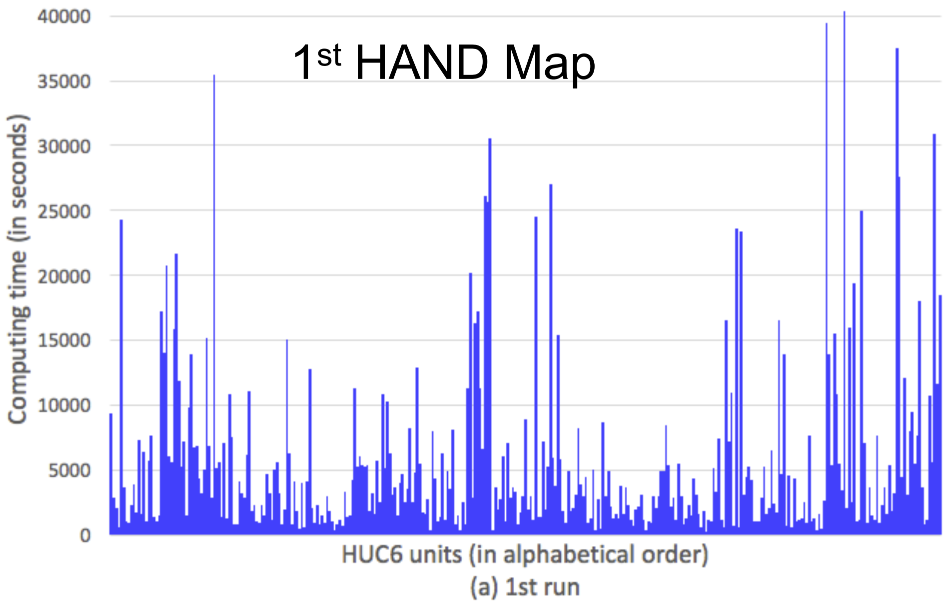
```
${HUCID}-wbd.shp - Boundary polygon shape file of the HUC unit, extracted from USGS WBD
${HUCID}-wbdbuf.shp - Buffered boundary polygon shape file of the HUC unit (10km buffer)
${HUCID}-flows.shp - Flowline shape file of the HUC unit, extracted from NHDPlus medium resolution
${HUCID}-inlets0.shp - Point shape file of all the inlets in the HUC unit, in NHDPlus projection
${HUCID}-inlets.shp - Point shape file of all the inlets in the HUC unit, in NED projection EPSG:4269
${HUCID}-weights.tif - Weight grid of the rasterized inlet points, used by TauDEM aread8 as weight
${HUCID}.tif - clipped DEM from USGS 10m National Elevation Dataset (NED) with a buffer size (10km)
${HUCID}fel.tif - pit removed DEM; output of TauDEM pitremove
${HUCID}p.tif - D8 flow direction raster; output of TauDEM d8flowdir
${HUCID}sd8.tif - D8 slope raster; output of TauDEM d8flowdir
${HUCID}ang.tif - Dinfinity flow direction raster; output of TauDEM dinfflowdir
${HUCID}slp.tif - Dinfinity slope raster; output of TauDEM dinfflowdir
${HUCID}ssa.tif - Contributing area raster; output of TauDEM aread8
${HUCID}src.tif - Stream grid; output of TauDEM threshold (threshold=1)
${HUCID}dd.tif - Bufferred HAND raster; output of TauDEM dinfdistdown (vertical distance down)
${HUCID}hand.tif - HAND raster, buffer removed, final result
```

Computation Summary

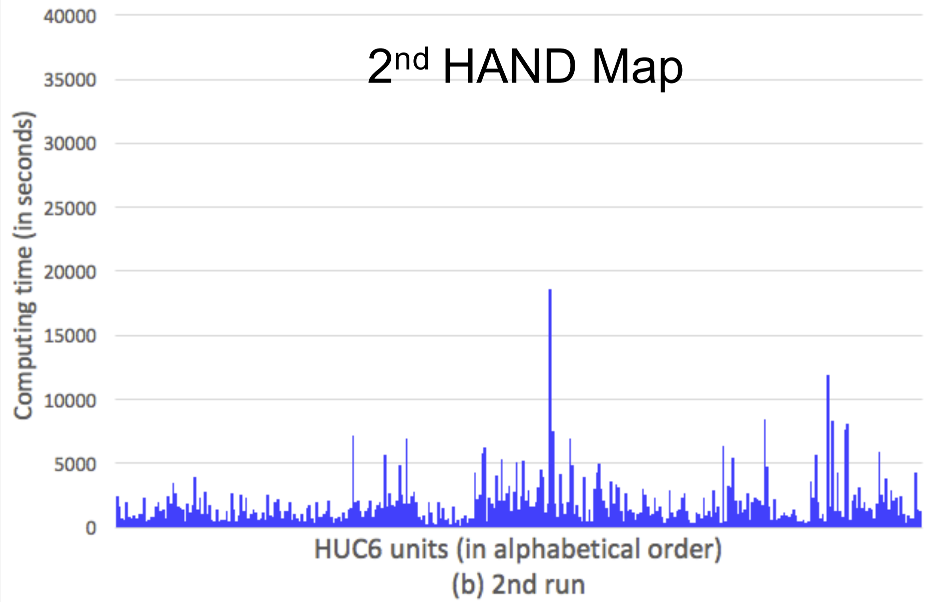
- *36 hours in total for 331 HUC6 units*
- 1.34 CPU years
- On average, each job used 65.26 cores and took 0.54 hours to compute
- The two flow direction algorithms only took 12.65% computing time, on average
- The rest of the time was spent on transformation and pre- & post-processing of geospatial data

Computing Time Comparison

Old flow direction algorithms



New flow direction algorithms



5.76 CPU Years

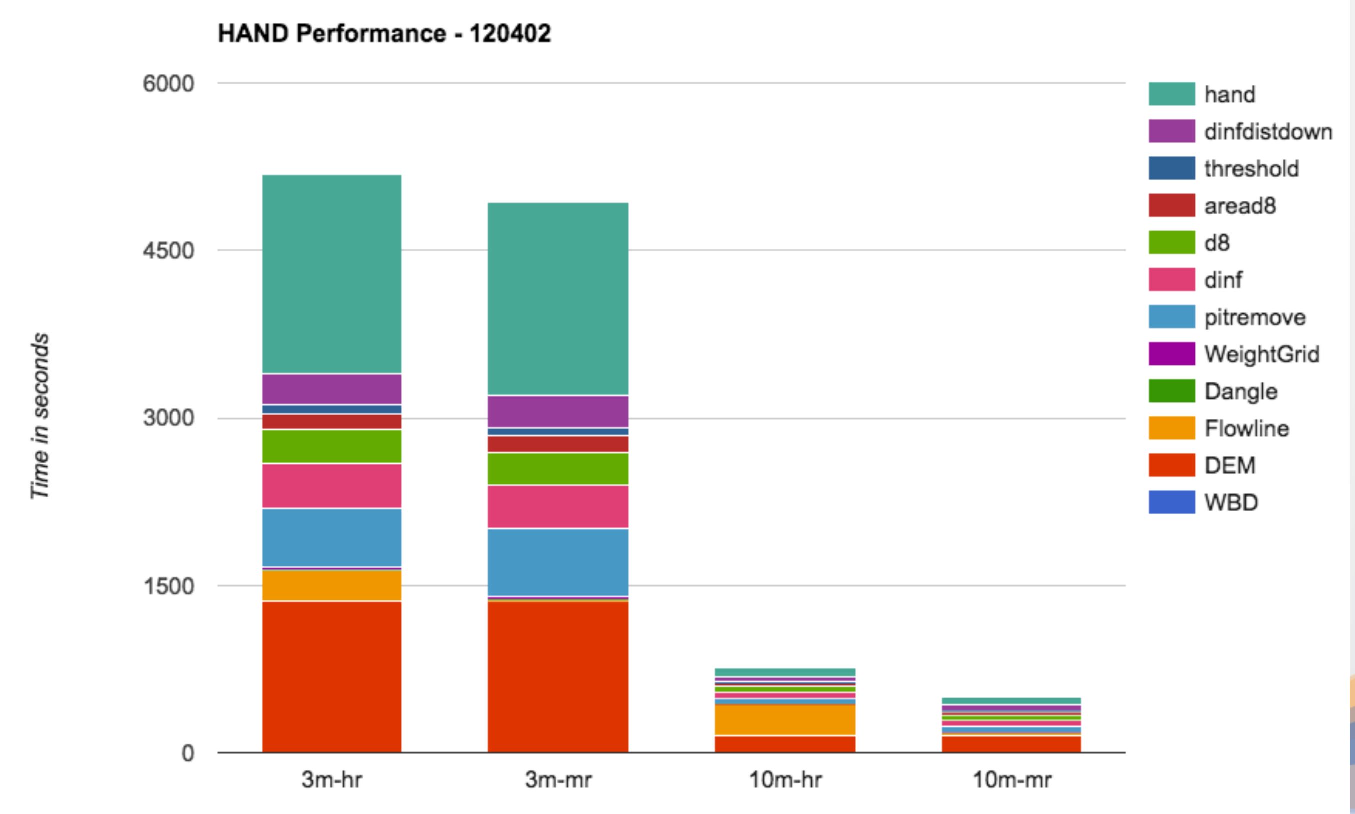
8 days of elapsed time

1.34 CPU Years

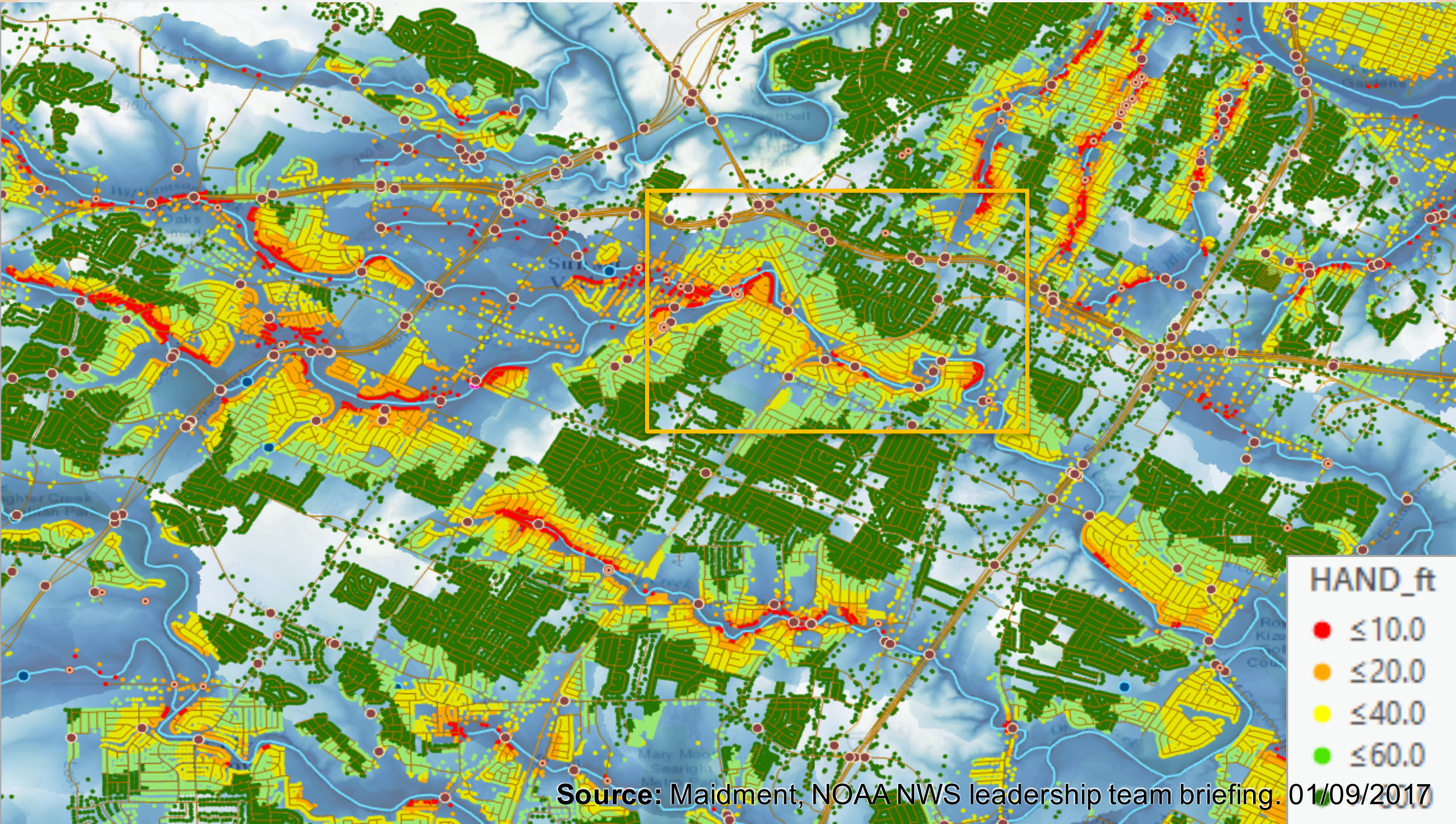
1.5 days of elapsed time

Scalability - HAND Computation

- Experiment on HUC6 unit 120402
  - 6 nodes, 120 cores
- 3m *vs.* 10m DEM
  - 5.4GB *vs.* 595MB
- NHDPlus MR *vs.* NHDPlus HR



Williamson Creek, South Austin: address points attributed with HAND



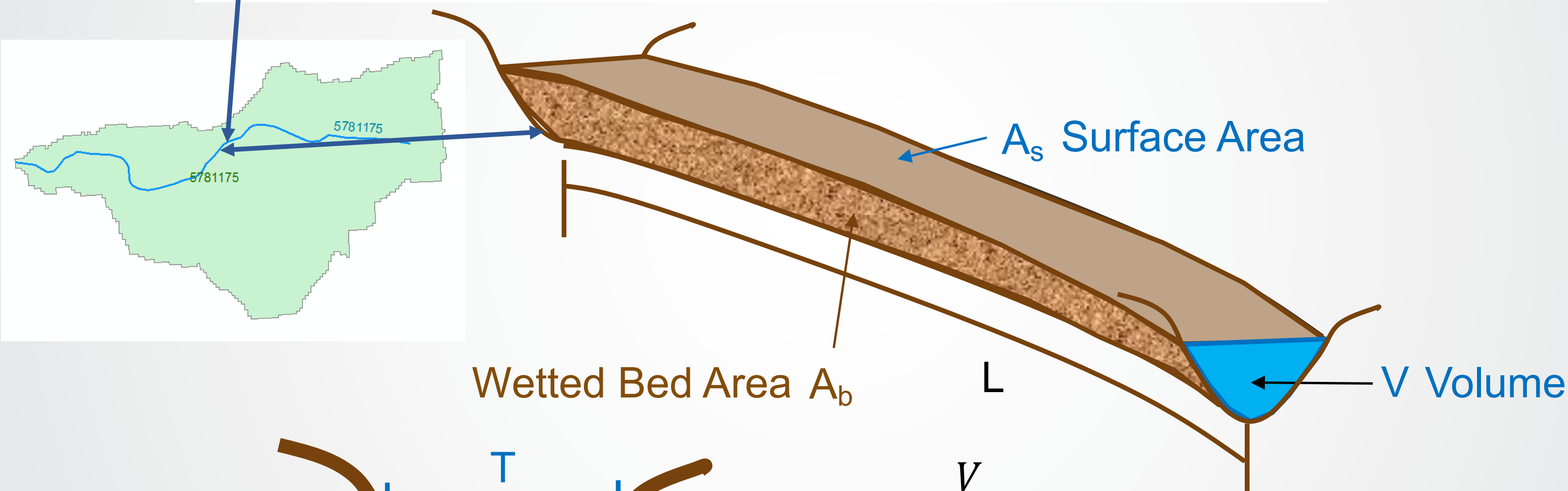
Source: Maidment, NOAA NWS leadership team briefing. 01/09/2017

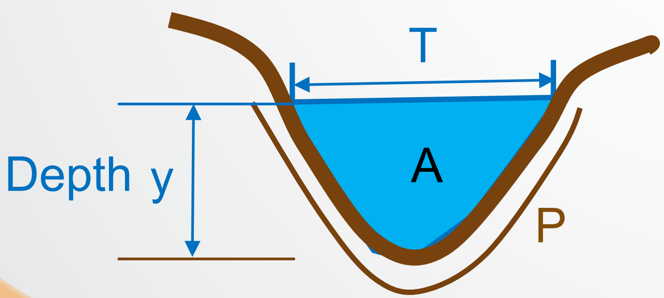
Inundation Mapping

- Create hydro property table for CONUS
  - 2.7m river reaches
  - 82 water depth stages
- Inundation forecast mapping
  - Inundation forecast table for CONUS
  - Inundation map for each HUC6 unit
  - Inundation map for CONUS
  - Visualization

Reach Hydraulic Parameters

Comid	y	A	R	P	T	V	Ab	As
5781175	3							
5781175	4							





$A = \frac{V}{L}$ Cross Section Area

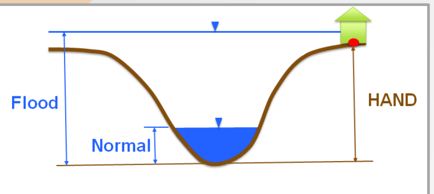
$P = \frac{A_b}{L}$ Wetted Perimeter

$T = \frac{A_s}{L}$ Top Width

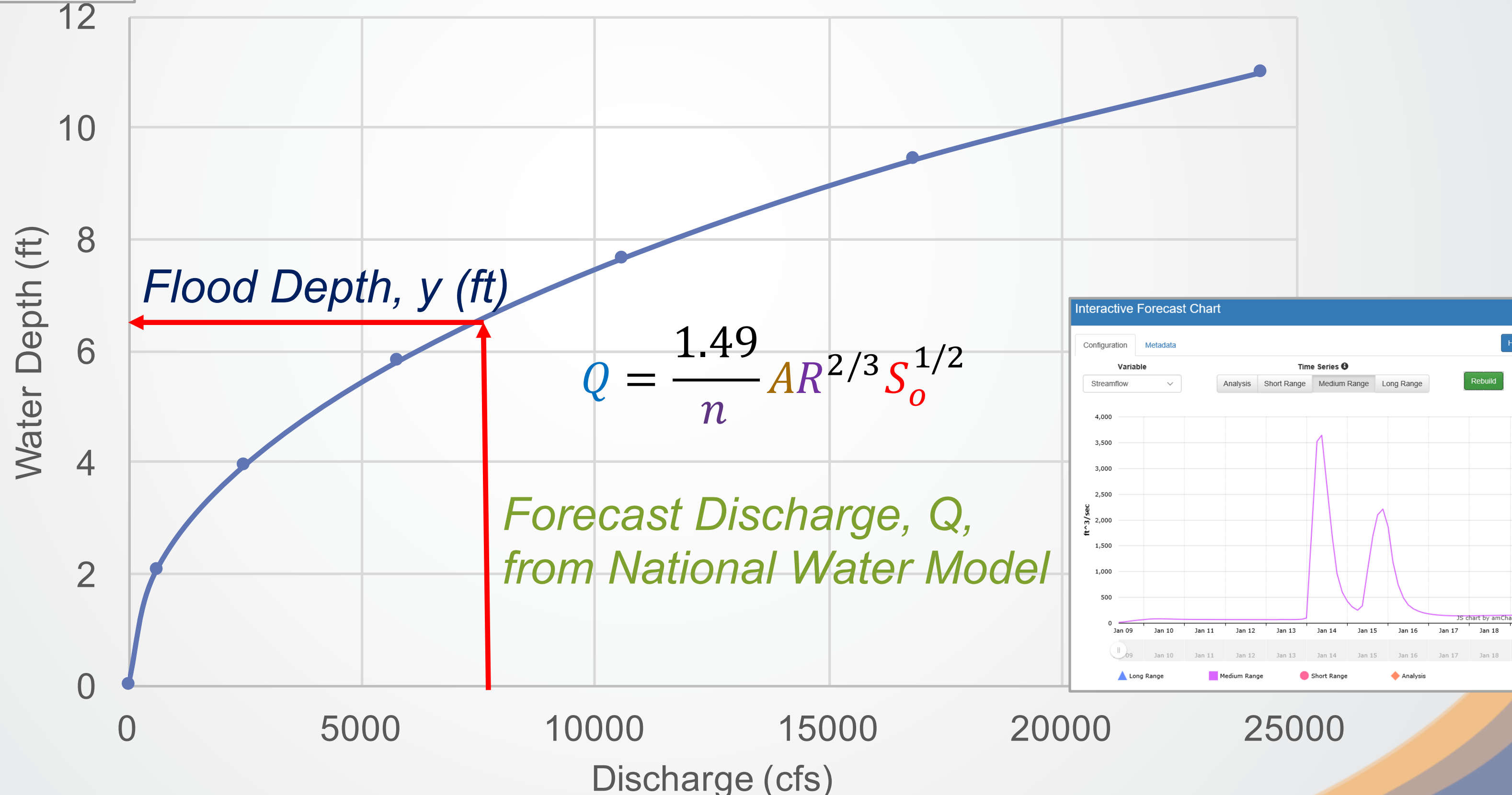
$R = \frac{A}{P}$ Hydraulic Radius

Source: Maidment, NOAA NWS leadership team briefing. 01/09/2017

Rating Curve – Connects Discharge with Depth

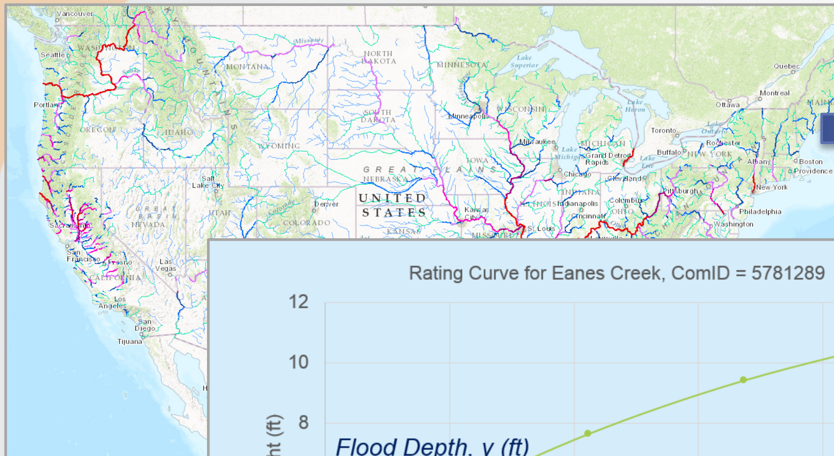


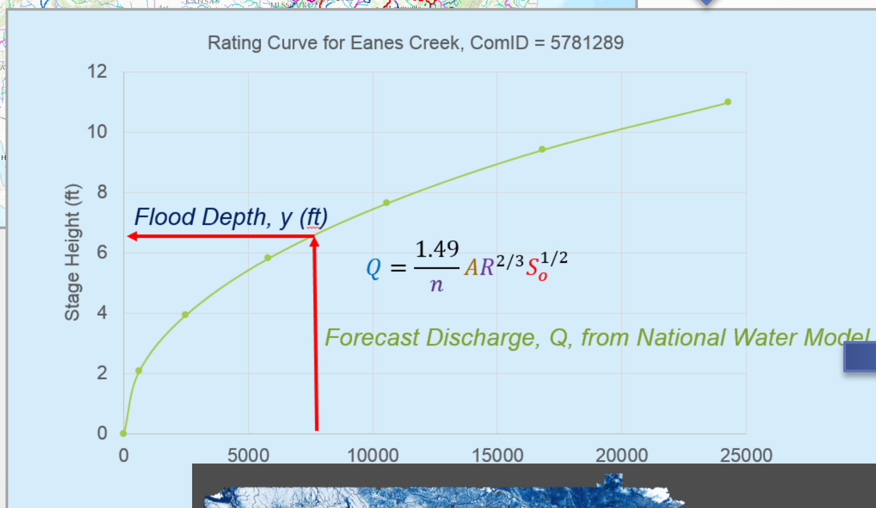
Rating Curve for Eanes Creek, ComID = 5781289

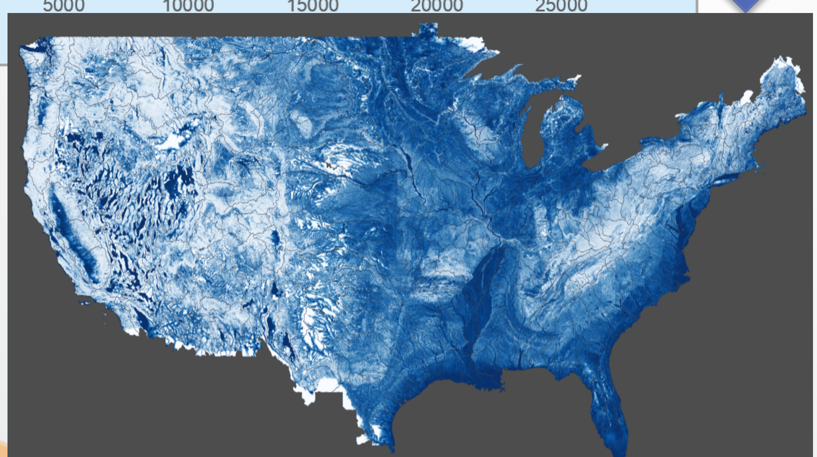


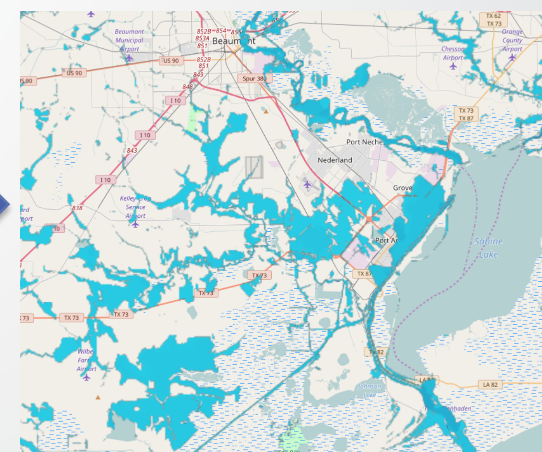
Source: Maidment, NOAA NWS leadership team briefing. 01/09/2017

Continental-Scale Flood Inundation Mapping









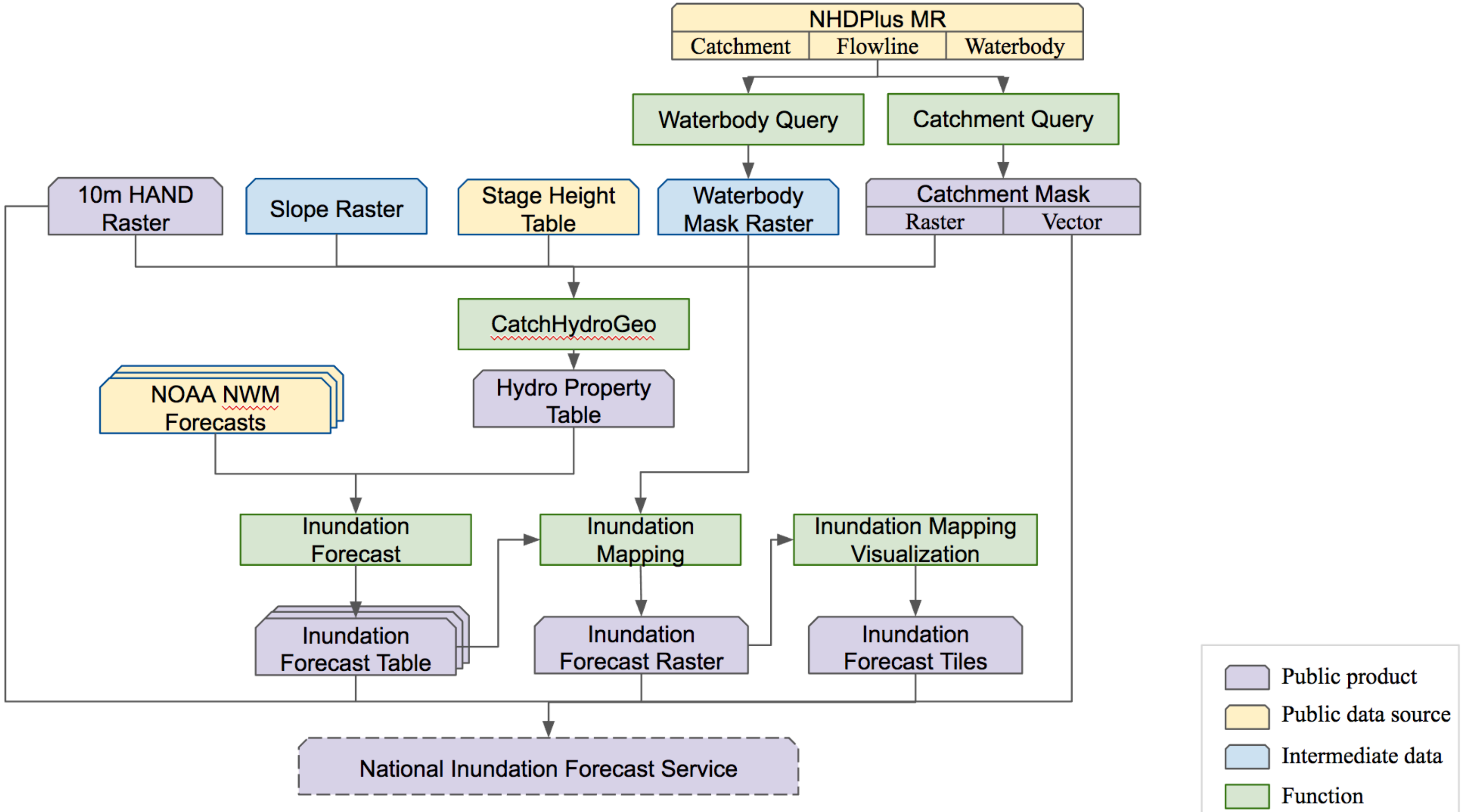
1. Forecast **discharge** with National Water Model

2. Convert discharge to **depth** using rating curve

3. Convert depth to **inundation** using HAND

Source: Maidment, NOAA NWS leadership team briefing. 01/09/2017

Inundation Forecast for CONUS



NWM Operational Configuration

Running Continuously on WCOSS since May 9th

	Analysis & Assimilation	Short-Range	Medium-Range	Long-Range
Cycling Frequency	Hourly	Hourly	Daily at 06Z	Daily Ens (16 mem)
Forecast Duration	- 3 hrs	0-15 hours	0-10 days	0-30 days
Forecast Latency (latency of external forcing data accounts for most of delay)	1 hour	1 hour 45 mins	6 hours	19 hours
Meteorological Forcing	MRMS blend/ HRRR/RAP bkgnd.	Downscaled HRRR/RAP blend	Downscaled GFS	Downscaled & bias-corrected CFS
Spatial Discretization & Routing	1km/250m/NHDPlus Reach	1km/250m/NHDPlus Reach	1km/250m/NHDPlus Reach	1 km/NHDPlus Reach

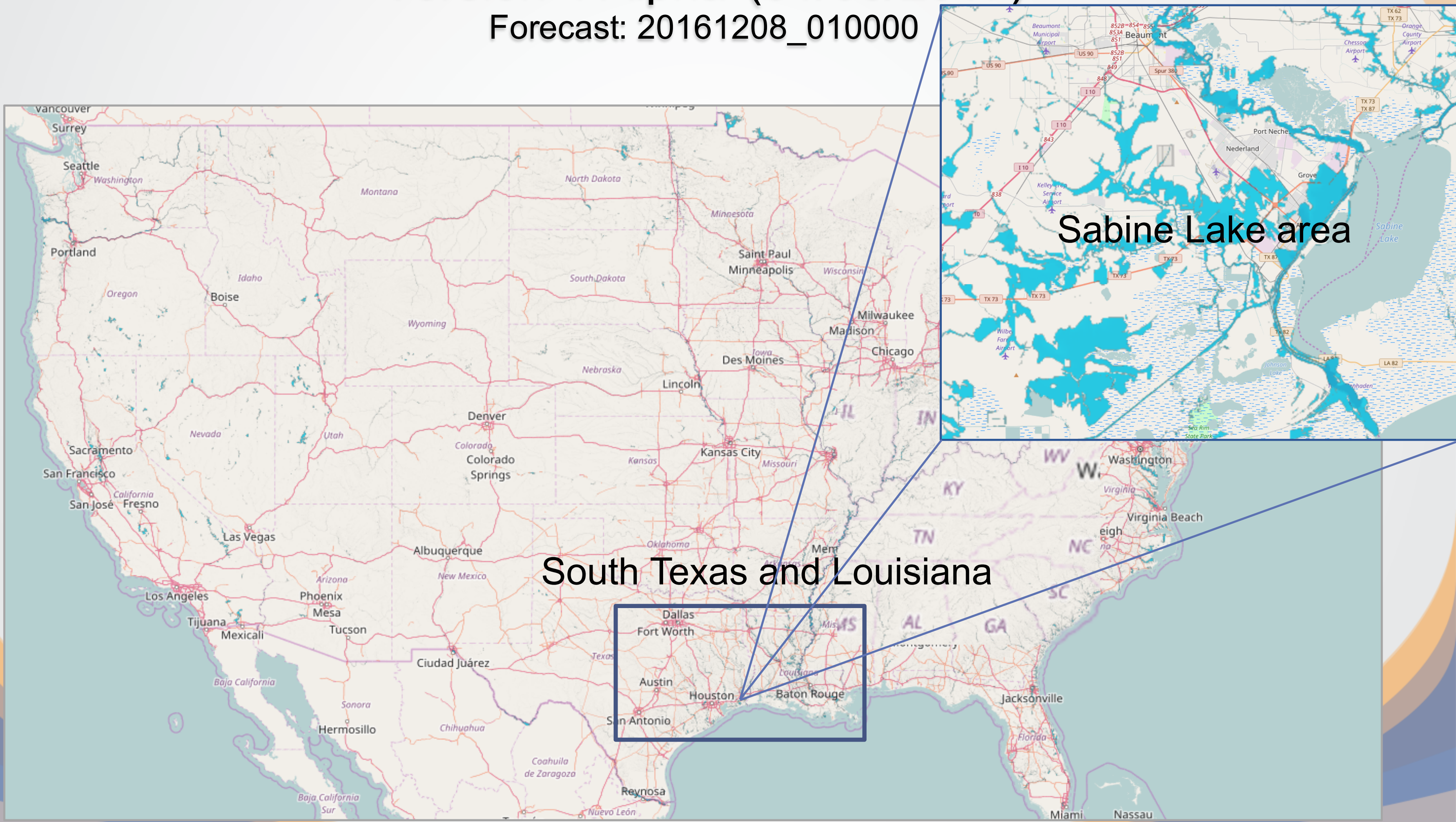
Assimilation of USGS Obs

Reservoirs (1260 water bodies parameterized with level pool scheme)

Source: http://www.nws.noaa.gov/com/weatherreadynation/files/NationalWaterModel.pdf

Continental-Scale Flood Inundation Mapping Process, Version 1 Alpha (01/06/2017)

Forecast: 20161208_010000



Source: Maidment, NOAA NWS leadership team briefing. 01/09/2017

CONUS-scale Computation Cost

Product	Size	Input	Performance on ROGER	Frequency
10m HAND raster - HUC6	635939 x 282122 (180b cells); 718GB uncompressed geotiff	USGS NED 10m: same size NHDPlus MR: 2.69m reaches 18GB FileGDB NHDPlus HR: 30m+ reaches	36 hours for 331 HUC6 units; 1.34 CPU years; 65.26 processors/unit; 0.54 hours/unit	Once
10m catchment mask raster - HUC6	same as HAND	NHDPlus MR: Catchment layer Flowline layer Each with ~3m vectors	330 HUC6 units: 16 hours; 1 processor	Once
Catchment mask vector - HUC6	2.69m catchment polygons			
Hydro property table - CONUS	2638933x83 rows 14 columns indice: COMID and stage height	HAND raster: Slope raster: Catchment mask: same size as HAND Stage height table: 83 height values	330 HUC6 units: 10.5 hours; 1 node, 20 processors	Once
Inundation forecast table - CONUS	2615396 rows 3 columns (comid, H, Q)	NWM forecast per Time: 2699225 stations 6 columns Hydro property table	3hr15min using 330 HUC6 hydroprc 8min using CONUS netcdf; one processor	Per NWM forecast
Inundation map raster - HUC6	HAND raster size / 4 * 2 byte datatype; pyramids	HAND raster; HUC6 comid raster; Waterbody mask raster; Inundation forecast table	5hrs; 1 node, 1 processor	Per NWM forecast
HUC6 Inundation map viz: TMS tiles generation	zoom level 5-10: 300MB	Inundation map raster; Coloring table	6.5hrs: 1 node, 10 processors 1hr: 10 nodes, 10 processors/node	Per NWM forecast
CONUS Inundation map viz: TMS tilemerge for CONUS	zoom level 5-10: 151MB	TMS tiles	28m; 1 node, 1 processor	Per NWM forecast

Ongoing Work

- Methodology improvements
  - Soft burnin
  - River channel geometry
  - Flat and coastal area handling (non-DEM approaches)
  - High-resolution LiDAR-derived DEM
- A big data visualization challenge
  - Exploring industrial and open source solutions
  - Big raster+vector visualization
    - Animation
    - Data partitioning, indexing, querying, and online streaming
- Software code polishing and release
- EOT
  - UCGIS summer school (May 2017)
  - NWC summer institute (June 2017)

Conclusion

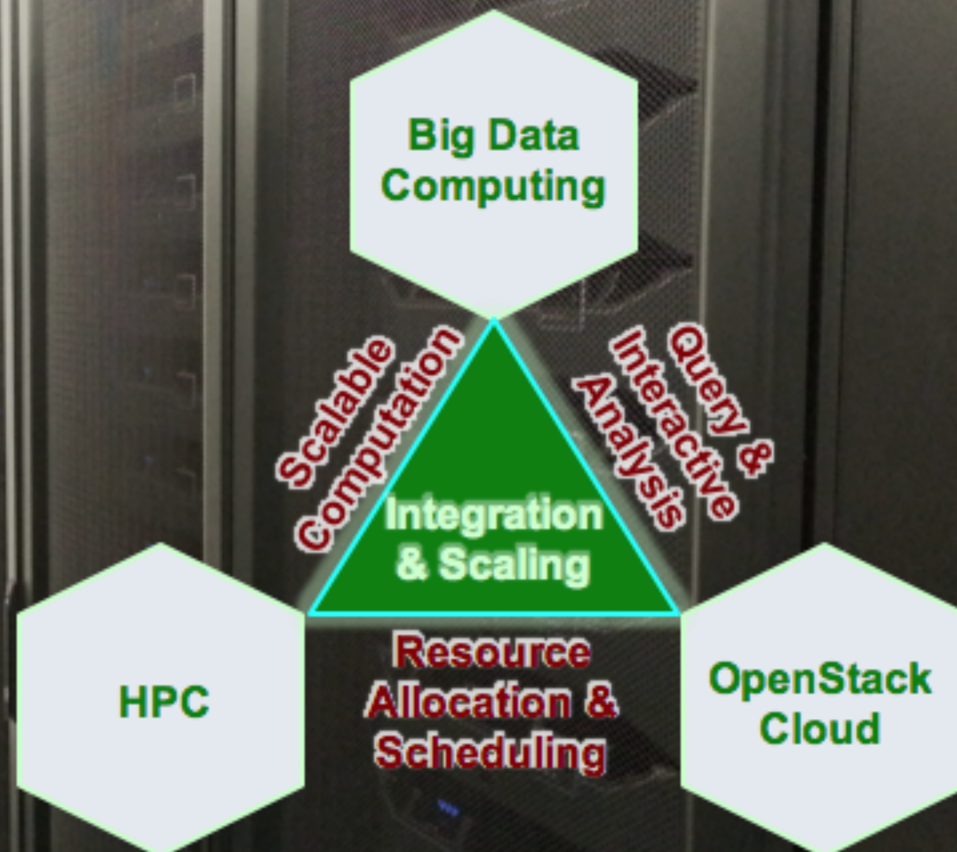
- Built all the technical pieces for continental inundation mapping
- Produced the 1st 10m HAND map for CONUS
  - May 29, 2016
- TauDEM accelerated to handle high-resolution DEMs for CONUS
- Produced the 1st real-time inundation map for CONUS at 10m raster resolution and 1km NHDPlus vector resolution
  - Jan 06, 2017
- Community feedback
  - USGS and NOAA NWS leadership team briefings, 07/18/2016 and 01/09/2017
  - Conference presentations
    - AWRA'16 summer, CyberGIS'16
  - EOT
    - NWC summer institute'16
    - Online class Fa'16

Discussion

- ECSS
  - Sometimes, benchmarks and profiling may not be enough
    - New parallel computing algorithms need to be developed
  - Community engagement
    - Computation as a way to engage domain science community expertise
      - Supercomputing makes it possible to present data, results, process of unprecedented sizes to end users
- XSEDE resource usage models for end users
  - Advantages of hybrid supercomputing architecture
    - Sandbox solution
      - Resources: HPC, cloud, storage
      - User space: data, software, online services, user environment
  - End-to-end solutions
    - Lower the barrier for users to pick up the sandbox
    - Usability
      - Short turnaround on issue resolving
      - Regular user engagement
    - Sustainability and portability
    - Measures
      - Data products
      - Software products
      - Online services
      - Science use cases

CyberGIS ROGER

- Dell PowerEdge r730
- 24 Compute Nodes
  - 2x Xeon 2.6GHz
  - 10 cores each
  - 128GB RAM
  - 500GB HD
  - 10 Gb network
- 12 GPU Nodes
  - Nvidia Tesla K40M
- 16 High Memory Nodes
  - 256GB RAM
  - 800GB SSD
- 2 GridFTP Nodes
  - 40 Gb network



Resourcing Open Geospatial Education and Research
after Roger Tomlinson
~45 TFLOPS

ROGER

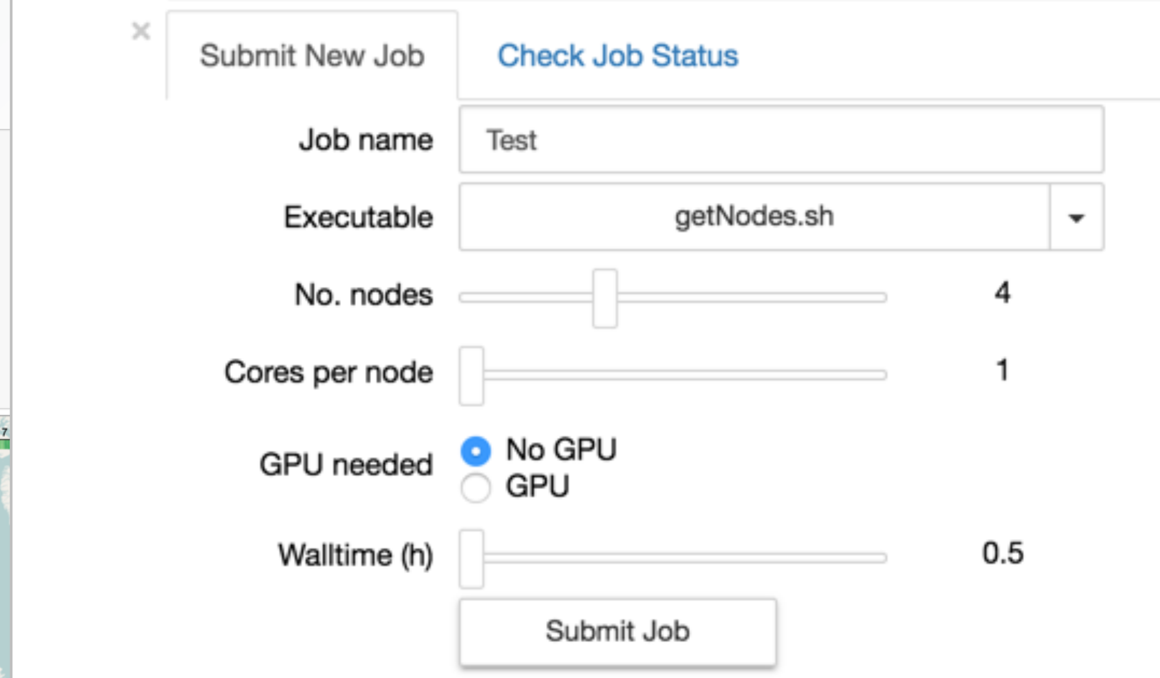
- NSF MRI for dedicated cyberGIS research
- 3rd-tier XSEDE resource at NCSA
- HPC component
  - Geospatial software environment
    - GDAL, GEOS, PROJ4, GRASS, SPATIALITE, GEOTOOLS, NETCDF/HDF (parallel)
  - 5PB storage via GPFS
- Data-intensive computing component
  - Hadoop, Spark, Accumulo, Cassandra, Redis
- Cloud component
  - OpenStack

NFIE and TauDEM on ROGER

- HUC-based computation on ROGER HPC
  - NFIE code: /projects/nfie/nfie-floodmap
  - TauDEM development version: /gpfs_scratch/taudem/
  - Software environment
    - GDAL, NetCDF4/HDF5, Python libs, GEOS, PROJ4, ESRI FileGDB, Spatialite/SQLite, MPI
- 3DEP and NHDPlus data deployment on GPFS
  - /gpfs_scratch/usgs/ned{1|3|10}m
  - /gpfs_scratch/usgs/nhd
  - /gpfs_scratch/usgs/wbd
- Data & visualization server in ROGER cloud
  - http://nfie.roger.ncsa.illinois.edu/nfiedata/
- Gateway environment
  - Jupyter interactive analysis environment
    - http://hydroroger.ncsa.illinois.edu/
  - CyberGIS Gateway
    - http://sandbox.cigi.illinois.edu
- iRODS and GridFTP
  - iRODS NCSA vault: cg-hm02.cigi.illinois.edu:/projects/nfie/irods/nfie_vault
  - External access
    - RENCI iRODS
    - TACC Wrangler

Lowering Access Barrier - JupyterHub

In [1]:
```
import roger
roger.Roger()
```



ROGER job submission cell

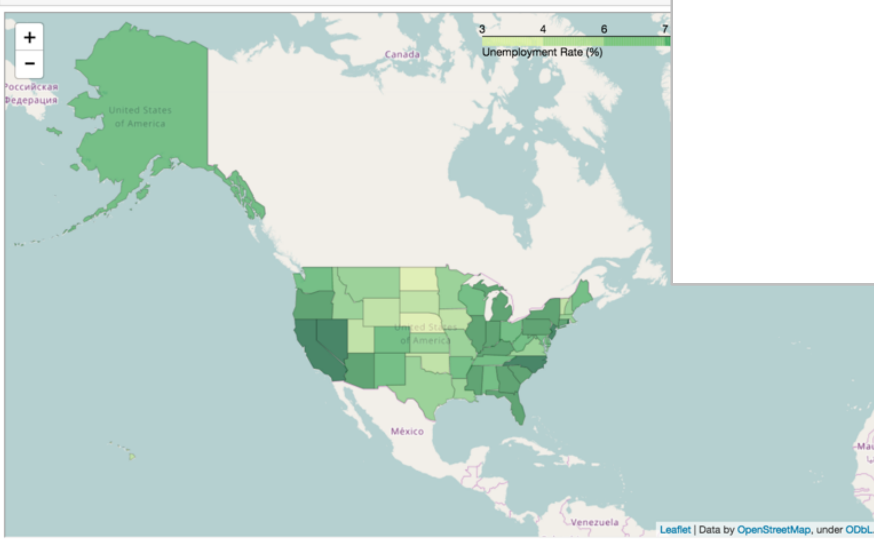
In [19]:
```
state_data = pd.read_csv(US_Unemployment_Oct2012)

m = folium.Map(location=[48, -102], zoom_start=3)
m.choropleth(
    geo_path=us_states,
    data=state_data,
    columns=['State', 'Unemployment'],
    key_on='feature.id',
    fill_color='YlGn',
    fill_opacity=0.7,
    line_opacity=0.2,
    legend_name='Unemployment Rate (%)',
    highlight=True
)

m.save(os.path.join('results', 'GeoJSON_and_choropleth_11.html'))

m
```

Out[19]:



Mapping cells using Folium

Resources

- Code:
  - https://github.com/dtarb/TauDEM
  - https://github.com/cybergis/nfie-floodmap
- Data: http://nfie.roger.ncsa.illinois.edu/nfiedata/
- Maps:
  - HAND raster: http://nfie.roger.ncsa.illinois.edu/nfiedata/maps/maps.html#source=..%2Fyanliu%2Fviz%2Fhuc6-mercator-new-singlelayer.json&extent=-14392000_2436200_-7279500_6594375
  - Inundation map – 20161208_010000: http://nfie.roger.ncsa.illinois.edu/nfiedata/maps/maps.html#source=..%2Fyanliu%2Fviz%2Finunmap-mercator-singlelayer-1color.json&extent=-14392000_2436200_-7279500_6594375

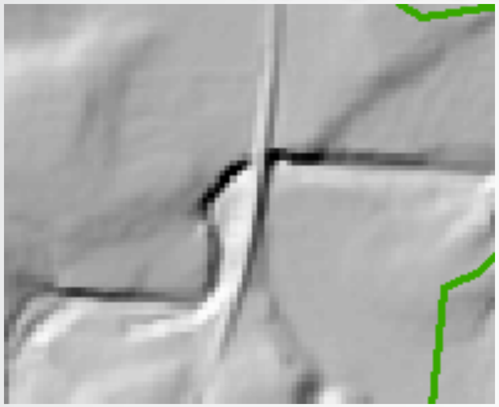
Acknowledgments

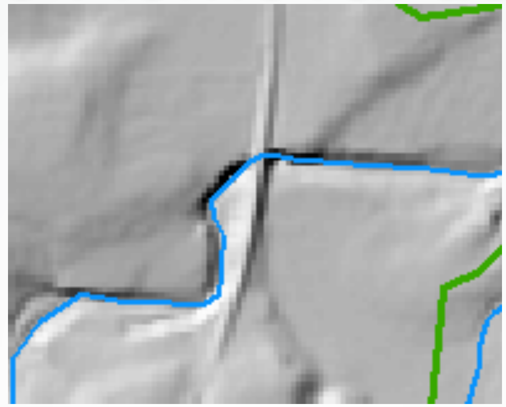
- **Federal Agencies**
  - **National Science Foundation**
    - ACI-1443080
    - ACI-1429699
    - ACI-1047916
    - 1343785
    - 1148453
    - XSEDE
  - **US Geological Survey**
    - G14AC00244
- **Government Agencies**
  - **NOAA NWS**
  - **USGS**
  - **City of Austin**
  - **Texas Division of Emergency Management**
- **Industry**
  - **Environmental Systems Research Institute (Esri)**

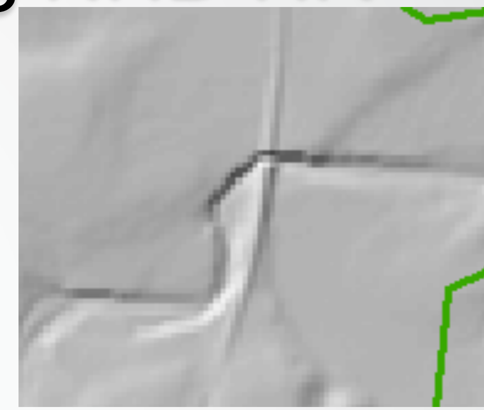
Appendix: HAND and Inundation Mapping Methodology

NHD-HAND Workflow

1. Obstacle Removal using NHD HR







48	48	49	48	47
48	46	50	46	45
47	45	51	43	39
44	40	51	47	47
41	45	52	48	48

48	48	49	48	47
48	46	50	46	45
47	45	51	43	39
44	40	51	47	47
41	45	52	48	48

48	48	49	48	47
48	46	50	46	45
47	45	40	40	39
44	40	51	47	47
41	45	52	48	48

Obstacles, like this railroad that are present in DEM are removed tracing down NHD HR streams ensuring that elevation never increases along a stream.

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

2. Remove (Fill) Sinks

Raw DEM (Obstacles removed)

36	35	36	39	42	48	50
39	32	33	35	38	46	51
47	40	36	32	38	43	46
52	48	39	36	31	35	34
52	50	47	38	34	31	30
51	50	49	45	40	35	33
52	51	52	50	47	41	37

Pit-Removed DEM (Hydrologically conditioned)

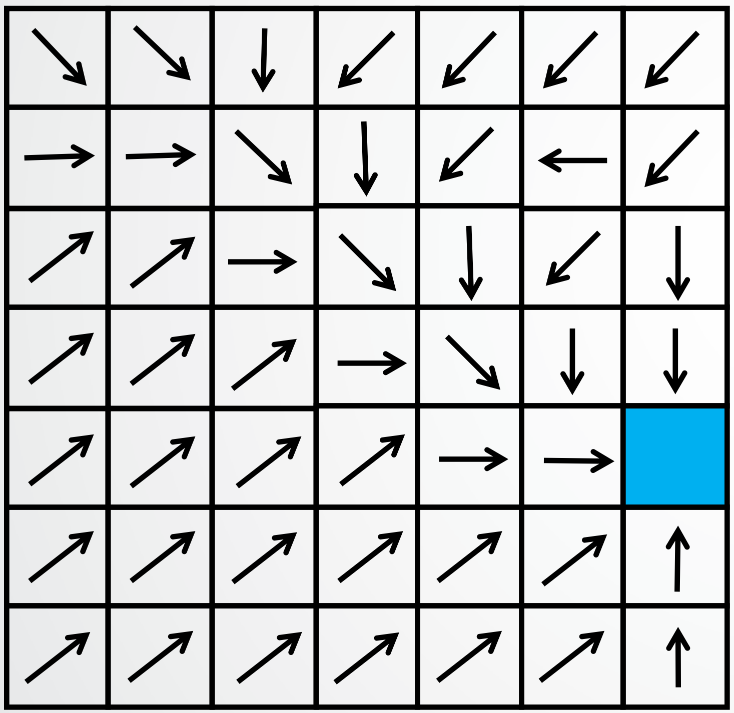
36	35	36	39	42	48	50
39	33	33	35	38	46	51
47	40	36	32	38	43	46
52	48	39	36	31	35	34
52	50	47	38	34	31	30
51	50	49	45	40	35	33
52	51	52	50	47	41	37

○ —— Sinks

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

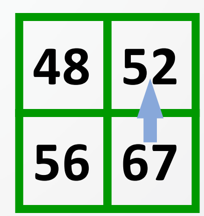
3. D8 Flow Direction

D8 Flow Direction



Flow direction from each grid cell set in the steepest downslope direction to one of eight adjacent neighbors

Routing across flats away from high terrain and towards low terrain (Garbrecht and Martz, 1997)

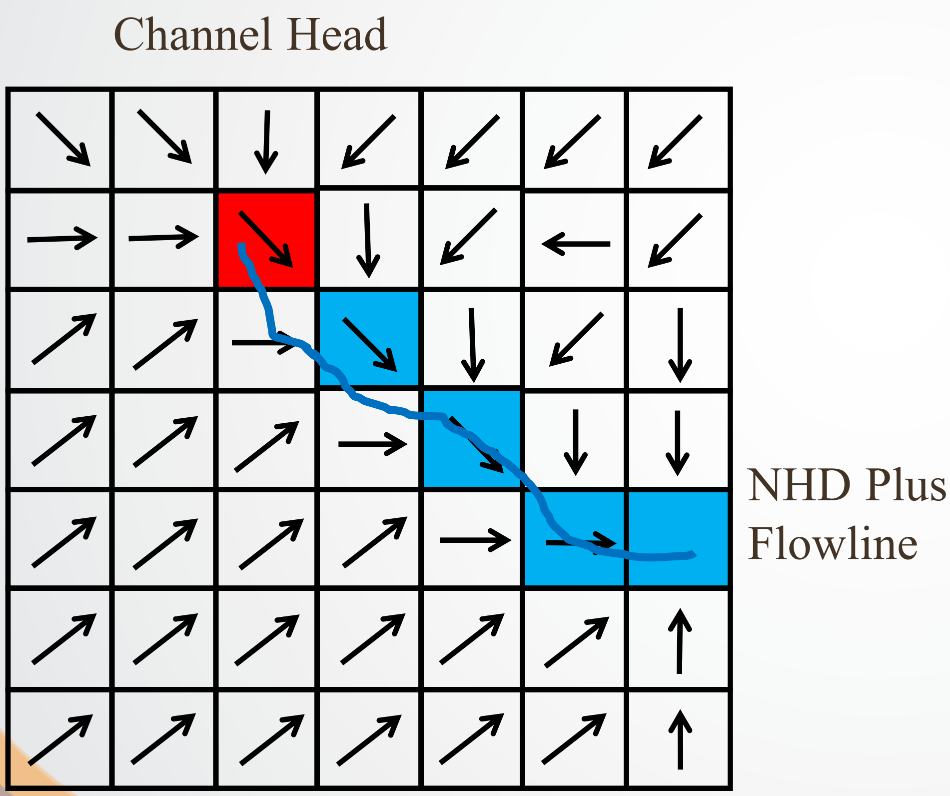


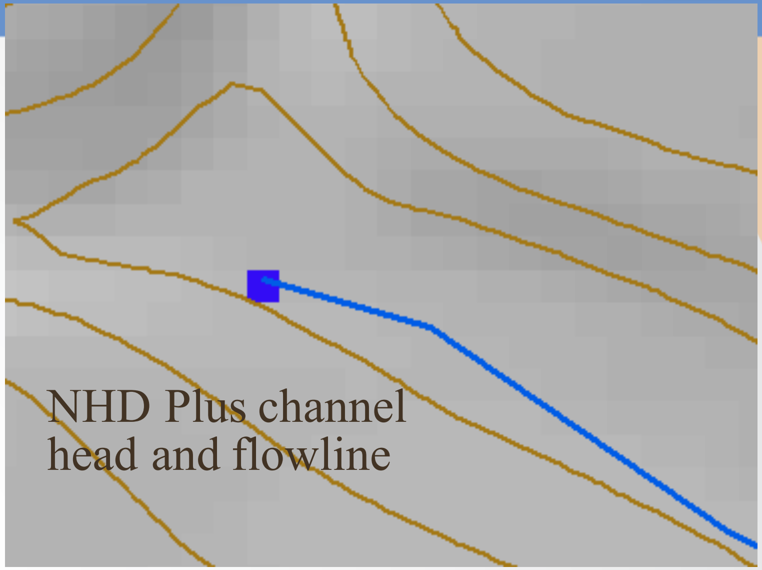
$$\frac{67-52}{30} = 0.50 \qquad \frac{67-48}{30\sqrt{2}} = 0.45$$

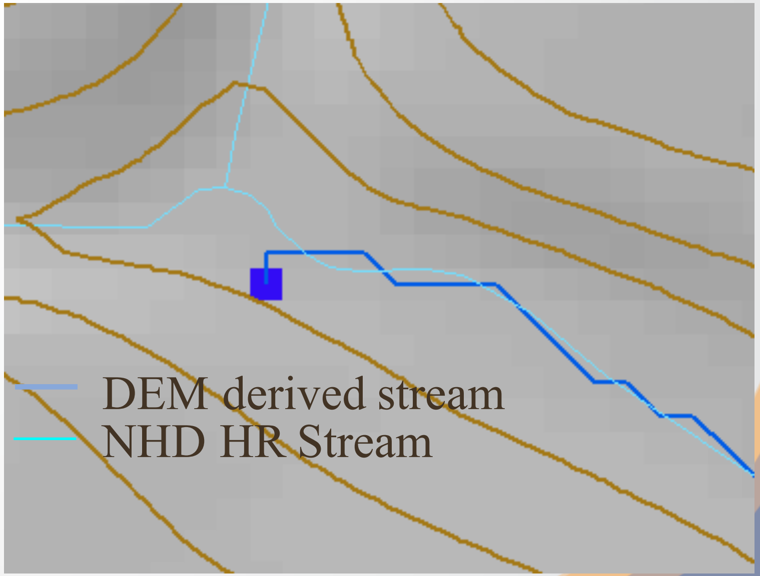
Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

4. Stream Definition

Grids from the NHD Plus channel head cell to the outlet cell are defined as Stream Cells



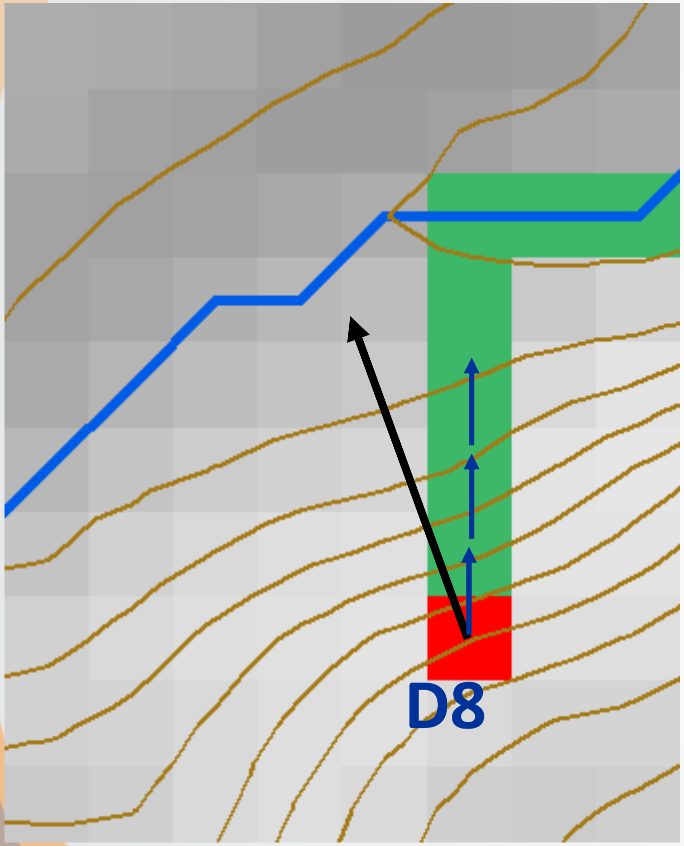


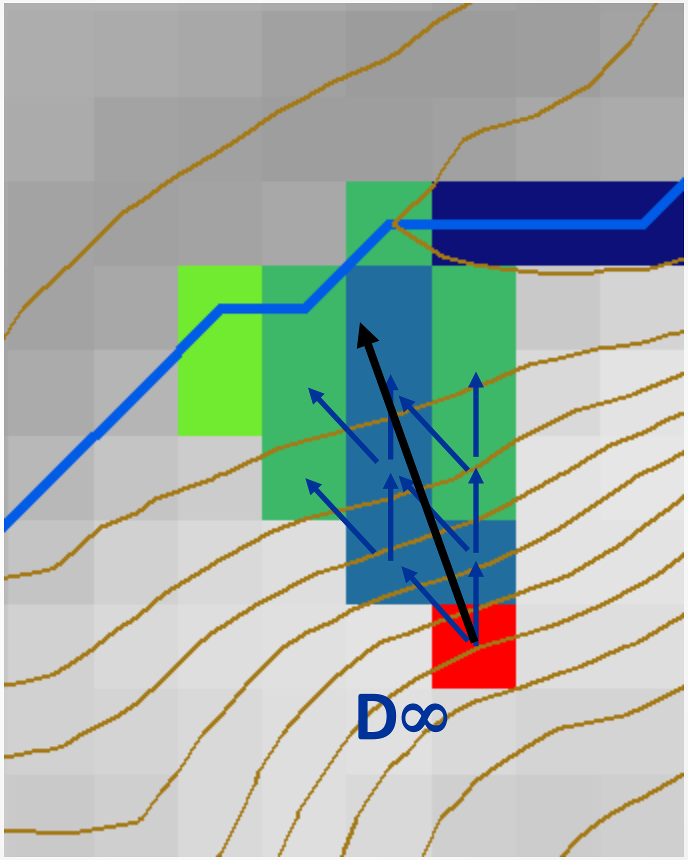


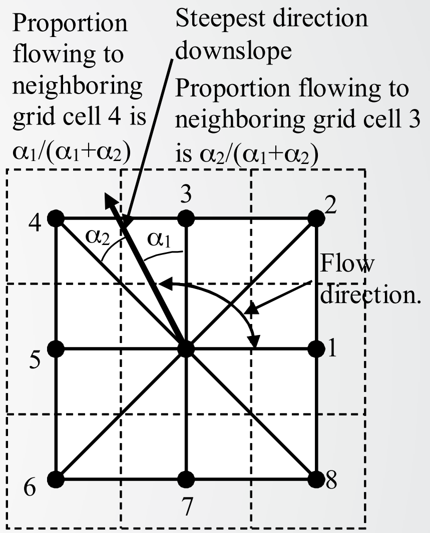
DEM derived stream network is consistent with the DEM but at the scale (drainage density) of NHD Plus.

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

5. TauDEM Dinfinity Representation of Flow Field





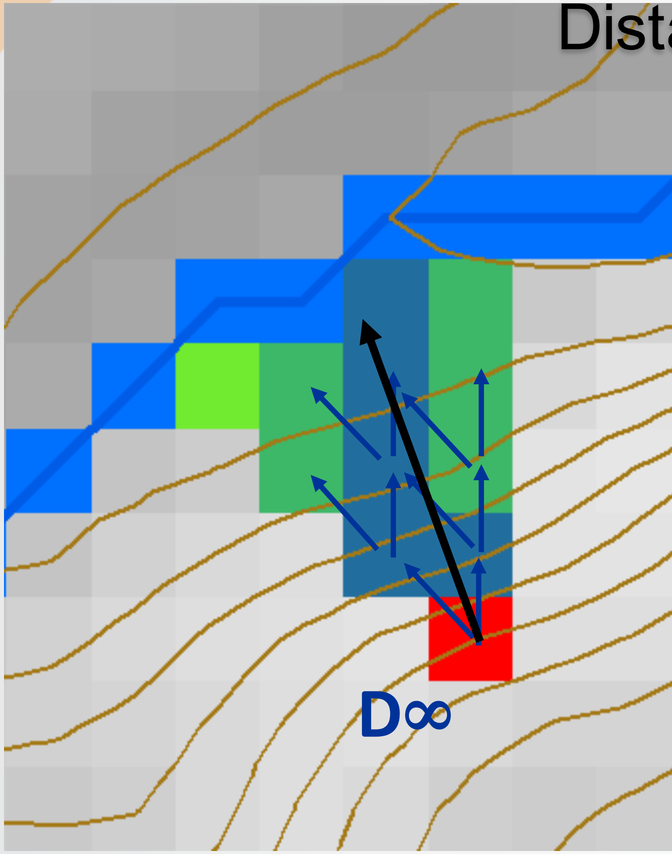


Tarboton, D. G., (1997), "A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models," Water Resources Research, 33(2): 309-319.)

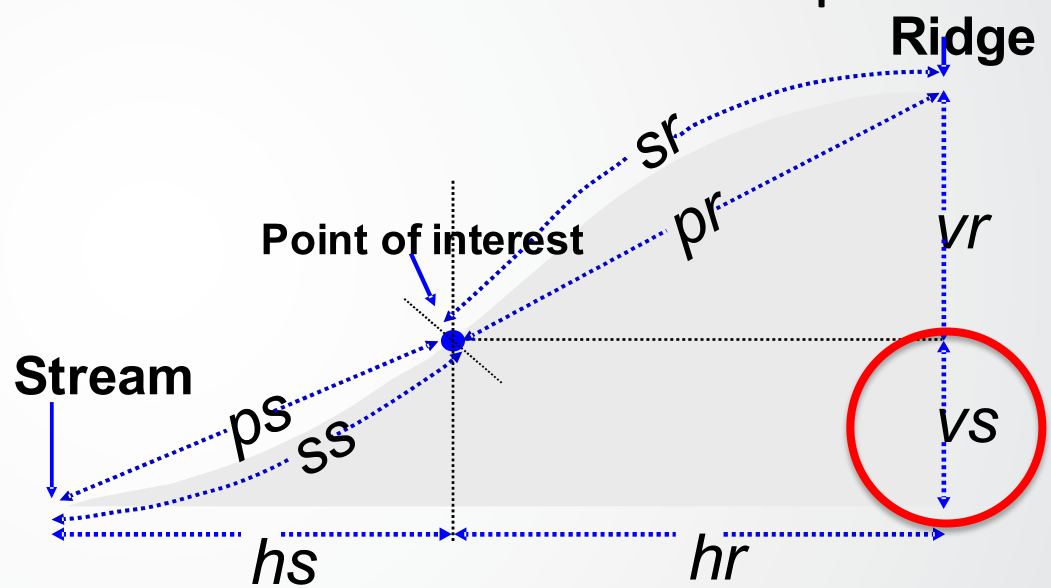
http://hydrology.usu.edu/taudem

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

6. HAND evaluated using TauDEM Dinfinity Vertical Distance Down function



TauDEM Distance Down and Distance Up



Vertical distance to stream evaluated as weighted average over multiple flow paths. This results in a "smooth" height above nearest stream layer

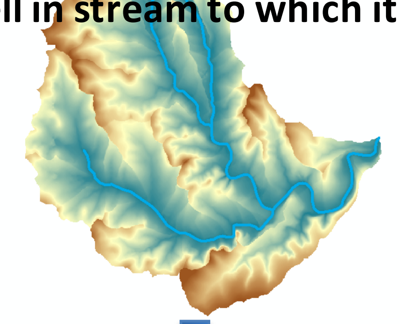
Tesfa, T. K., D. G. Tarboton, D. W. Watson, K. A. T. Schreuders, M. E. Baker and R. M. Wallace, (2011), "Extraction of hydrological proximity measures from DEMs using parallel processing," Environmental Modelling & Software, 26(12): 1696-1709, http://dx.doi.org/10.1016/j.envsoft.2011.07.018.

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

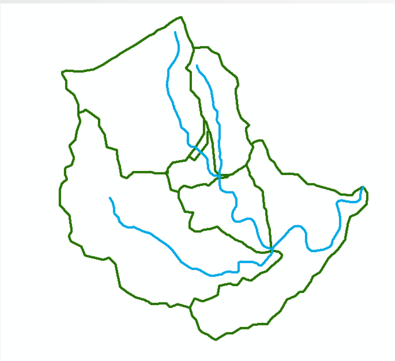
Implementation based on NHDPlus catchments and reaches

Height above stream

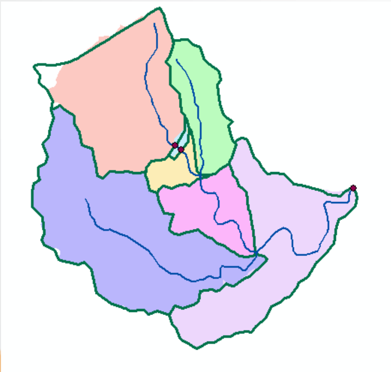
(relative elevation of land surface cell above cell in stream to which it flows)



Flowlines



Catchments



Reach Scale Flood Depth

Comid	Depth (ft)
5781365	8
5781381	9
5781405	10
5781401	15
5781399	14
5781383	12
5781933	11

Inundation map

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016

7. Mapping from height above nearest drainage to flood extent

Reach and Watershed id

$h_w(1) = 3.5$

			1						
		1	1	1	1				
	1	1	1	1	1	1			
	2	2	2	3	3	3	3		
2	2	2	2	3	3	3	3	3	
2	2	2	2	3	3	3	3	3	
2	2	2	2	3	3	3	3	3	
2	2	2	2	3	3	3	3	3	3
2	2	2	2	3	3	3	3	3	3
	2				3	3	3	3	

Height above nearest drainage raster h_{HAND}

			0						
		2	0	2	4				
	4	3	0	3	4	5			
	3	2	0	0	1	2	4		
3	1	0	1	1	0	1	2	5	
2	1	0	1	1	1	0	1	5	
2	0	1	2	2	1	0	1	4	
2	0	1	3	3	2	1	0	1	4
2	2	2	3	4	3	2	1	1	4
	3				3	2	3	4	

$h_w(2) = 2.5$

$h_w(3) = 1.5$

Source: Maidment *et. al*, USGS leadership team briefing. 07/18/2016