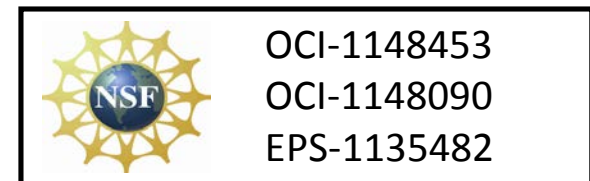


Clearing your Desk! Software and Data Services for Collaborative Web Based GIS Analysis

David Tarboton, Ray Idaszak, Jeffery Horsburgh, Dan Ames, Jon Goodall, Larry Band, Venkatesh Merwade, Alva Couch, Rick Hooper, David Maidment, Pabitra Dash, Michael Stealey, Hong Yi, Tian Gan, Tseganeh Gichamo, Ahmet Yildirim, Yan Liu

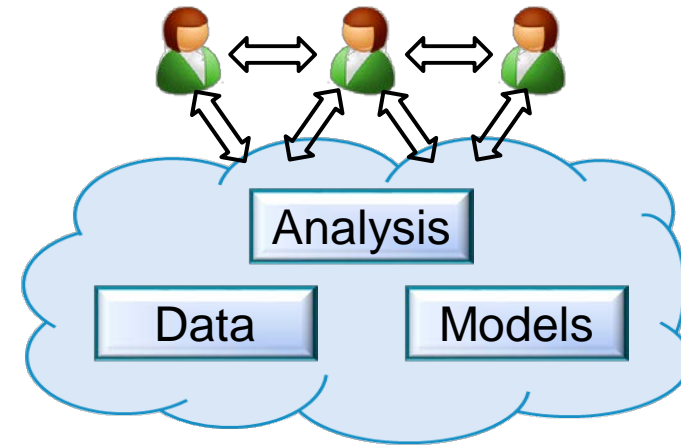
USU, RENCi, BYU, UNC, UVA, CUAHSI, Tufts, Texas, Purdue, Cactus

<http://www.hydroshare.org>



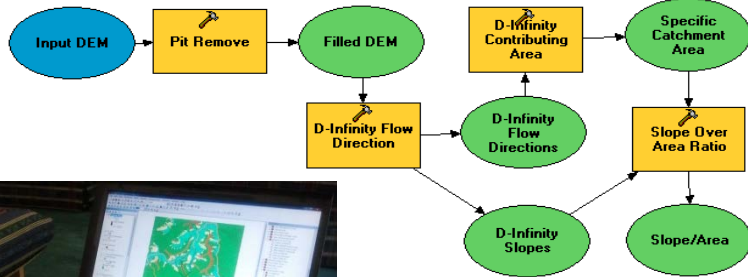
Outline

- Data and computational challenges
- HydroShare
 - Goals
 - Resource data model
 - Architecture
- Terrain analysis and TauDEM in CyberGIS
- Data services for hydrologic modeling
- Summary

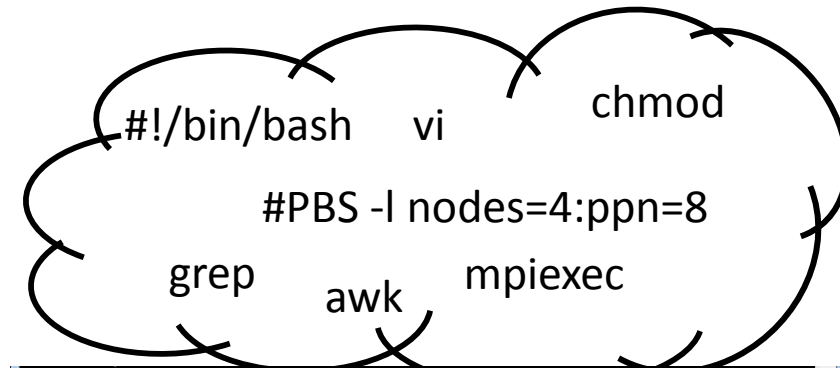


Do you have the access or know how to take advantage of advanced computing capability?

Hydrologic Experimentation and Modeling



Data Intensive High Performance Computing



```
-bash-3.2$ ls tddata
logan LoganOutlet.shn LoganOutlet.shp LoganOutlet.shx
LoganOutlet.dbf LoganOutlet.sbx LoganOutlet.shp.xml
-bash-3.2$ ls tddata/logan
logan.tif
-bash-3.2$ ls
eric logMffel run.bash taudem.bash taudem_submit.sh
logMf run_all.bash run_taudem.sh taudem_041959 tddata
-bash-3.2$ run_taudem.sh pitremove -z logan -fel loganfel
#3008.ip-net
-bash-3.2$
```

A digital divide

Data and Software Services

The challenge of increasing Digital Elevation Model (DEM) resolution

1980's DMA 90 m

10^2 cells/km²

1990's USGS DEM 30 m

10^3 cells/km²

2000's NED 10 m

10^4 cells/km²

2010's LIDAR ~1 m

10^6 cells/km²



e.g. 50,000 km²
Watershed

27 MB

240 MB

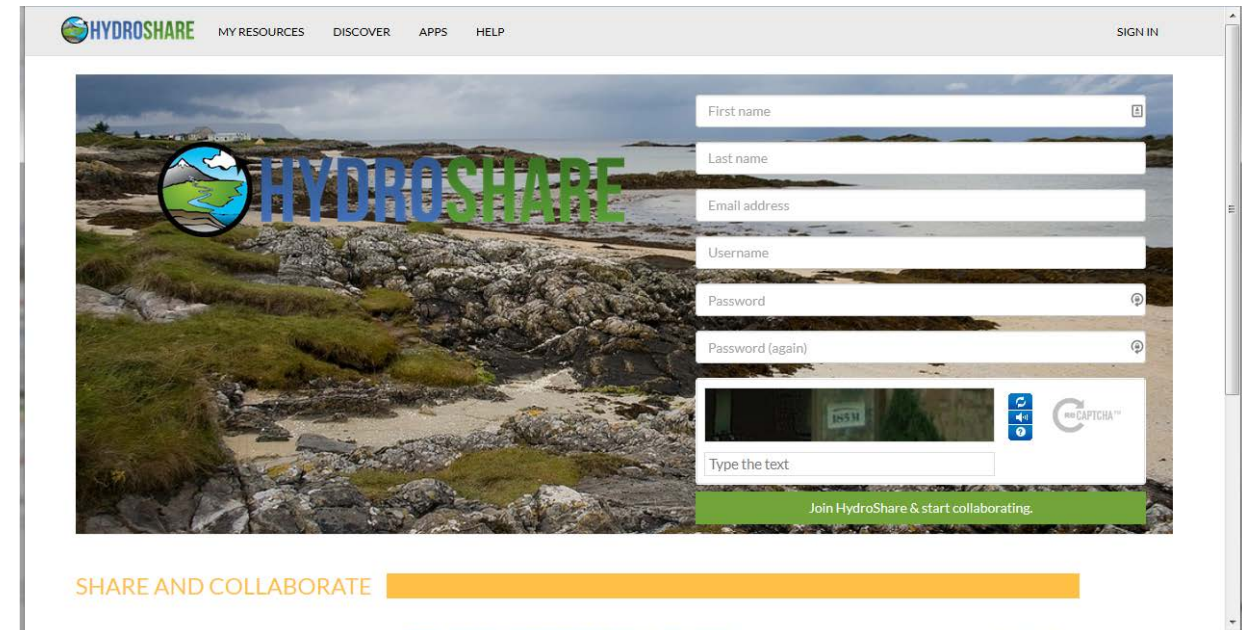
2 GB

200 GB

HydroShare web based collaboration environment


- Share your data and models with colleagues
- Manage who has access to the content that you share
- Share, access, visualize and manipulate a broad set of hydrologic data types
- Sharing and execution of models
- Web services API to facilitate automated and client access to almost all functionality
- Access to and use of high performance computing
- Publication of data and models with a DOI

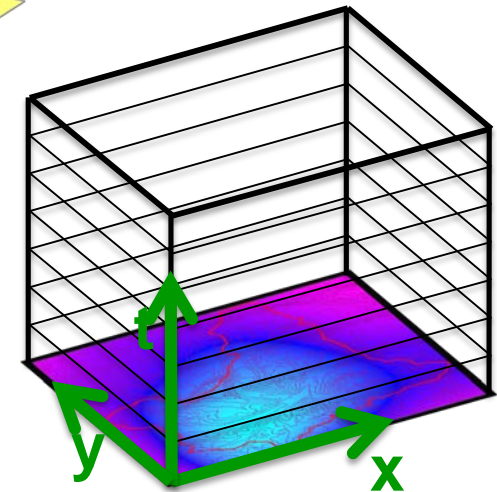
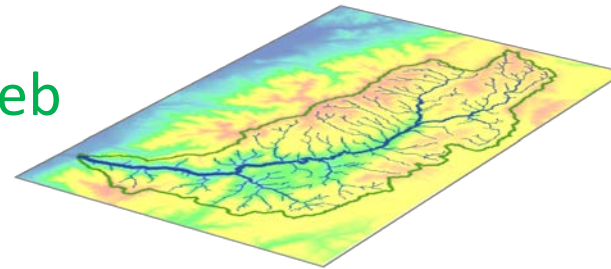
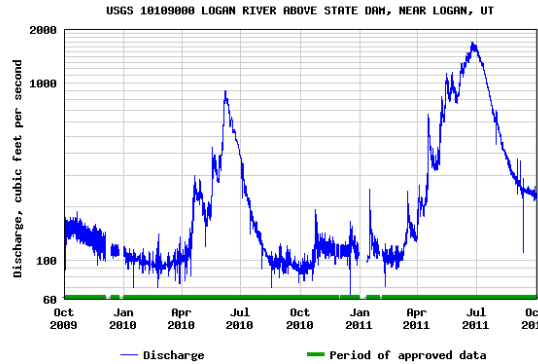
www.hydroshare.org



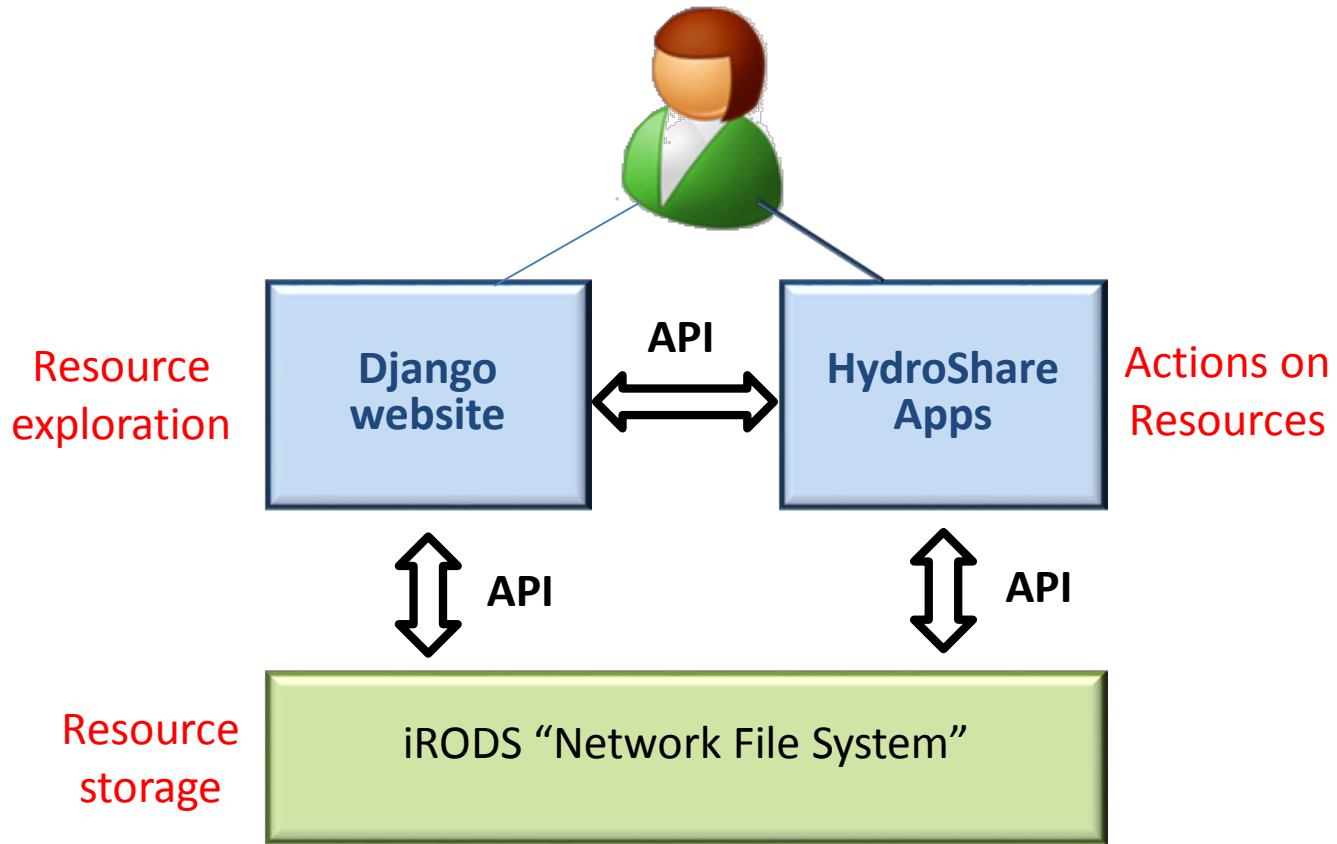
Enable more rapid advances in hydrologic understanding through collaborative data sharing, analysis and modeling.

Types of data supported as resources

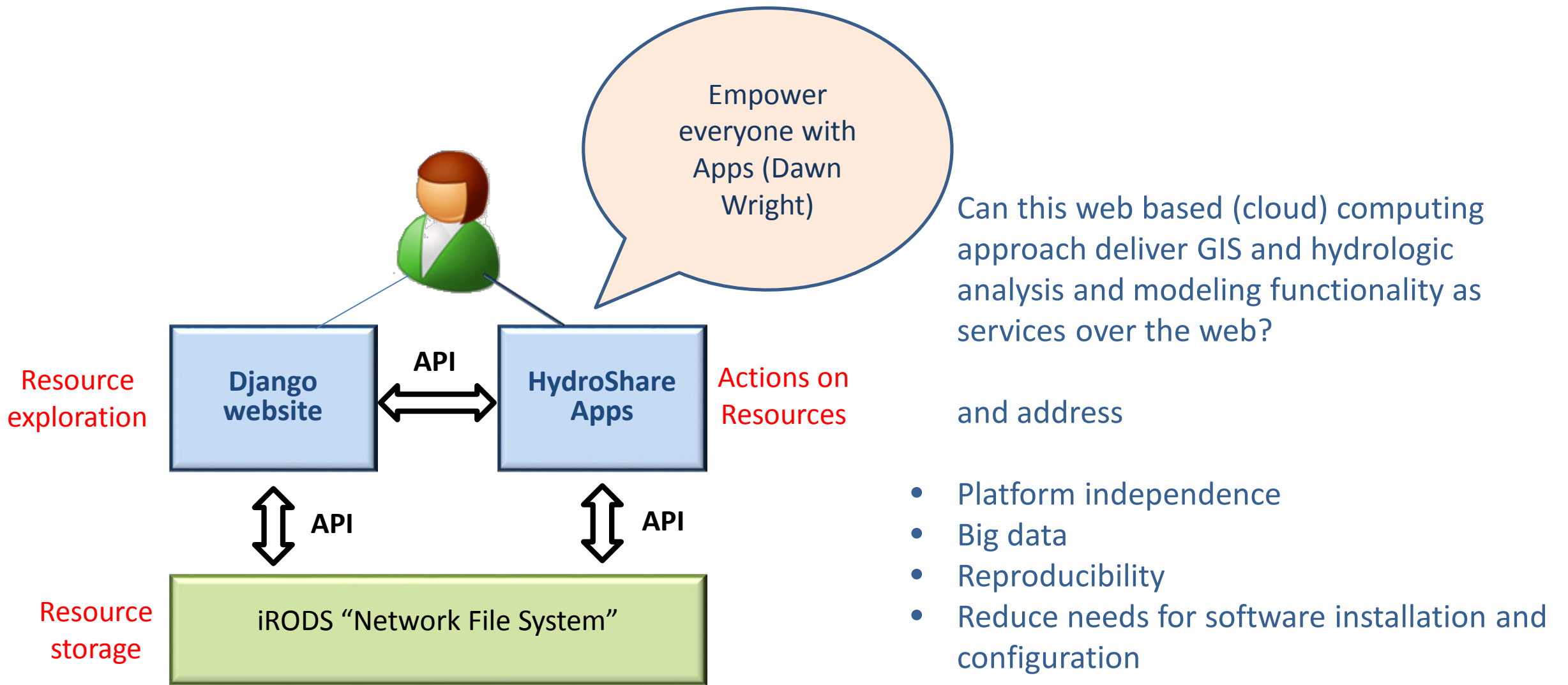
- Generic ✓
- Geographic Raster (GeoTIFF) ✓
- Time Series ✓
- Multidimensional (netCDF) ✓
- Model program ✓
- Model instance ✓
 - SWAT Model Instance ✓
- Web App ✓
- Geographic Feature (Shapefile) ✓
- Referenced Time Series (CUAHSI HIS web service link) 
- River Geometry
- Sample based observations (ODM2 and CZO)
- Composite resources (Collections of resources)



Architecture



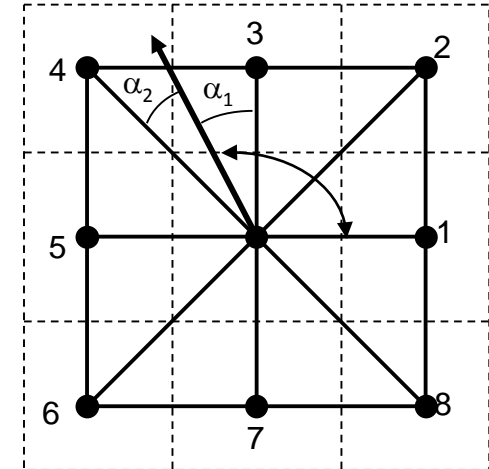
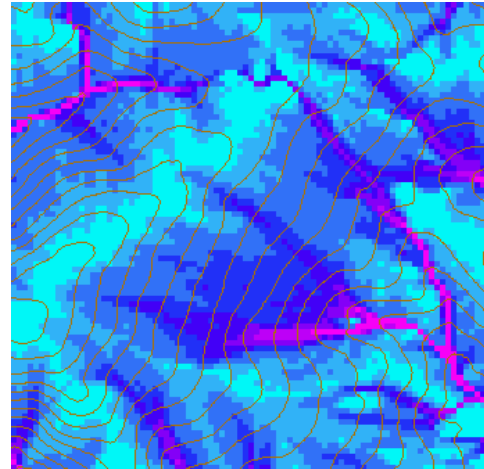
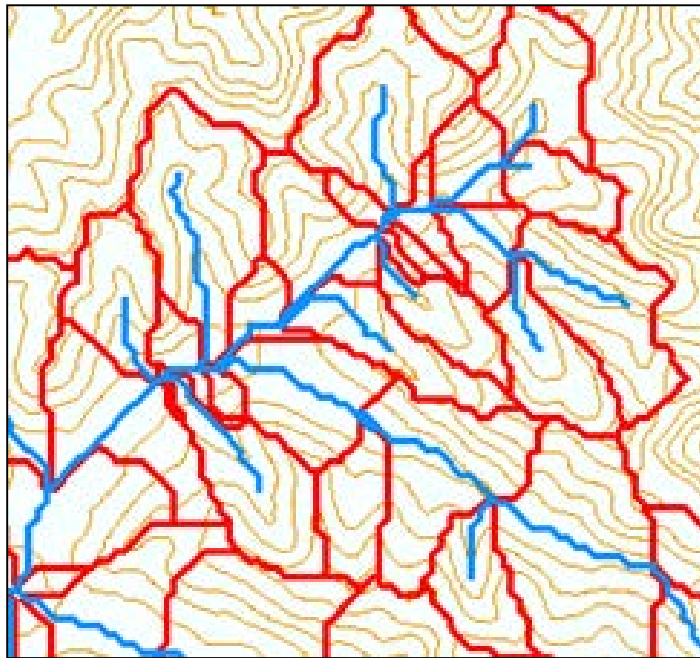
- Can this web based (cloud) computing approach deliver GIS and hydrologic analysis and modeling functionality as services over the web?
and address
- Platform independence
- Big data
- Reproducibility
- Reduce needs for software installation and configuration



Anyone can set up a server/app platform (software service) to operate on HydroShare resources through iRODS and API

Terrain Analysis as a use case



- Topography is fundamental to hydrology
- Watersheds are the most basic hydrologic landscape elements



TauDEM <http://hydrology.usu.edu/taudem/> is software for deriving hydrologically useful information from Digital Elevation Models

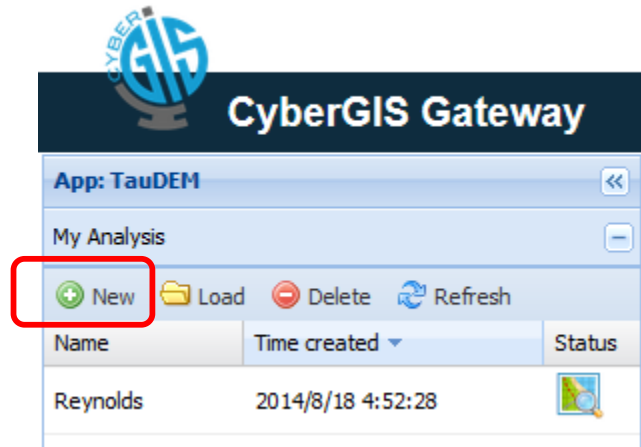
- Stream and watershed delineation
- Multiple flow direction flow field
- Calculation of flow based derivative surfaces
- MPI Parallel Implementation for speed up and large problems
- Open source platform independent C++ command line executables for each function
- Deployed as an ArcGIS Toolbox with python scripts that drive command line executables

Using TauDEM today requires


- Expertise in Hydrologic DEM analysis
- The software
 - ArcGIS licenses (for ArcGIS plugin)
 - The ability to install software
 - TauDEM command functions with MPI installation
 - Compilation for other platforms  
- Sufficient Hardware (RAM and Disk)
- The data (GDAL formatted rasters with consistent grid size and spatial reference)

Moving TauDEM to the cloud, CyberGIS TauDEM App

<http://gateway.cigi.illinois.edu/>

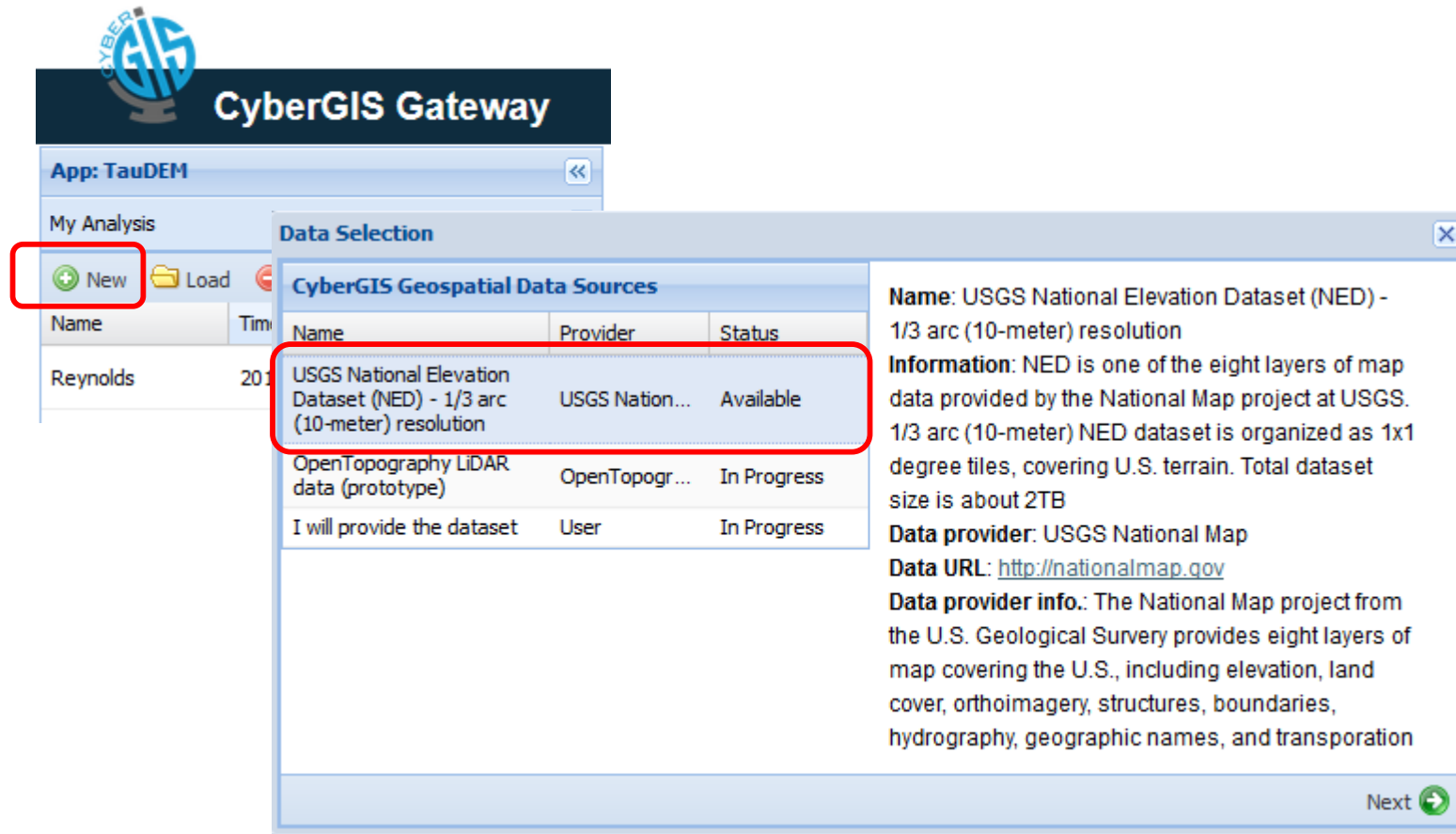


The screenshot displays the CyberGIS Gateway interface. At the top left is the CyberGIS logo, and the title "CyberGIS Gateway" is centered. Below the title, the current application is identified as "App: TauDEM". A section titled "My Analysis" contains a toolbar with four icons: a green plus sign labeled "New" (highlighted with a red box), a folder icon labeled "Load", a red minus sign labeled "Delete", and a circular arrow icon labeled "Refresh". Below the toolbar is a table with three columns: "Name", "Time created", and "Status". The table contains one row with the name "Reynolds", the creation time "2014/8/18 4:52:28", and a small map icon in the status column.

Name	Time created	Status
Reynolds	2014/8/18 4:52:28	

Moving TauDEM to the cloud, CyberGIS TauDEM App

<http://gateway.cigi.illinois.edu/>



The screenshot displays the CyberGIS Gateway interface. At the top, the logo and text "CyberGIS Gateway" are visible. Below this, the application is identified as "App: TauDEM". A "My Analysis" section contains a table with one entry: "Reynolds" with a time of "20". A red box highlights the "+ New" button in this section. A "Data Selection" dialog box is open, showing a table of "CyberGIS Geospatial Data Sources". A red box highlights the first row of this table, which is "USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution". To the right of the table, detailed information for this dataset is provided, including its name, information about the dataset, the data provider (USGS National Map), and the data URL (<http://nationalmap.gov>). A "Next" button with a right arrow is located at the bottom right of the dialog box.

CyberGIS Gateway

App: TauDEM

My Analysis

Name	Time
Reynolds	20

Data Selection

Name	Provider	Status
USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution	USGS Nation...	Available
OpenTopography LiDAR data (prototype)	OpenTopogr...	In Progress
I will provide the dataset	User	In Progress

Name: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution

Information: NED is one of the eight layers of map data provided by the National Map project at USGS. 1/3 arc (10-meter) NED dataset is organized as 1x1 degree tiles, covering U.S. terrain. Total dataset size is about 2TB


Data provider: USGS National Map

Data URL: <http://nationalmap.gov>

Data provider info.: The National Map project from the U.S. Geological Survey provides eight layers of map covering the U.S., including elevation, land cover, orthoimagery, structures, boundaries, hydrography, geographic names, and transportation

Next

File Edit View History
 CyberGIS: high-perform <http://cybergis.cigi.uiuc.edu/>
 gateway.cigi.illinois.edu/home/apps.php?app=taudem

 CyberGIS Sandbox

dtarb's Profile | Logout


CyberGIS Gateway Home Apps Visualization Community Help

App: TauDEM <<
 My Analysis: Reynolds +
 Data and Parameters -

Data Source Study Area Workflow



Data Source: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution
Provider: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution
Coverage:
 Lower-left: [-125.001, 23.999]
 Upper-right: [-65.999, 50.001]
Coverage in Native Projection:
 Lower-left: [-125.001, 23.999]
 Upper-right: [-65.999, 50.001]
Coordinate System: [EPSG: 4269](#)
Vertical Unit: m

Results +



1000 km
500 mi

LonLat: -47.60166, 56.66228

 Copyright © 2010-2014 CIGI. CyberGIS Gateway, powered by GISolve.
 CyberGIS Gateway is based upon work supported in part by the National Science Foundation under Grant Numbers: 0848655 and 1047916. Any opinions, findings, and conclusions or recommendations expressed in the Gateway are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. 

File Edit View History
CyberGIS: high-perform <http://cybergis.cigi.uiuc.edu/>
gateway.cigi.illinois.edu/home/apps.php?app=taudem

CyberGIS Sandbox

dtarb's Profile | Logout

CyberGIS Gateway

Home | Apps | Visualization | Community | Help

App: TauDEM

My Analysis: Reynolds

Data and Parameters

Data Source Study Area Workflow

Data Source: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution
Provider: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution
Coverage:
Lower-left: [-125.001, 23.999]
Upper-right: [-65.999,50.001]
Coverage in Native Projection:
Lower-left: [-125.001, 23.999]
Upper-right: [-65.999,50.001]
Coordinate System: [EPSG: 4269](#)
Vertical Unit: m

Results

CyberGIS 0848655

App: TauDEM

My Analysis: Logan River

Data and Parameters

Data Source Study Area Workflow

Data Source: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution
Provider: USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution
Coverage:
Lower-left: [-125.001, 23.999]
Upper-right: [-65.999,50.001]
Coverage in Native Projection:
Lower-left: [-125.001, 23.999]
Upper-right: [-65.999,50.001]
Coordinate System: [EPSG: 4269](#)
Vertical Unit: m

Select the products you want

TauDEM Workflow Wizard

Select Products

Filter

<input type="checkbox"/>	Name	RID
Common TauDEM Products		
<input type="checkbox"/>	Hydrologically Conditioned Elevation Grid	1
<input type="checkbox"/>	D8 Flow Direction	3
<input type="checkbox"/>	D8 Slope	2
<input type="checkbox"/>	D8 Contributing Area	6
<input type="checkbox"/>	Dinfinity Flow Direction	5
<input type="checkbox"/>	Dinfinity Slope	4
<input checked="" type="checkbox"/>	Dinfinity Specific Catchment Area	7
<input type="checkbox"/>	Contributing Area Stream Raster	14
<input checked="" type="checkbox"/>	Peucker Douglas Stream Raster	15
<input checked="" type="checkbox"/>	Stream Network And Subwatersheds	21
<input type="checkbox"/>	Gage Subwatersheds	27
<input type="checkbox"/>	Topographic Wetness Index	22
Specialized TauDEM Products		
<input type="checkbox"/>	Grid Strahler Order	8
<input type="checkbox"/>	Grid Path Length	9
<input type="checkbox"/>	Grid Total Length	10
<input type="checkbox"/>	D8 Flow Accumulation Options	11
<input type="checkbox"/>	Dinfinity Flow Accumulation Options	12

The wizard configures the sequence of functions to run to get the result

Results displayed in browser

The screenshot shows a web browser window with the URL `gateway.cigi.illinois.edu/home/apps.php?app=taudem`. The page title is "CyberGIS Gateway". The user is logged in as "dtarb's Profile" and can click "Logout". The main navigation menu includes "Home", "Apps", "Visualization", "Community", and "Help".

The left sidebar shows the application "App: TauDEM" and a list of results. The "Results" section is expanded to show a legend for "Dinfinity Contributing Area (metre)".

Color	Value Range (metre)
Lightest Blue	88.8441860009 - 266.532558003
Light Blue	266.532558003 - 1599.19534802
Medium Blue	1599.19534802 - 14392.7581321
Dark Blue	14392.7581321 - 172713.097586
Very Dark Blue	172713.097586 - 2590696.46379
Black	2590696.46379 - 46632536.3481
Dark Purple	> 2590696.46379

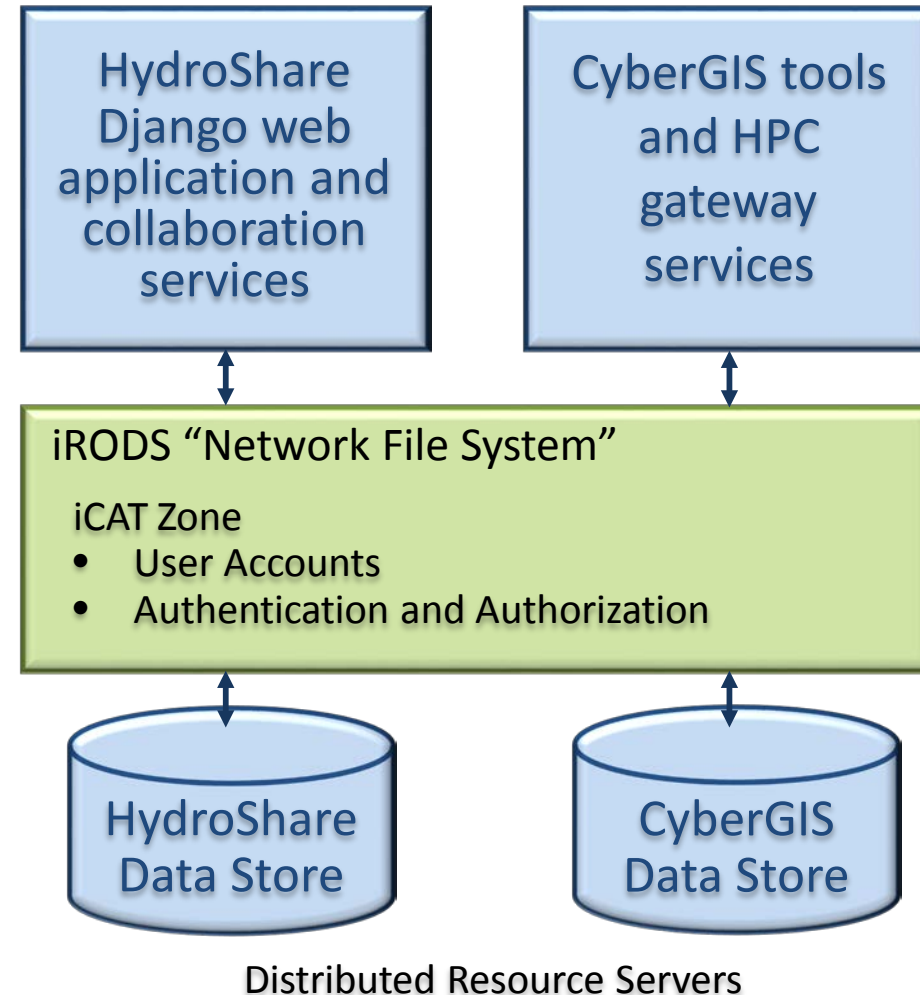
The main map area displays a stream network with a color scale from light blue to black. A scale bar at the bottom indicates 1000 m and 2000 ft. The coordinates are LonLat: -111.75596, 41.91792.

And saved in HydroShare (in progress)

Copyright © 2010-2014 CIGI. CyberGIS Gateway, powered by GISolve. CyberGIS Gateway is based upon work supported in part by the National Science Foundation under Grant Numbers: 0846655 and 1047916. Any opinions, findings, and conclusions or recommendations expressed in the Gateway are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

CyberGIS HydroShare Coupling

- CyberGIS data holdings shared as HydroShare resources
- HydroShare user resources staged at CyberGIS for input to computation
- Results persisted as resources in HydroShare to support
 - Collaboration
 - Input for further analysis and models



Advancing Data Services for Modeling and Analysis

Assumptions

1. GIS and hydrologic modelers have to learn and become comfortable using a modern scientific programming language (e.g. Python or R)
2. Modeling is increasingly data intensive (large datasets from a range of sources) demanding more data and computing resources than is in most PC's
3. Reproducibly installing and configuring models on different platforms is a challenge
4. Should not have to become expert in HPC systems and learning them is a barrier to using HPC and research with big data and computationally intensive models

Computation via Python Client calling Data and Modeling Services

Input

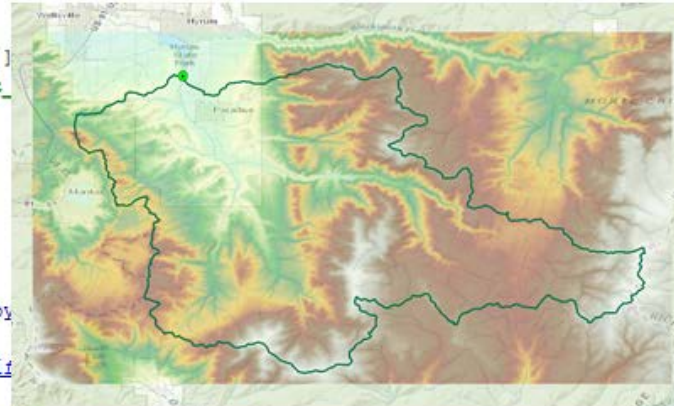
```
1  __author__ =
2  import ...
5  """ Illustration of Watershed Delineation using CI-WATER data services """
6  HDS = HydroDS(username=settings.USER_NAME, password=settings.PASSWORD)
7  subsetDEM = HDS.subset_raster(input_raster='nedWesternUS.tif', left=-111.97, top=41.629,
8                               right=-111.48, bottom=41.36, output_raster='MyDEM.tif')
9  projectDEM = HDS.project_resample_raster(input_raster_url_path=subsetDEM['output_raster'],
10                                           cell_size_dx=30, cell_size_dy=30, epsg_code=26912,
11                                           output_raster='MyDemProj.tif', resample='bilinear')
12
13 # Create outlet point
14 outlet = HDS.create_outlet_shapefile(point_x=-111.855, point_y=41.596,
15                                     output_shape_file_name='Outlet.shp')
16 outletProj = HDS.project_shapefile(outlet['output_shape_file_name'], 'OutletProj.shp',
17                                   epsg_code=26912)
18
19 ## Delineate watershed
20 Watershed = HDS.delineate_watershed(projectDEM['output_raster']
21                                   input_outlet_shapefile_url_path=outletProj['output
22                                   threshold=1000, epsg_code=26912,
23                                   output_raster='Watershed.tif',
24                                   output_outlet_shapefile='movedOutlet.shp')
25 ListMyFiles()
```

Python session on desktop but data and analysis on server with results pushed to iRODS/HydroShare for storage and collaboration

Result

```
C:\Python27\ArcGIS10.3\python.exe D:/Scratch/CI-WATERDemos/demo.py
http://hydro-ds.uwrl.usu.edu:20199/files/data/user\_4/MyDEM.tif
http://hydro-ds.uwrl.usu.edu:20199/files/data/user\_4/MyDemProj.tif
http://hydro-ds.uwrl.usu.edu:20199/files/data/user\_4/Outlet.zip
http://hydro-ds.uwrl.usu.edu:20199/files/data/user\_4/OutletProj.zip
http://hydro-ds.uwrl.usu.edu:20199/files/data/user\_4/movedOutlet.zip
http://hydro-ds.uwrl.usu.edu:20199/files/data/user\_4/Watershed.tif
```

Process finished with exit code 0



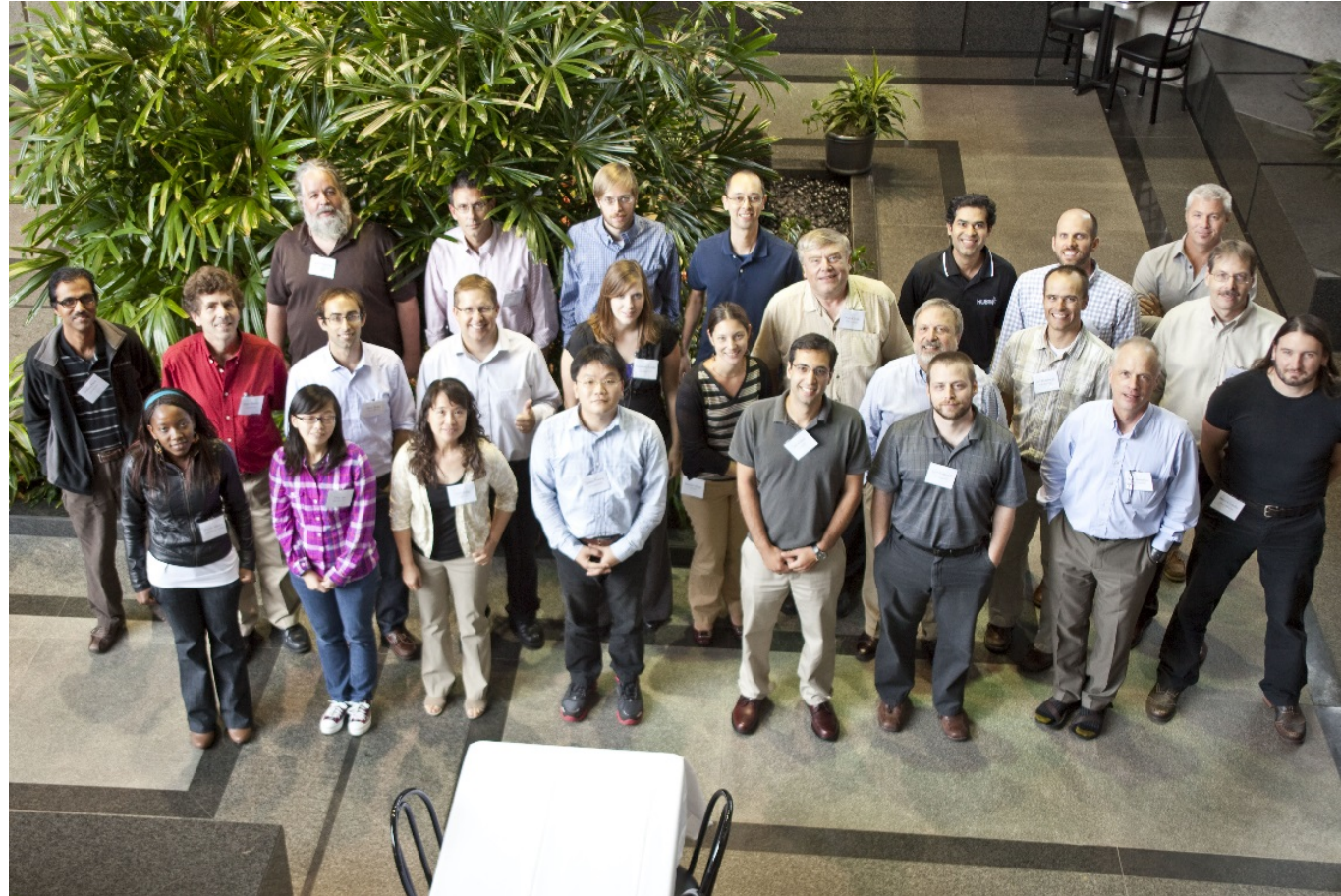
Details in <https://www.hydroshare.org/resource/cfb8d71b7f1f4e75a44f5e634f4730d4/> or search for HydroGate in HydroShare

Summary

1. Web based Cyberinfrastructure for GIS and Hydrologic Data and Modeling is emerging to support
 - Large datasets
 - Collaboration
 - Reproducible workflows and results
 - And reduce software installation and configuration limitations.
2. HydroShare and CyberGIS are part of this.
 - Interoperability is key to leveraging full potential of multiple emerging cloud cyberinfrastructure systems

Thanks to the HydroShare, CyberGIS and CI-WATER teams!

- USU
- RENCI/UNC
- CUAHSI
- BYU
- Tufts
- UVA
- Texas
- Purdue
- SDSC



<http://www.hydroshare.org>

