



A Utah-Wyoming Cyberinfrastructure
Water Modeling Collaboration

Enhance access to data- and computationally- intensive modeling.

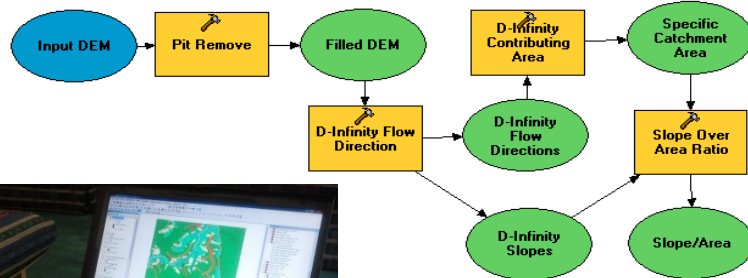
Team 2 – David Tarboton
November 2014



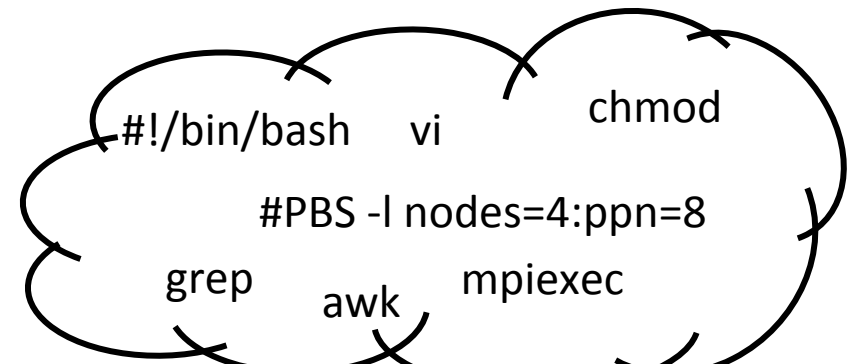
A digital divide

Researchers

- Experimentalists
- Modelers



Big Data and HPC



```
-bash-3.2$ ls tddata
logan      LoganOutlet.sbn  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
logMffel  run_all.bash  run_taudem.sh  taudem_041959  tddata
-bash-3.2$ run_taudem.sh pitremove -z logan -fel loganfel
43058.1b-nec
-bash-3.2$
```

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

Gateways, Web Interfaces, Software services

Goals

1. “Provide hydrologic researchers, modelers, water managers, and users access to HPC resources without requiring them to become HPC and CI experts”
2. “Reduce the amount of time and effort spent in finding and organizing the data required to execute the models”

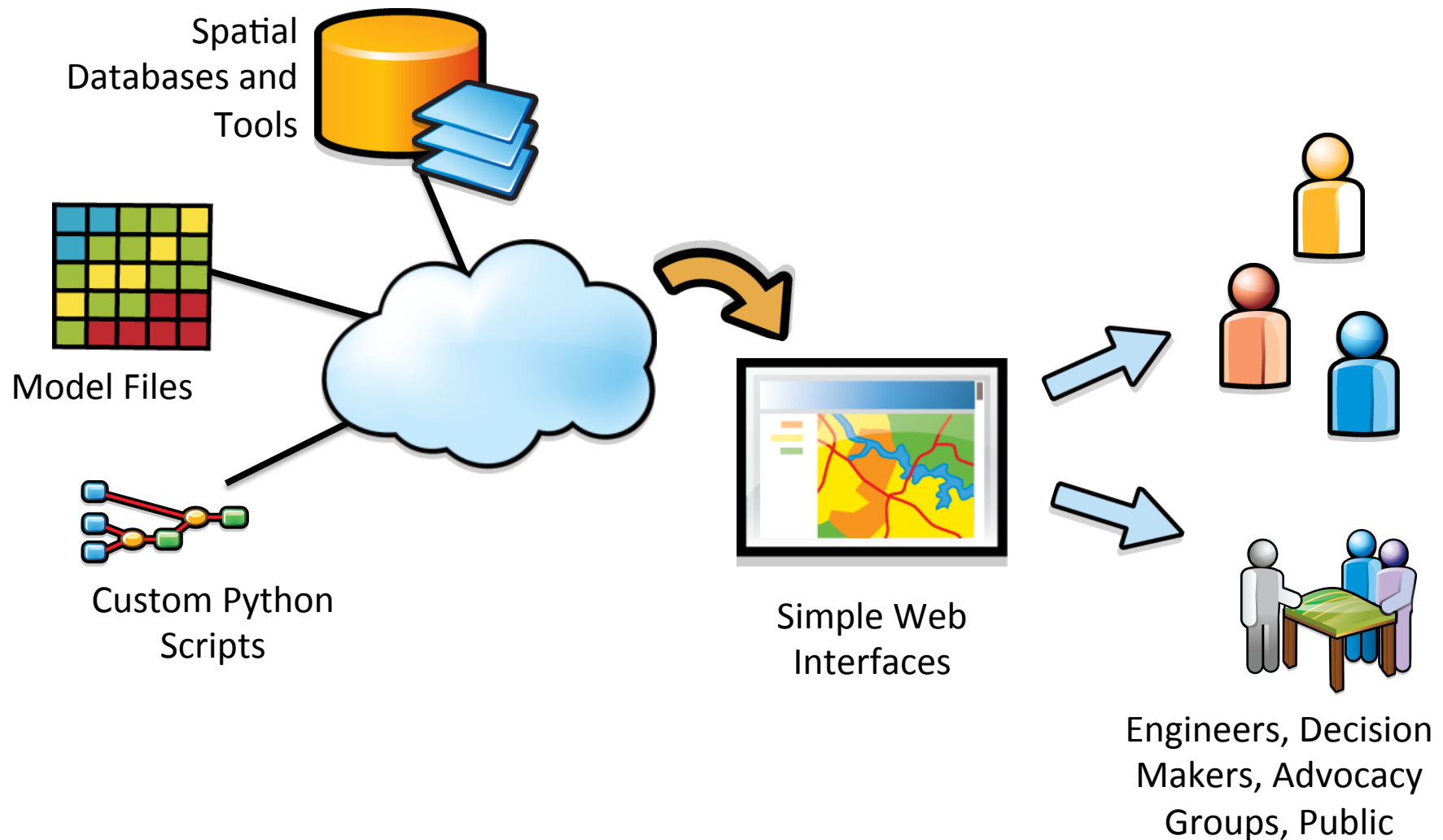
Overview

- Data access services for modeling (USU)
- Python Client for HPC access via web services gateway (USU)
- Climate and Urban Water Systems (UU)
- Toolkit for cloud based water resources modeling (BYU)

Proposal Timeline and Milestones

CI-WATER Milestones and Timeline	Year 1				Year 2				Year 3			
(assumes 9/1/11 start)	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
Component 2. Enhance Access to Data- and Computationally-Intensive Modeling												
Water management components for data services												
Develop data model design	█	█										
Implement prototype data services			█	█								
Test and refine prototype data services					█	█						
Implement operational data services							█	█				
Urban data components for data services												
Develop data model design	█	█										
Implement prototype data services			█	█								
Test and refine prototype data services					█	█						
Implement operational data services							█	█				
Develop and deploy data services for datasets required by models	█	█	█	█	█	█	█	█	█	█	█	█
Develop a HubZero instance that interfaces with Utah/Wyoming HPC resources	█	█	█	█								
Develop HubZero functionality			█	█	█	█						
Develop appropriate user interfaces within HubZero to enable user access to models					█	█	█	█				
Develop model output post processing and visualization tools							█	█	█	█		
Develop community model collaboration capabilities									█	█	█	█

Cloud-Based Modeling for Decision Support



USU Team

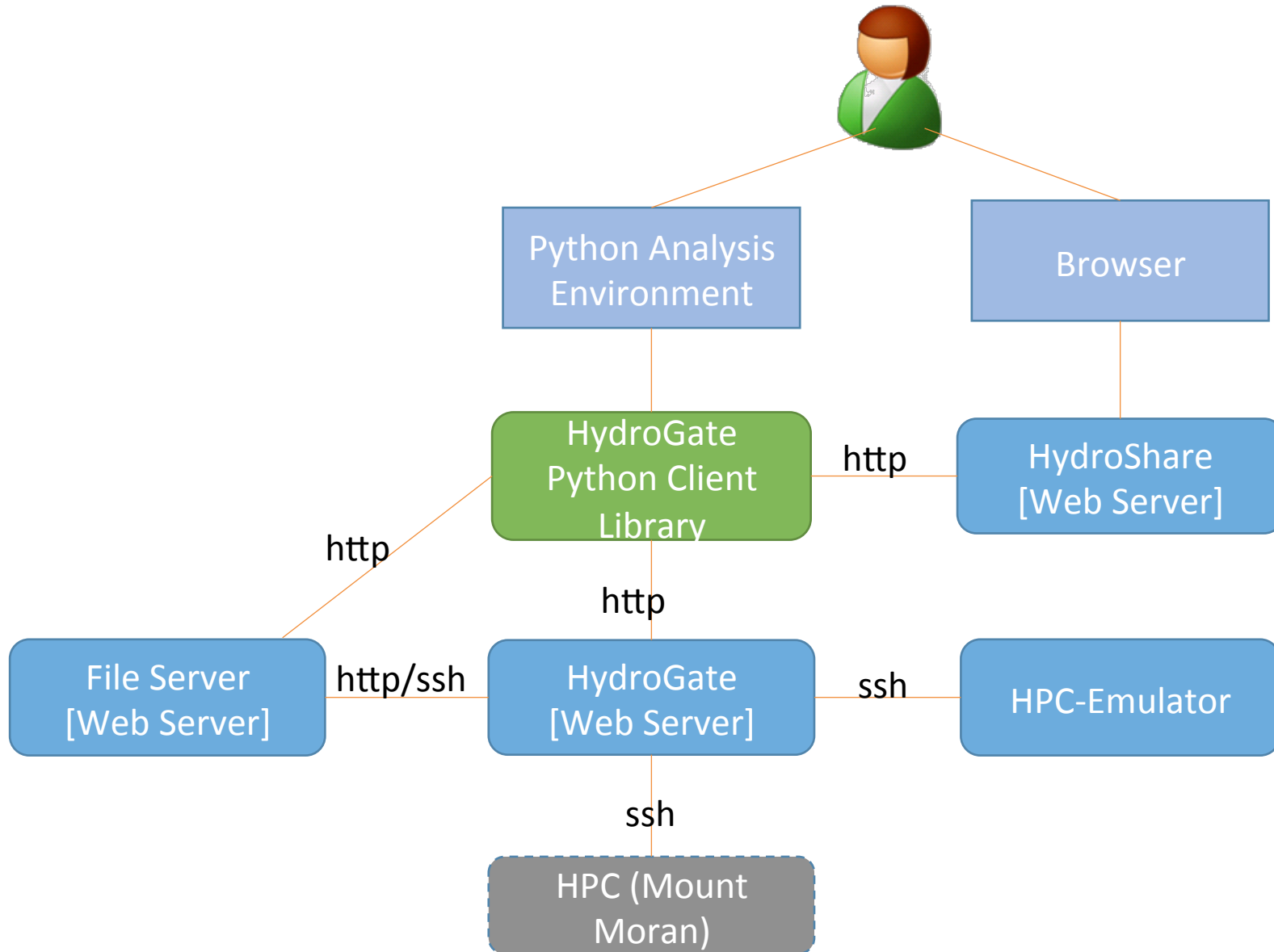
- David Tarboton, Jeff Horsburgh, David Rosenberg (co-PI's)
- Pabitra Dash (software engineer)
- Tseganeh Gichamo (CEE PhD student in hydrologic modeling)
- Ahmet Yildirim (Computer Science PhD student in parallel programming and gateways)
- Adel Abdallah (CEE PhD student in Water Resources Management)



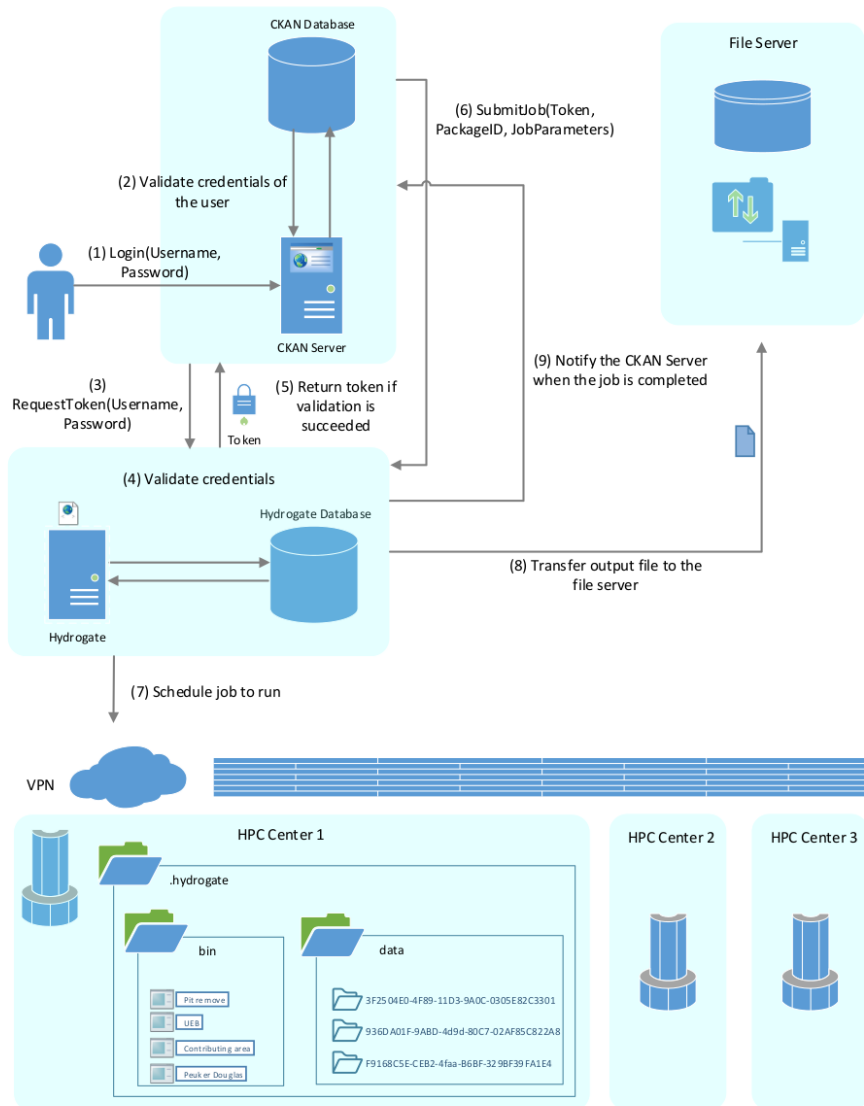
Delivering Hydrologic Modeling functionality as a service over the web

- Services oriented scripting
 - Data services
 - Modeling services
- Web service gateway to HPC (HydroGate)
- Water Management Data Model (WamDam)
- Leverage (and contribute to) other related systems (HydroShare, CyberGIS)

CIWATER Service Oriented Architecture



HydroGate: An API for Authenticated Access to HPC Resources (Mt Moran)



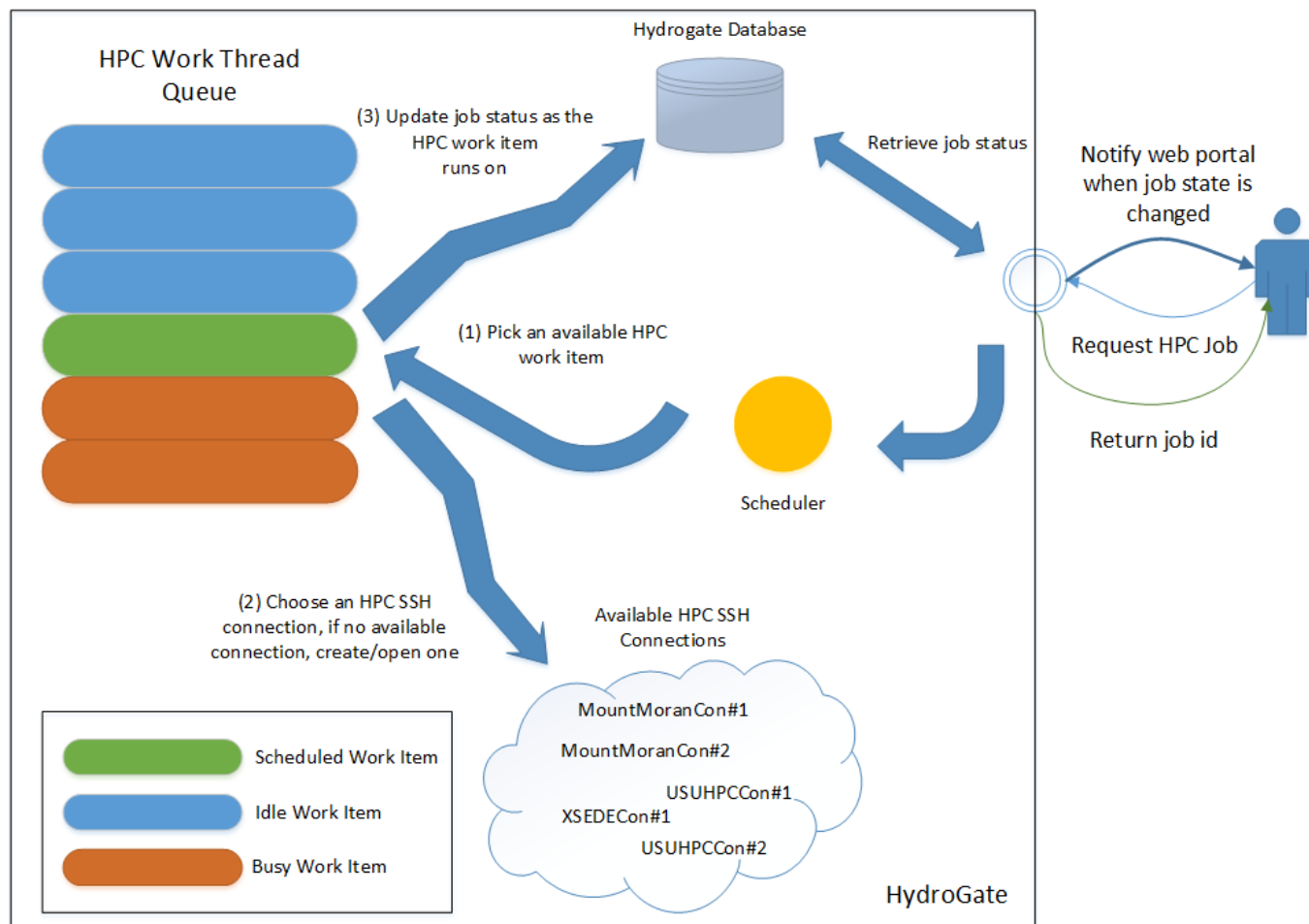
Uses standard secure shell (SSH) for communication so no software installation is needed.

Mt Moran access requires UWYO VPN

Methods

- RequestToken
- IsTokenValid
- SubmitPackage
- CheckPackageStatus
- DeletePackage
- SubmitJob
- CheckJobStatus
- DeleteJob
- GetProgramInformation

Asynchronous HPC Job Execution



Yildirim, A. A., D. Tarboton, P. Dash and D. Watson, (2014), "Design and Implementation of a Web Service-Oriented Gateway to Facilitate Environmental Modeling using HPC Resources," in D. P. Ames, N. W. T. Quinn and A. E. Rizzoli (eds), Proceedings of the 7th International Congress on Environmental Modelling and Software, San Diego, California, USA, June 16-19, 2014, International Environmental Modelling and Software Society (iEMSSs), ISBN: 978-88-9035-744-2, <http://www.iemss.org/society/index.php/iemss-2014-proceedings>.

Data Sets Supported via Data Services

Dataset	Server location
Elevation (USGS NED)	CI-WATER Server (Static)
Terrain Derivatives (Slope, Aspect, Flow Direction, Contributing area etc.)	CI-WATER Server (Static)
Land Cover (NLCD)	CI-WATER Server (Static)
Daymet Weather	CI-WATER Server (Dynamic, periodically updated)
NLDAS Weather	Dynamically retrieved from NASA

Prototype services for Greater Salt Lake area complete.

Operational Services for Western US still to be done.

Data Services Example

Input

demo.py x

```
from hydrogate import Client
client = Client()
client.subset_dem(left_top_x=432760, left_top_y=4662453, right_bottom_x=461700, right_bottom_y=4615000)
subset_dem_request = client.get_most_recent_request(service_name='subset_dem')
print subset_dem_request.file_path
client.download_file(file_url_path=subset_dem_request.file_path, save_as=r'C:\Users\dtarb\Desktop\HydroGateDemo\sub_dem.zip')
```

Result

C:\Python27\ArcGIS10.2\python.exe C:/Users/dtarb/Desktop/HydroGateDemo/demo.py
subset_dem execution was successful.

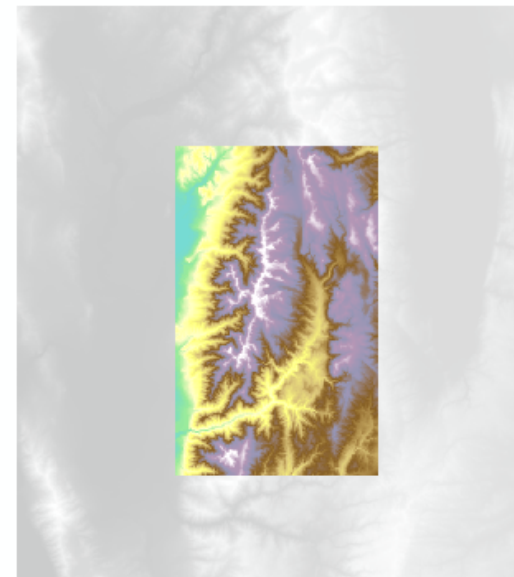
Output file URL path: <http://129.123.41.158:8080/dem/dem3504668463149717179.tif.zip>

```
{
  "Output file path": "http://129.123.41.158:8080/dem/dem3504668463149717179.tif.zip",
  "Service status": "success",
  "Request time": "2014-11-15 23:53:20.538000",
  "Service name": "subset_dem",
  "Service ID name": "",
  "Service ID value": ""
}
```

<http://129.123.41.158:8080/dem/dem3504668463149717179.tif.zip>

Downloaded file saved successfully at: C:\Users\dtarb\Desktop\HydroGateDemo\sub_dem.zip

Process finished with exit code 0



Modeling example

Input

```
modeldemo.py x demo.py x
from hydrogate import Client
client = Client()
client.subset_dem(left_top_x=410000, left_top_y=4682453, right_bottom_x=481700, right_bottom_y=4600000)
# Previously created shapefile
outlet_shapefile_url = "http://129.123.41.158:8080/dem/LoganOutlet.shp.zip"
subset_dem_request = client.get_most_recent_request(service_name='subset_dem')
print subset_dem_request.file_path
input_raster_url = subset_dem_request.file_path
client.generate_watershed_raster(input_raster_url_path=input_raster_url,
                                outlet_shapefile_url_path=outlet_shapefile_url,
                                save_as=r'C:\Users\dtarb\Desktop\HydroGateDemo\WS_Logan.tif')
```

Result

C:\Python27\ArcGIS10.2\python.exe C:/Users/dtarb/Desktop/HydroGateDemo/modeldemo.py
subset_dem execution was successful.

Output file URL path: <http://129.123.41.158:8080/dem/dem3937079318519734987.tif.zip>

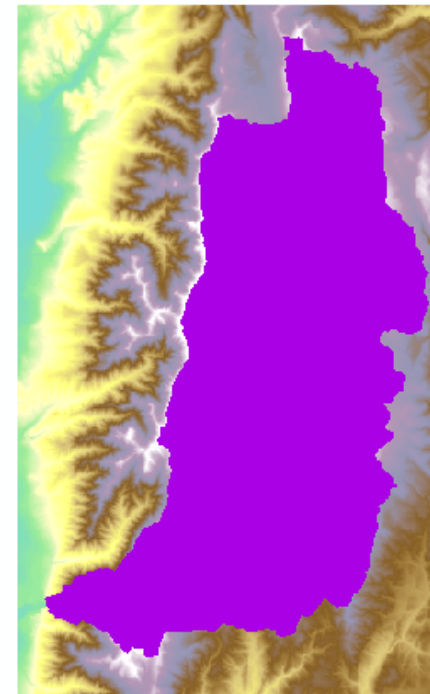
```
{
  "Output file path": "http://129.123.41.158:8080/dem/dem3937079318519734987.tif.zip",
  "Service status": "success",
  "Request time": "2014-11-16 00:26:20.548000",
  "Service name": "subset_dem",
  "Service ID name": "",
  "Service ID value": ""
}
```

<http://129.123.41.158:8080/dem/dem3937079318519734987.tif.zip>

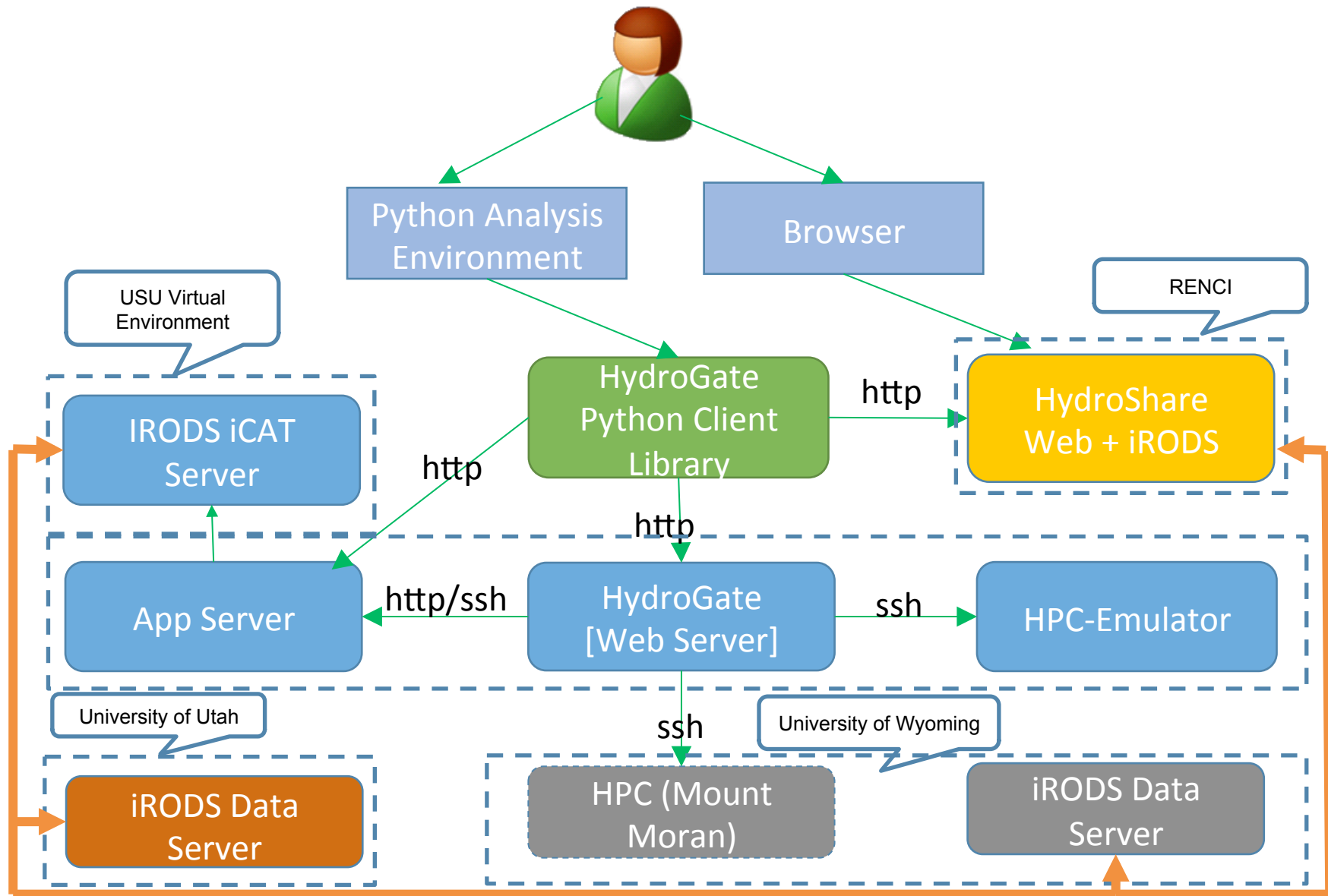
generate_watershed_raster execution was successful.

Output file URL path: <http://129.123.41.158:8080/dem/dembfa12a4fdf924fa6b0896f892eb837c8WS.tif.zip>

Downloaded file saved successfully at: C:\Users\dtarb\Desktop\HydroGateDemo\WS_Logan.tif



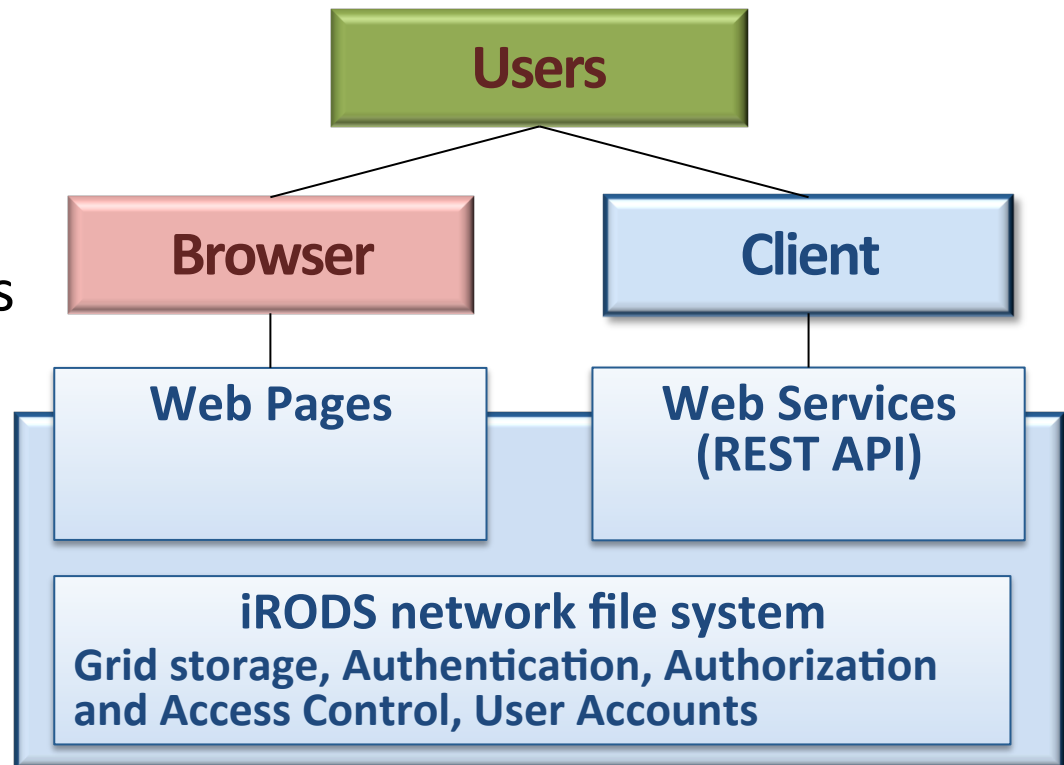
Planned Fully Built CIWATER Service Oriented Architecture



iRODS Federated Data Grid Network File System

HydroShare

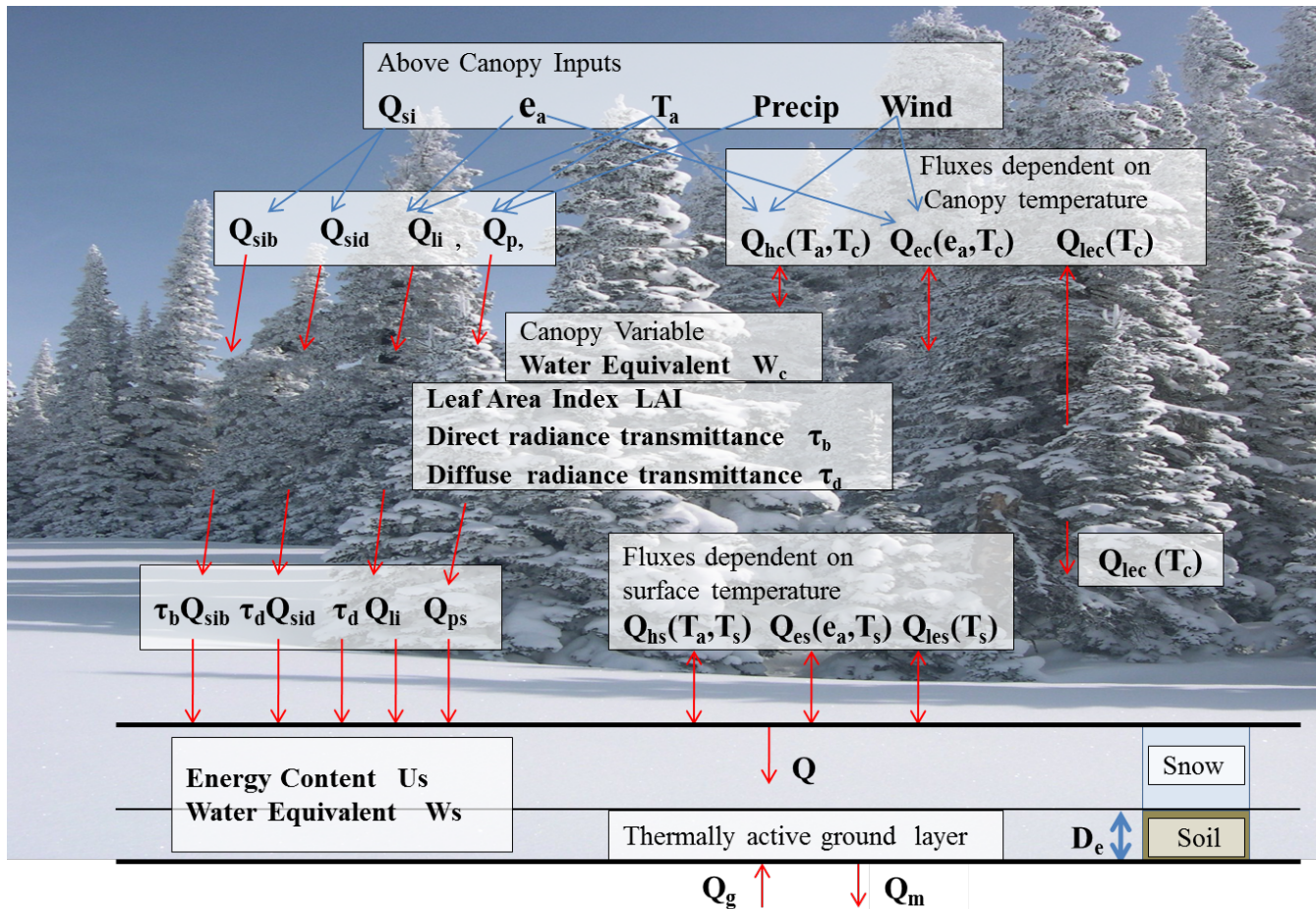
HydroShare is being developed as an online, collaborative environment for the sharing of hydrologic data and models



- CI-WATER will leverage HydroShare user accounts and web system for federated file management
- CI-WATER apps will contribute to HydroShare functionality

Utah Energy Balance Snowmelt Model

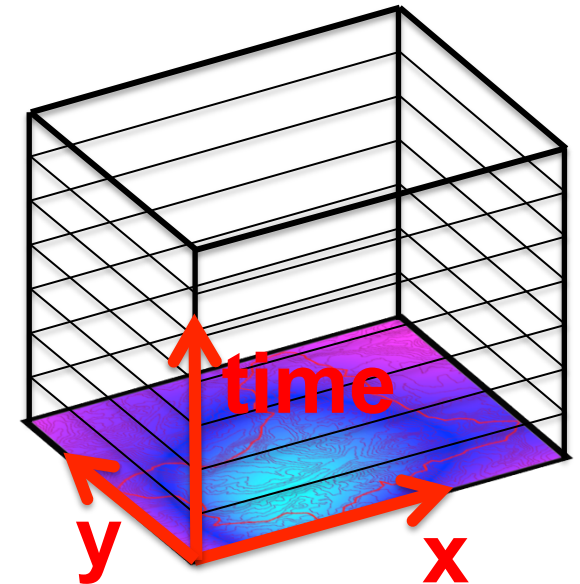
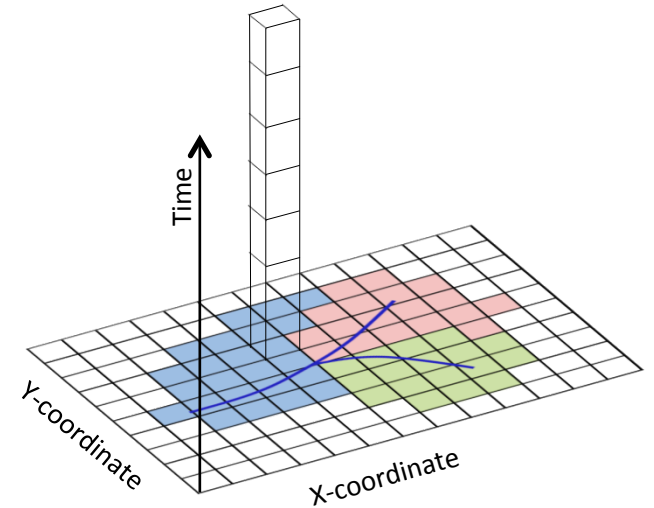
Used in CI-WATER to address what are the impacts of land cover change on watershed snowmelt inputs



Mahat, V. and D. G. Tarboton, (2012), "Canopy radiation transmission for an energy balance snowmelt model," Water Resour. Res., 48: W01534, <http://dx.doi.org/10.1029/2011WR010438>.

UEB use case

- Model run separately at each active grid cell
- Parallel implementation for large areas using HPC
- Data services to provide input data
- Data services to configure model inputs
- Modeling services to execute model



TauDEM is software for deriving Hydrologically Useful Information from Digital Elevation Models

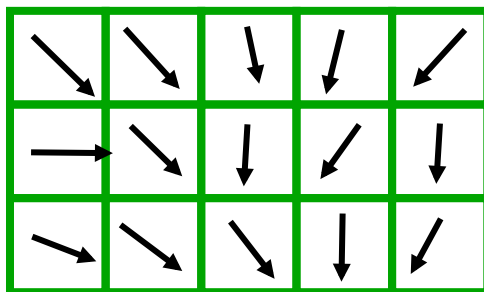
Raw DEM



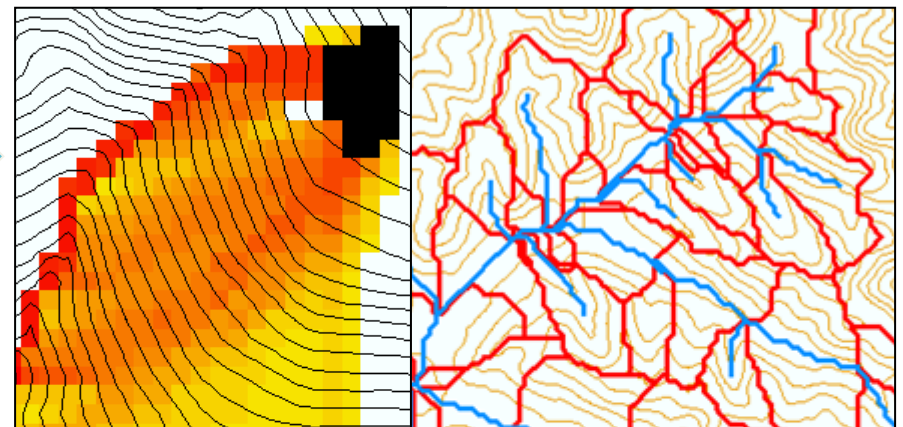
Pit Removal



Flow Field



Flow Related Terrain Information



Used in CI-WATER for terrain analysis and watershed delineation

CyberGIS Platform for Web GIS applications

The screenshot displays the CyberGIS Gateway web application interface. The browser address bar shows the URL: `gateway.cigi.illinois.edu/home/apps.php?app=taudem`. The page features a navigation menu with links for Home, Apps, Visualization, Community, and Help. The main content area is divided into two panels, both titled "App: TauDEM".

Left Panel (My Analysis: Reynolds):

- App:** TauDEM
- My Analysis:** Reynolds
- Data and Parameters:**
 - 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:** (empty)

Select the products you want

TauDEM Workflow Wizard

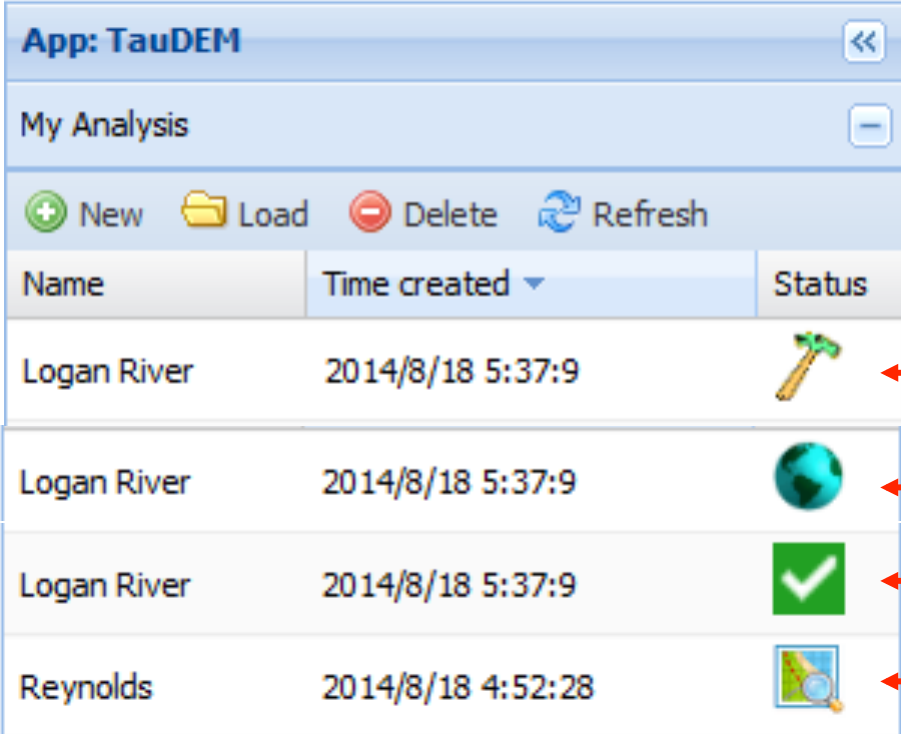
Select Products

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

The job progresses through the system



The screenshot shows the 'App: TauDEM' interface. It features a toolbar with 'New', 'Load', 'Delete', and 'Refresh' buttons. Below the toolbar is a table with columns for 'Name', 'Time created', and 'Status'. The table contains four rows of data, each with a corresponding icon in the 'Status' column. Red arrows point from the text labels on the right to the icons in the table.

Name	Time created	Status
Logan River	2014/8/18 5:37:9	 Analysis submitted
Logan River	2014/8/18 5:37:9	 Analysis running
Logan River	2014/8/18 5:37:9	 Results data created
Reynolds	2014/8/18 4:52:28	 Results Ready

Execution is on XSEDE behind the scenes

Results displayed in browser

File Edit View History Bookmarks Tools Help

CyberGIS: high-performance, d... x +

gateway.cigi.illinois.edu/home/apps.php?app=taudem

dtarb's Profile | Logout

CyberGIS Gateway Home Apps Visualization Community Help

App: TauDEM

My Analysis: 4668:Logan Stream Network

Data and Parameters

Results

Data

- TauDEM:4668_LoganStreamNetworkssa
- TauDEM:4668_LoganStreamNetworksd8
- TauDEM:4668_LoganStreamNetworksca

Legend

Dinfinity Contributing Area (metre)

- 88.8441860009 - 266.532558003
- 266.532558003 - 1599.19534802
- 1599.19534802 - 14392.7581321
- 14392.7581321 - 172713.097586
- 172713.097586 - 2590696.46379
- 2590696.46379 - > 2590696.46379

1.75596, 41.91792

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: 0846855 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.

GI Solve

The collaboration with CyberGIS has enhanced the capability to use TauDEM for large datasets

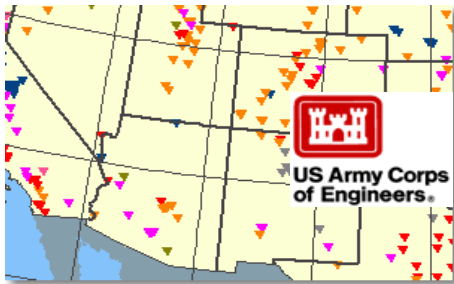
Water Management Problem

Water management data resides in different **data sources**, uses different **firmware**, **formats**, **terminology**, and applies to various domains and contexts with various available **metadata**

- What are the water system components and attributes in a geographic and domain area of interest?
- How are these components physically connected to each other?
- What data is available to run a particular model in a particular place?

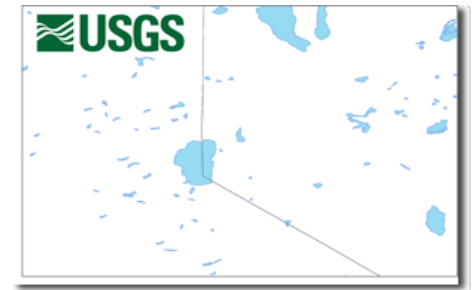


How to organize all these together?



US Dams dataset
23 attributes
8,121 instances

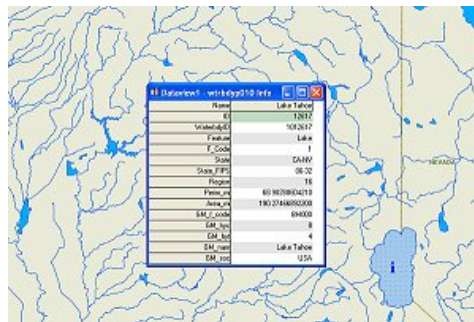
- Consistent semantics and syntactic structure
- Supportive metadata



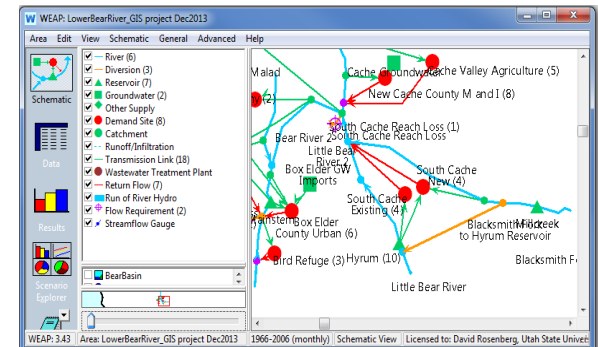
US Water Bodies and Wetlands Dataset
15 attributes
26,872 instances



Time Series Data
32 attributes



Streams Network
22 attributes



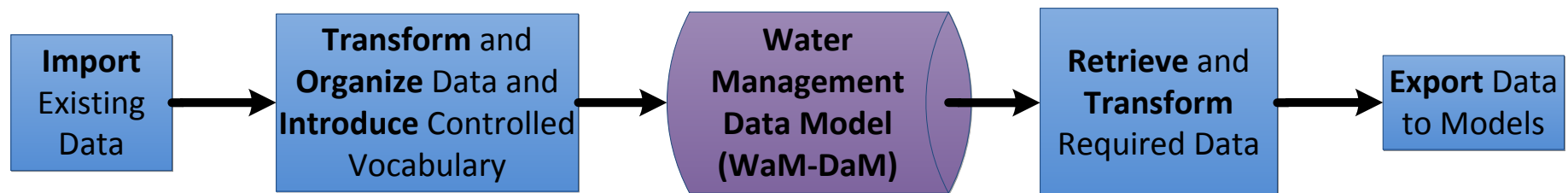
WEAP Model
Lower Bear River, UT
53 instances ²⁵

We a need a data model to supports all these common features

Model	Flexible and extensible	Networks	Scenarios	conditional query	Dynamic controlled vocabulary	Descriptive and explicit metadata	Multiple data formats	Open source envir.
Arc Hydro								
ODM								
HydroPlatform								
WEAP								
HEC-DSS								
WaDE								
WISKI Kisters								
GoldSim								
SWMM								
CALVIN								
ArcSWAT								
GSSHA								
MODSIM								
TOPNET								
AdHydro								

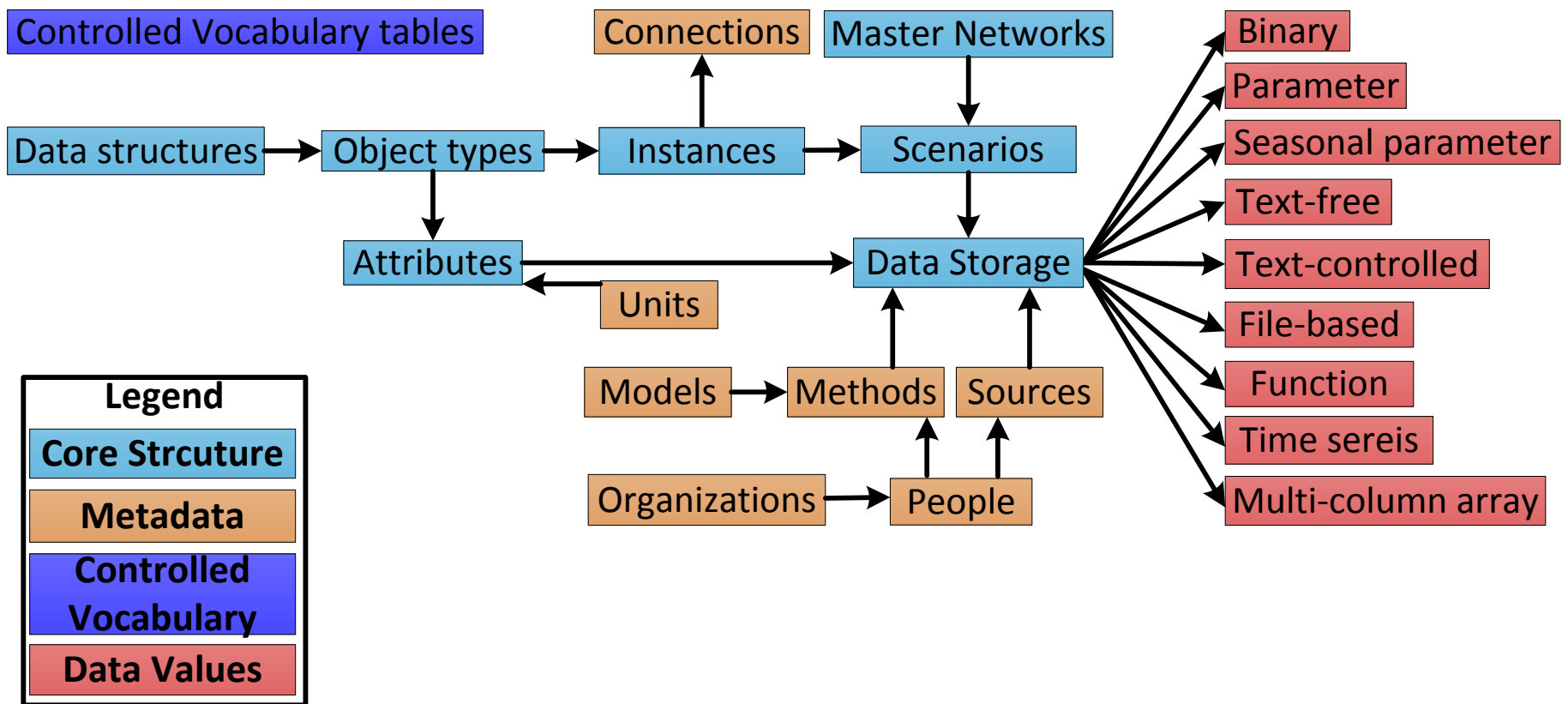
Water Management Data Model (WaM-DaM)

1. **Organize** water management data
2. **Synthesize** data across domains and sources
3. **Compare** data from different scenarios
4. **Serve** data to run models
5. **Publish** model data and share with others



Abdallah, A. M. and D. E. Rosenberg, (2014), "WaM-DaM: A Data Model to Organize and Synthesize Water Management Data," in D. P. Ames, N. W. T. Quinn and A. E. Rizzoli (eds), Proceedings of the 7th International Congress on Environmental Modelling and Software, San Diego, California, USA, International Environmental Modelling and Software Society (iEMSS), ISBN: 978-88-9035-744-2, <http://www.iemss.org/society/index.php/iemss-2014-proceedings>.

WaM-DaM Conceptual Design

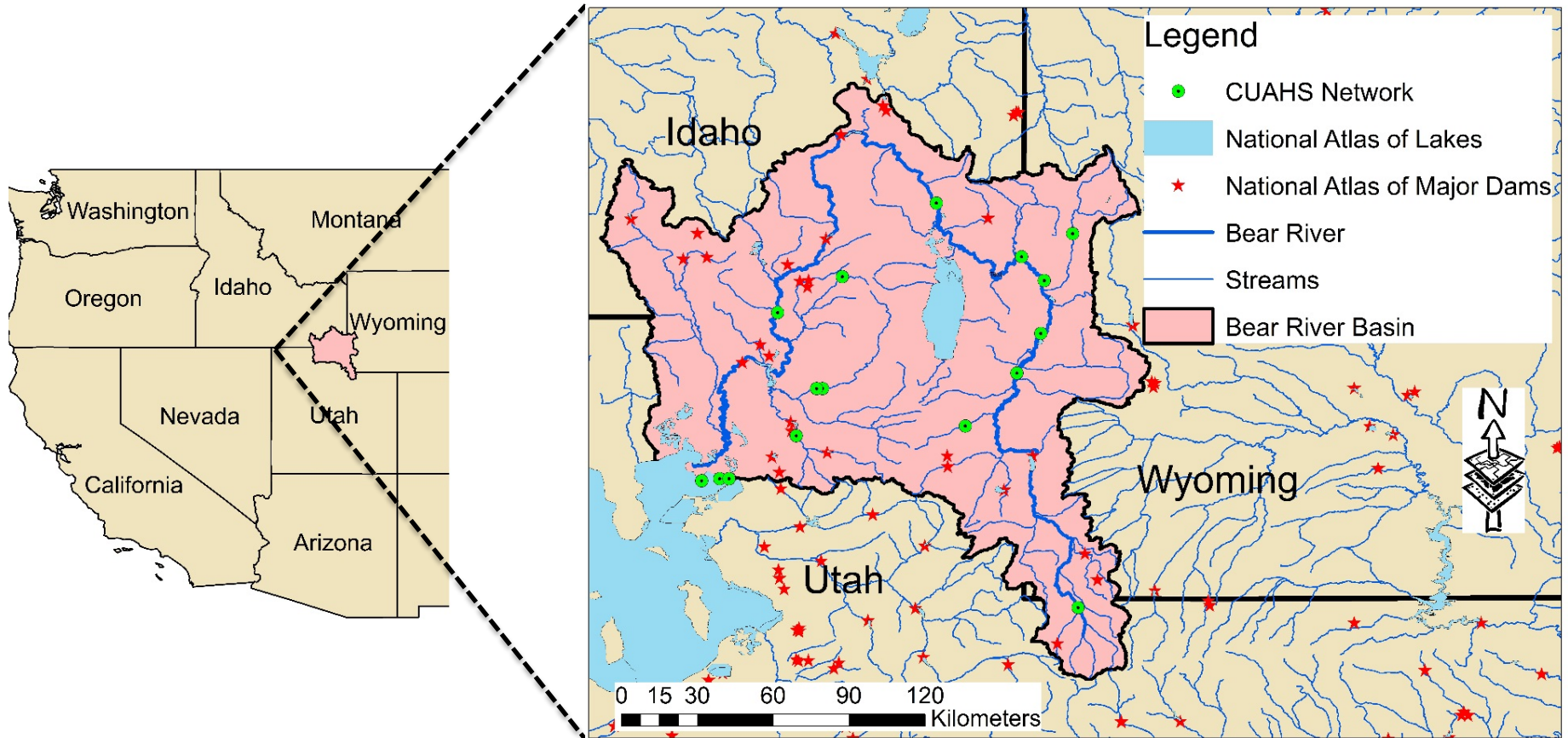


Follow WaM-DaM development @
<https://github.com/amabdallah/WaM-DaMv1.0>



Use cases

1. Integrate disparate water management data for the Bear River Basin, Utah



Use cases (cont.)

2. **Identify** differences in topology, metadata, and data between two WEAP scenarios in the lower Bear River basin
3. **Serve** data from WaM-DaM to WEAP, SWAT, and GoldSim models
4. **Use** HydroGate to run city water/energy simulation/optimization model on HPC resource

Accomplishments

- Prototype capability for supporting access to data required to support data intensive physically based distributed modeling
- HydroGate- Web based access to HPC resources
- Python Client Tool – Easy access to CI-WATER Web Services Tool and app prototypes and ongoing development

WaM-DaM Accomplishments

- Provides a **common** persistence model for water management data
- **Support** syntactic and semantic consistency
- Allow **interoperability** of data across models

Next Steps

- Complete data access and modeling services development
- Set up operational data services over Western US
- Use iRODS and federate across data services to provide “network file system” and transparent transport layer
- Use HydroShare for user management and access control (and thereby leverage other HydroShare capabilities available through federation with HydroShare iRODS data grid)
- Further cultivate partnerships to sustain development and functionality past end of grant (HydroShare, CyberGIS)