

HydroShare and iRODS: How iRODS Manages Data for a Hydrology Community of 1000's of Users

Ray Idaszak, David G. Tarboton (PI), Hong Yi, Chris Calloway, Shaowen Wang, Jeffery Horsburgh, Dan Ames, Martyn Clark, Jon Goodall, Alva Couch, Tony Castronova, Christina Bandaragoda, Martin Seul, Mark Henderson, Phuong Doan (*underlined names @ iRODS 2018 UGM in-person*)



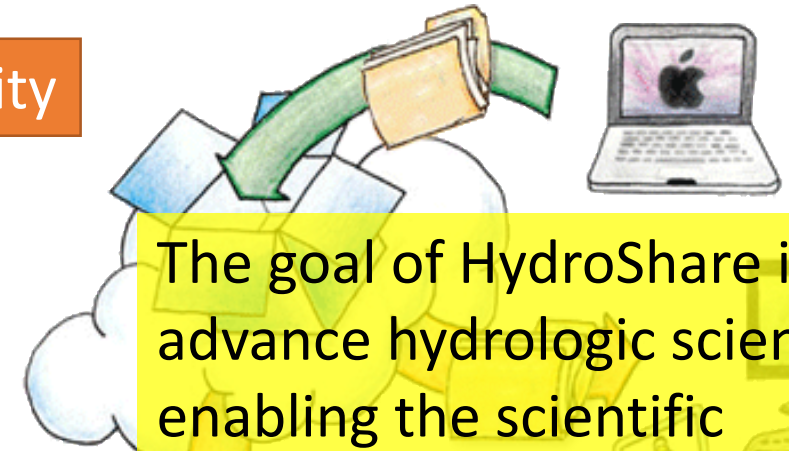
<http://www.hydroshare.org>



ACI 1148453, 1148090, 1664018,
1664061, and 1664119.
2012-2021

HydroShare is a platform for sharing Hydrologic Resources and Collaborating

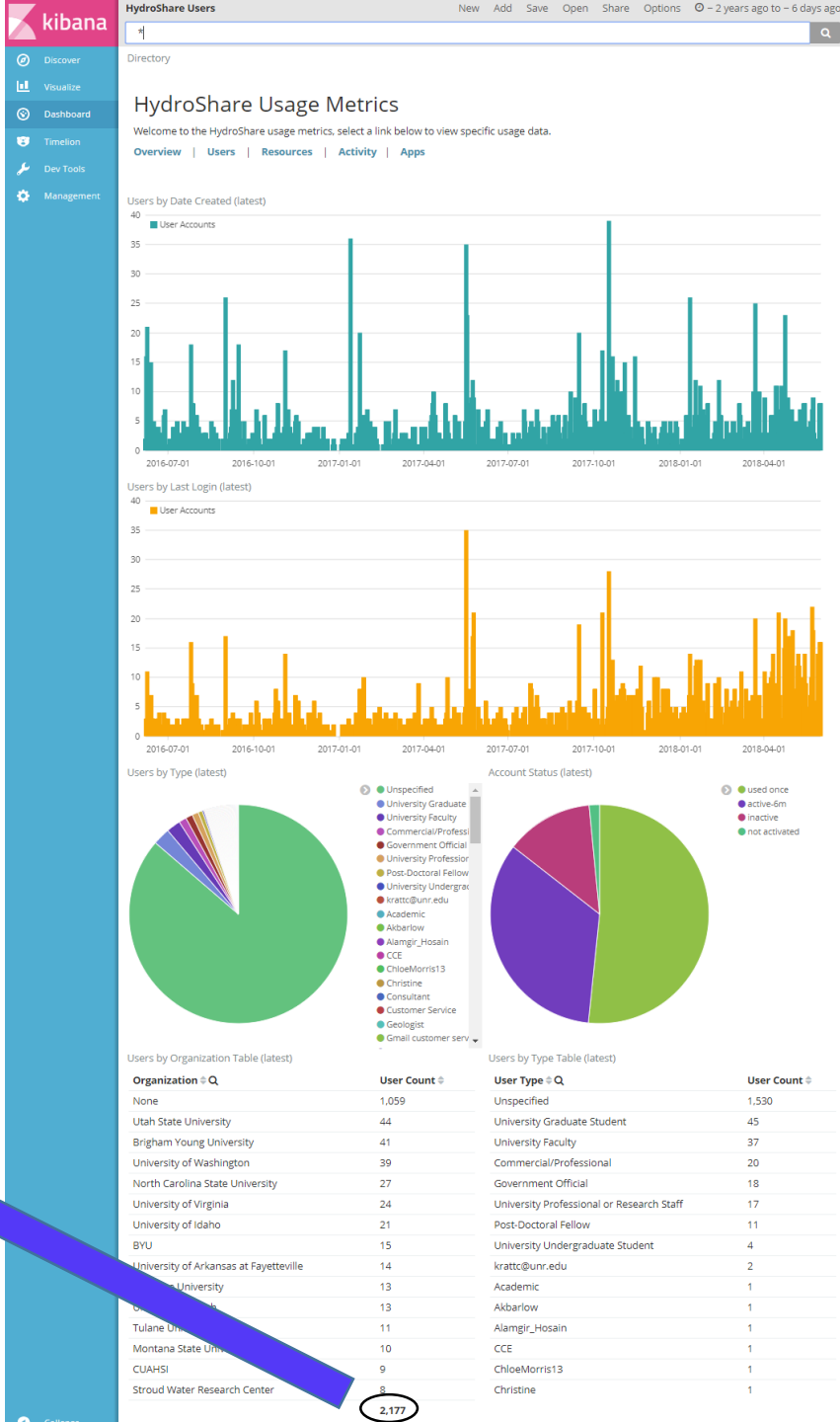
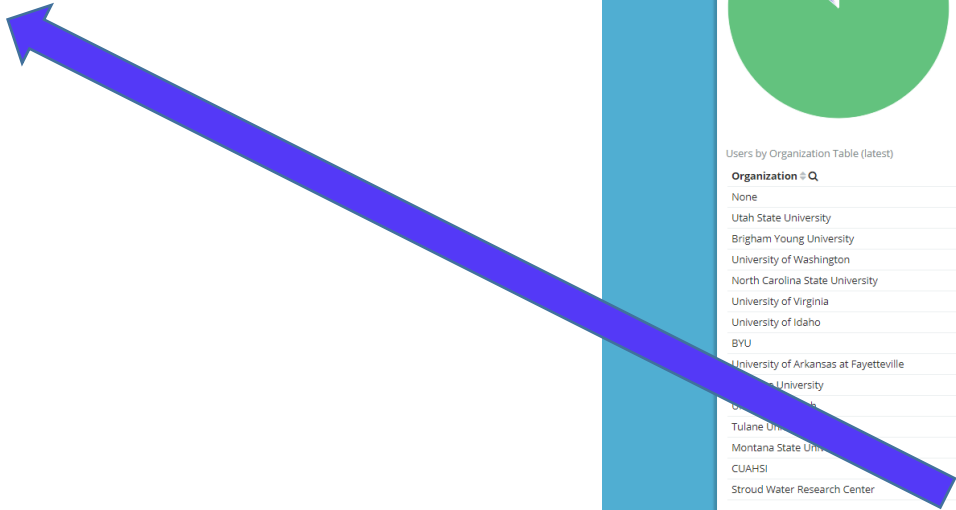
- File Storage DropBox-ish Functionality
- Meta Data Descriptions
- Data Access API
- Web Apps Value Added Functionality
- Social Functions
- DOI Data Publication



The goal of HydroShare is to advance hydrologic science by enabling the scientific community to more easily and freely share products resulting from their research - not just the scientific publication summarizing a study, but also the data and models used to create the scientific publication.

HydroShare Usage Metrics as of June 2018: > 2,000 users

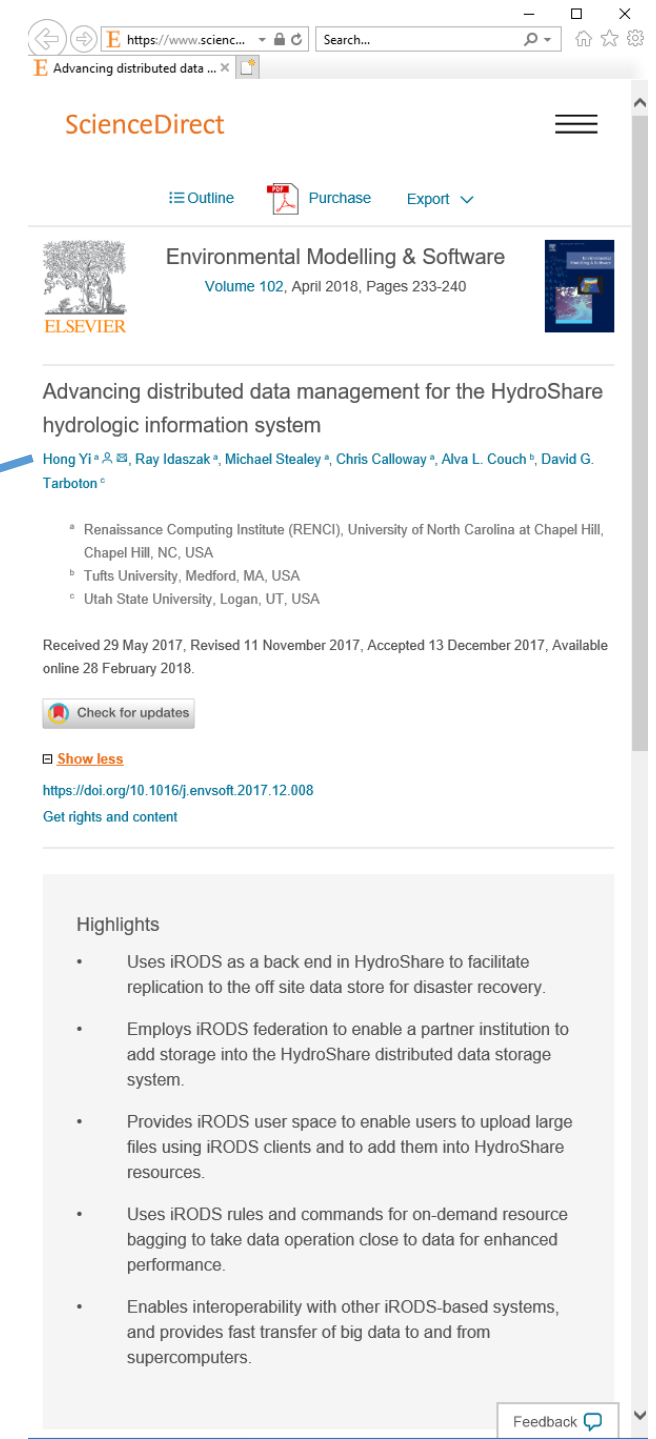
2,177



The best place to learn more about HydroShare and iRODS

Dr. Hong Yi et. al., Advancing distributed data management for the HydroShare hydrologic information system, Feb 2018,
<https://doi.org/10.1016/j.envsoft.2017.12.008>

<http://bit.ly/hydroshareandirops>



The screenshot shows the ScienceDirect website interface. At the top, the ScienceDirect logo is visible. Below it, the journal title 'Environmental Modelling & Software' and volume information 'Volume 102, April 2018, Pages 233-240' are displayed. The article title 'Advancing distributed data management for the HydroShare hydrologic information system' is prominently featured, followed by the authors 'Hong Yi^a, Ray Idaszak^a, Michael Stealey^a, Chris Calloway^a, Alva L. Couch^b, David G. Tarboton^c'. The affiliations for each author are listed below. A 'Check for updates' button is present. The DOI link 'https://doi.org/10.1016/j.envsoft.2017.12.008' is provided, along with a link to 'Get rights and content'. A 'Highlights' section at the bottom lists five key points about the article's content, including the use of iRODS for data management and federation.

ScienceDirect

Outline Purchase Export

Environmental Modelling & Software
Volume 102, April 2018, Pages 233-240

Advancing distributed data management for the HydroShare hydrologic information system

Hong Yi^a, Ray Idaszak^a, Michael Stealey^a, Chris Calloway^a, Alva L. Couch^b, David G. Tarboton^c

^a Renaissance Computing Institute (RENCI), University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
^b Tufts University, Medford, MA, USA
^c Utah State University, Logan, UT, USA

Received 29 May 2017, Revised 11 November 2017, Accepted 13 December 2017, Available online 28 February 2018.

Check for updates

Show less
<https://doi.org/10.1016/j.envsoft.2017.12.008>
Get rights and content

Highlights

- Uses iRODS as a back end in HydroShare to facilitate replication to the off site data store for disaster recovery.
- Employs iRODS federation to enable a partner institution to add storage into the HydroShare distributed data storage system.
- Provides iRODS user space to enable users to upload large files using iRODS clients and to add them into HydroShare resources.
- Uses iRODS rules and commands for on-demand resource bagging to take data operation close to data for enhanced performance.
- Enables interoperability with other iRODS-based systems, and provides fast transfer of big data to and from supercomputers.

Feedback

In HydroShare you can:

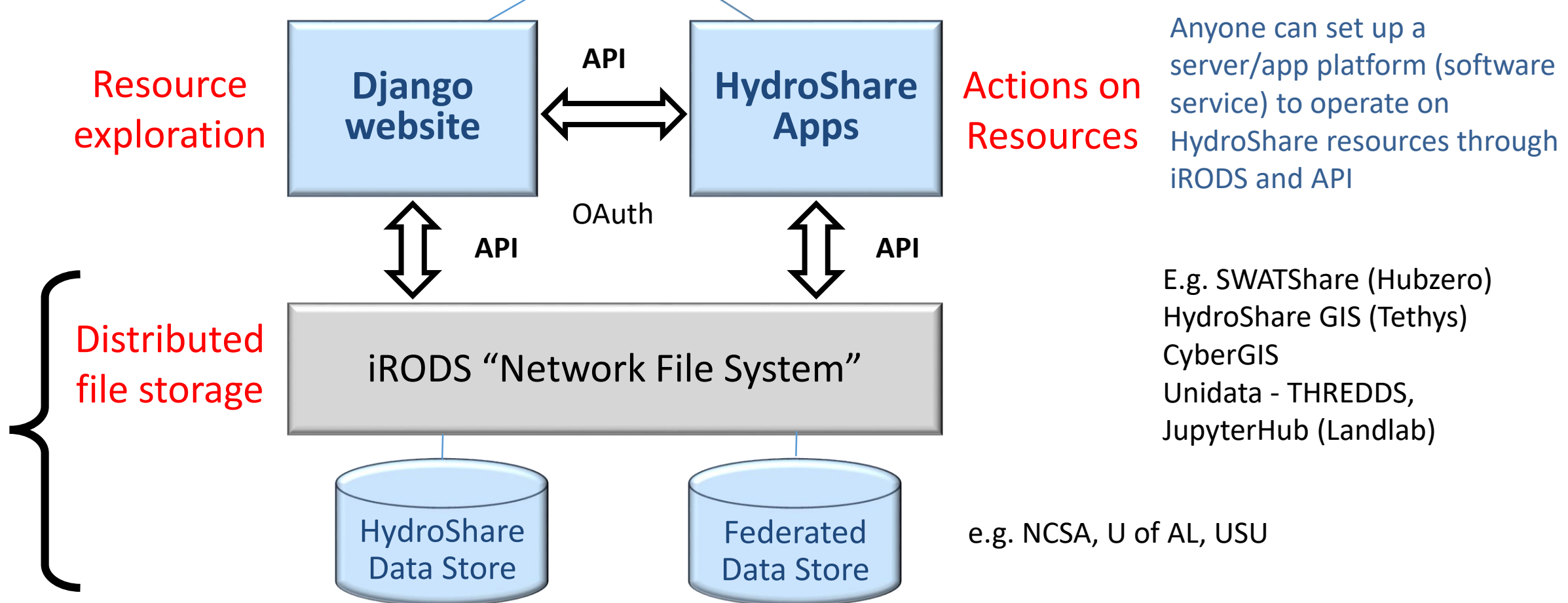
- Share your data and models with colleagues
- Manage who has access to the content that you share
- Share, access, visualize and manipulate a broad set of hydrologic data types and models
- Use the web services API to program automated and client access
- Publish data and models to document research findings supporting open data, reproducibility, transparency and trust in results (and meet the requirements of your data management plan and receive a citable digital object identifier (DOI) to get credit for your work)
- Discover and access data and models published by others
- Use web apps to visualize, analyze and run models on data in HydroShare

How HydroShare Works

Slide from Tarboton et. al. "HydroShare Present and Future: Advances in the Hydroshare Platform for Collaborative Data and Model Sharing," 2017 CUAHSI Conference on Hydroinformatics, July 25-27, <https://www.hydroshare.org/resource/6cb2da4dffa248c09bc4d7d883fdf4a1/>

- Organize and annotate your content
- Manage access

- Web software to operate on content you have access to (Apps)
- Extensibility

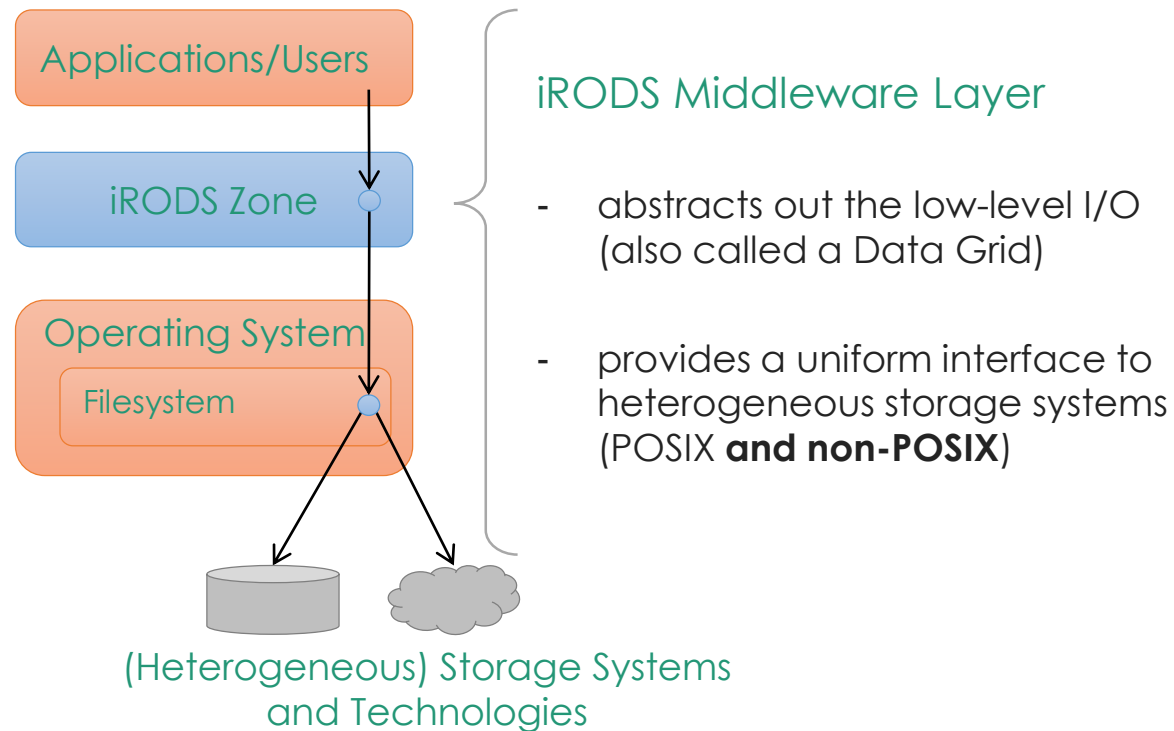


iRODS Data Virtualization

iRODS: The integrated Rule-Oriented Data System

iRODS is open source data grid middleware that implements...

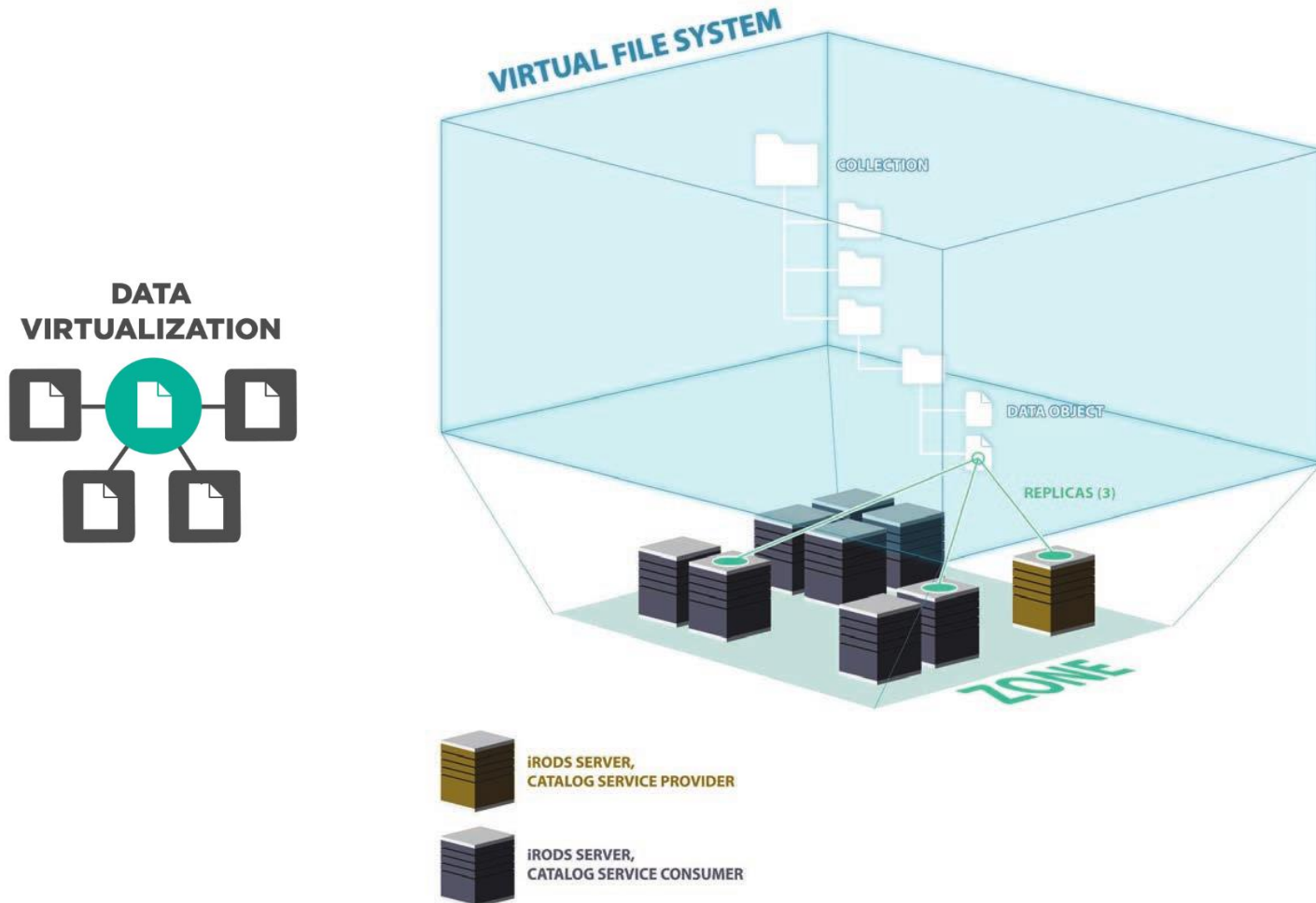
- Data Virtualization
- Automation of Data Operations
- A Robust Metadata Catalog
- Data Management Policy Enforcement and Compliance Verification



iRODS
<https://irods.org/>

IRODS Zone

IRODS provides a virtual system: logical representation of file hierarchies (called Collections) stored in distributed physical storage locations



iRODS presents centralized distributed storage systems under a unified namespace.

Administrators can control how the zone is presented to users and implement replication, load-distribution, and archiving policies that are completely transparent to the user.

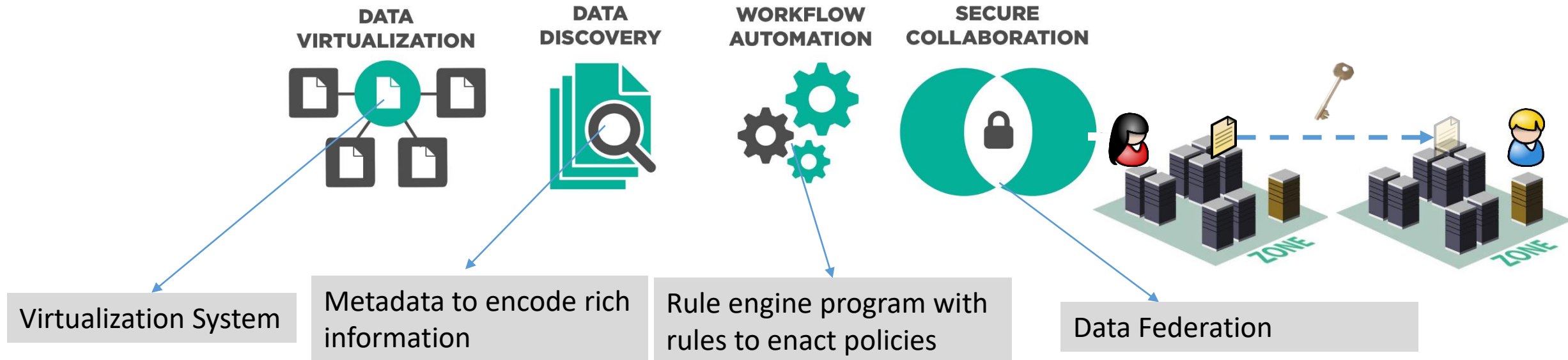
Independent zone can be federated with one another to allow controlled access to remote zones or zones operated by separate workgroups.

iRODS Key Features



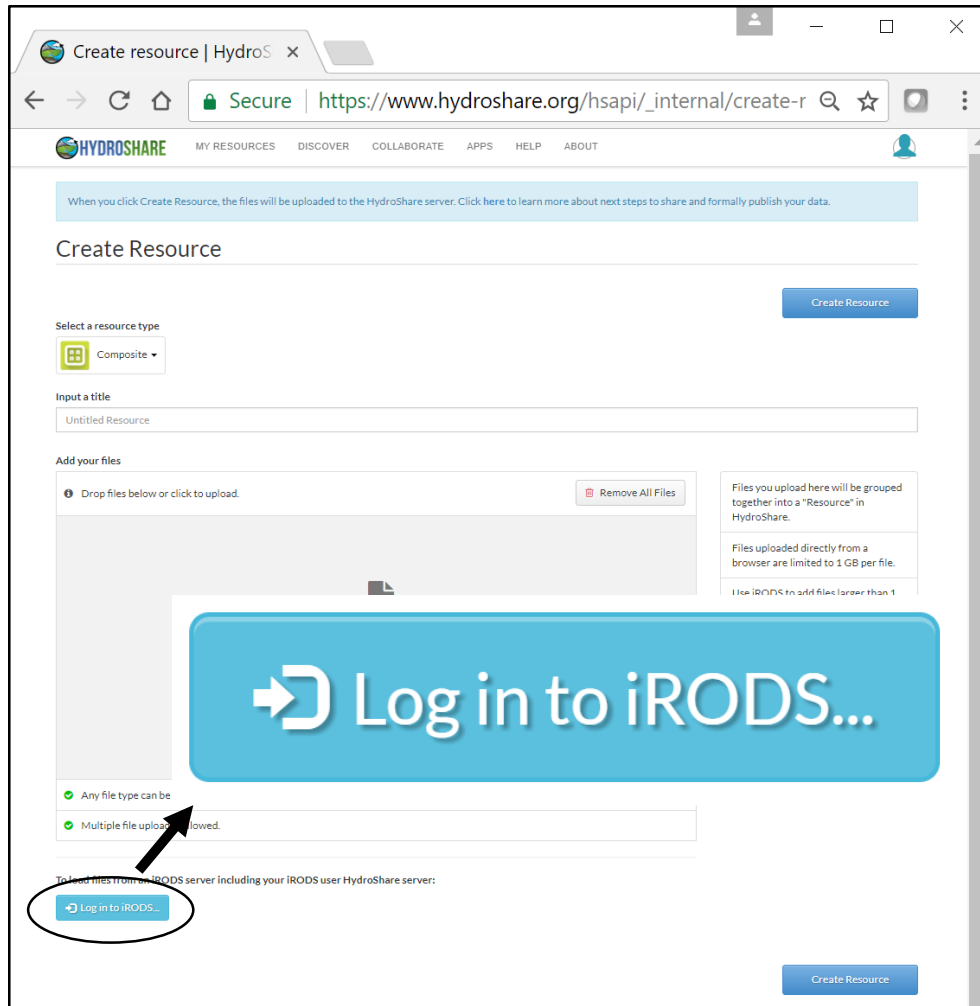
The Integrated Rule-Oriented Data System:

- Developed for working with massive collections of files
- Organizing, securing, preserving, and sharing data

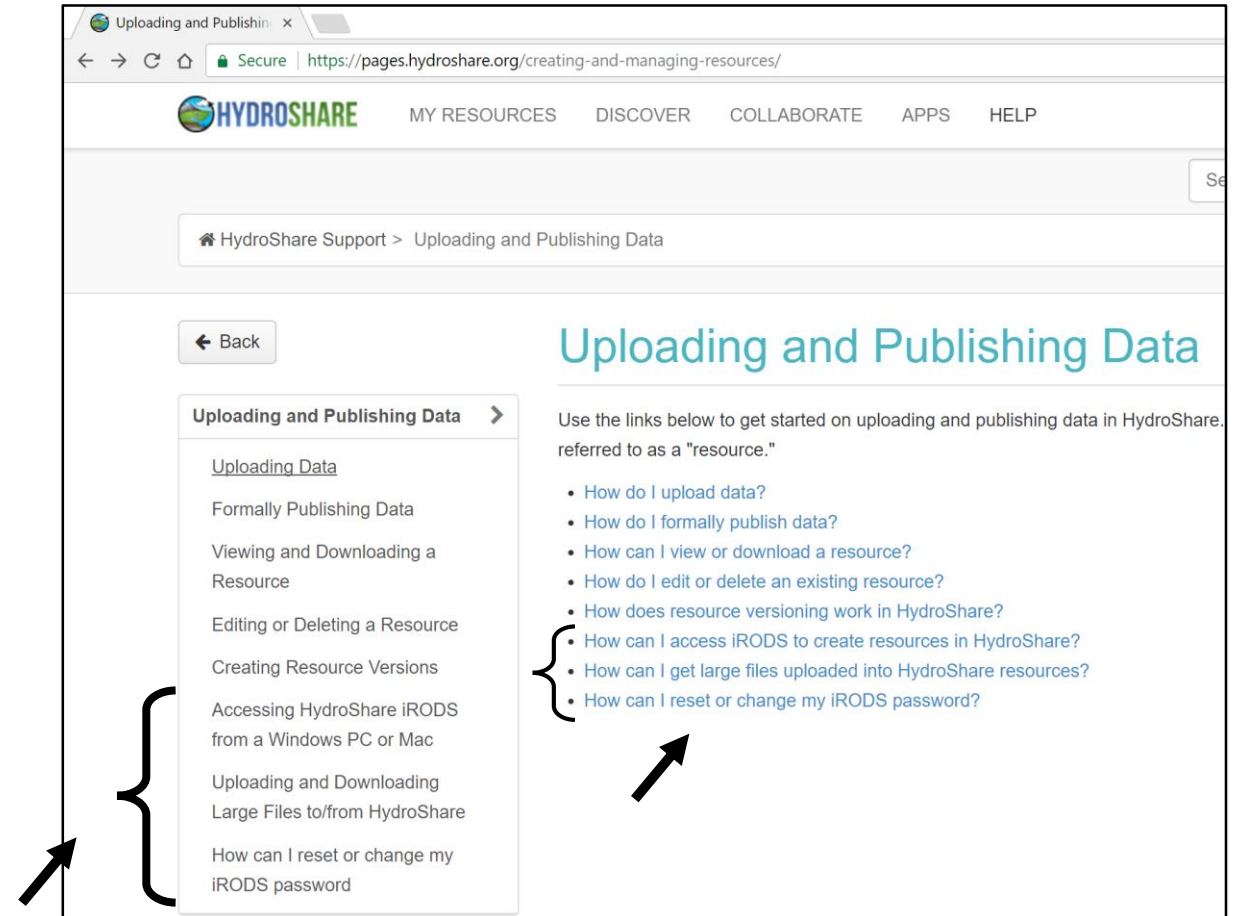


iRODS in the current HydroShare

- <https://help.hydroshare.org/creating-and-managing-resources/>



iRODS on the HydroShare resource landing page.



iRODS how-to discussed on the HydroShare Support pages.

NWM Forecast Viewer App

The image displays two browser windows from the HydroShare platform. The bottom window, titled "HydroShare Apps Portal", shows a grid of application tiles. The "National Water Model Forecast Viewer" tile is circled in black. An arrow points from this tile to a larger, detailed view of the app's interface shown in the top window.

HydroShare Apps Portal Interface:

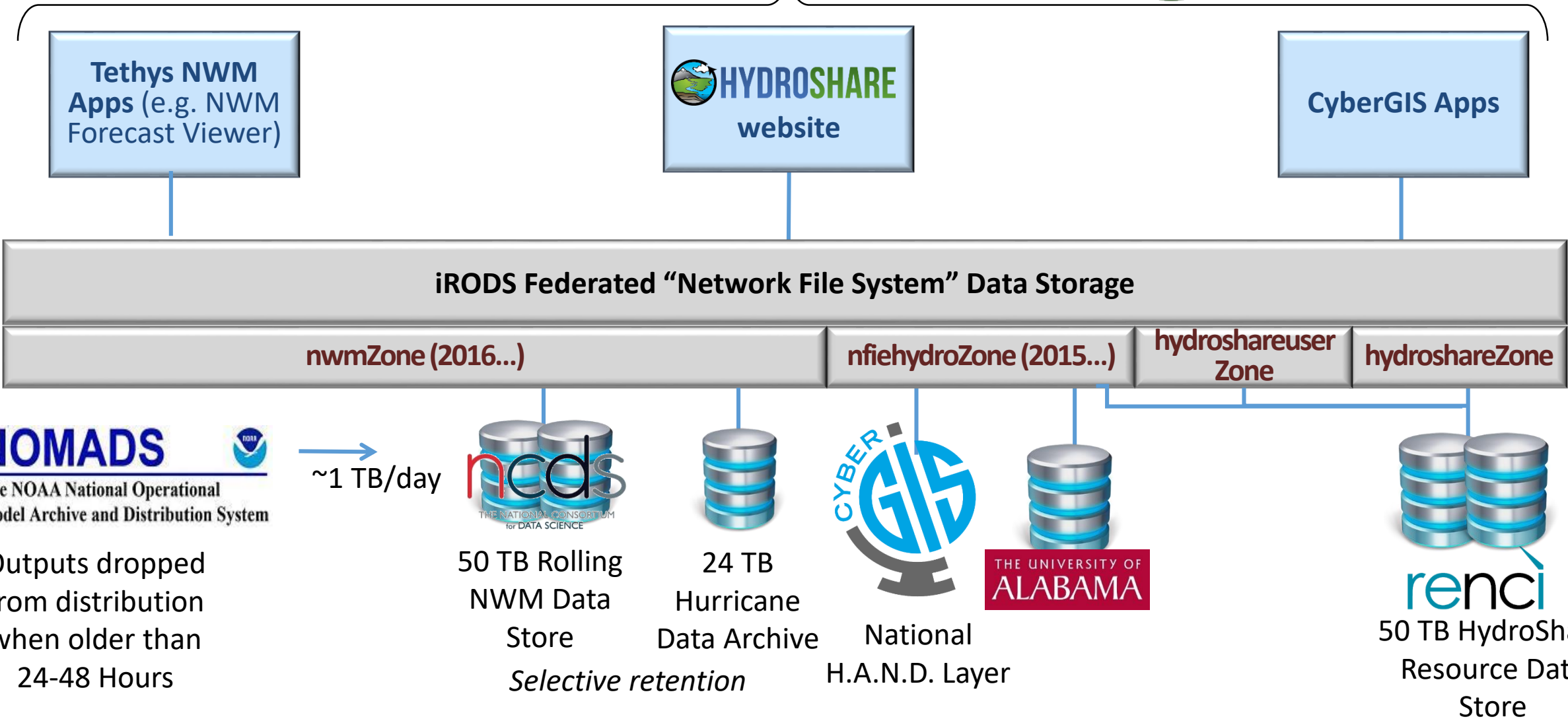
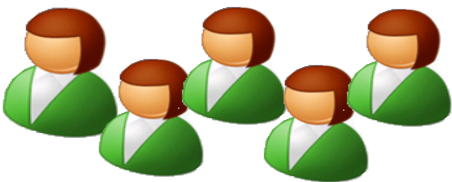
- Header: "HydroShare Apps Portal" with a logo and user profile "rayi".
- App Tiles:
 - Cybergis TauDEM App
 - HydroShare GIS
 - National Water Model Forecast Viewer** (highlighted)
 - Data Rods Explorer

National Water Model Forecast Viewer Interface:

- Header: "National Water Model Forecast Viewer" with a logo and navigation links: "Help...", "API", "Exit".
- Left Sidebar:
 - Home (red button)
 - Subset (text link)
 - Add Watershed (blue button)
- Main Content: A map of a watershed area with blue stream networks. The map includes labels for "CLEVELAND COUNTY", "Coke", "Northport", "Holt", "Brookwood", "Cotton", "Cualing", "ruscaloosa", and "WEST TUSCALOOSA".

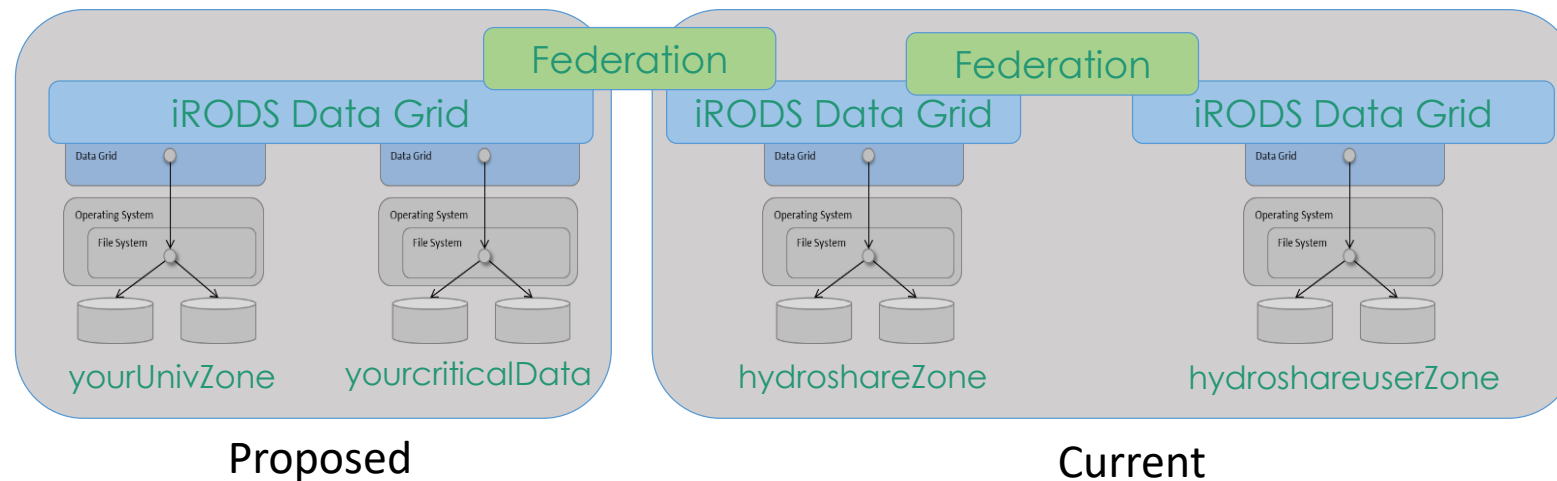
HydroShare: National Water Model Community Data Access Architecture

Community of users, developers, contributors and hydrologic science researchers



Exploring: HydroShare Extended Storage Ecosystem

- Potential benefits of this extended storage ecosystem for the current HydroShare include but are not limited to:
 - Use your own campus or organization's physical disk space towards HydroShare, especially if more than HydroShare's 50TB are needed
 - Have your own storage policies, e.g. quotas, archiving, replication
 - Host your own unique hydrology research data sets analogous to the National Water Model



To learn more

- <https://www.hydroshare.org/>
- <https://doi.org/10.1016/j.envsoft.2017.12.008>
- <https://help.hydroshare.org/>
- <http://youtube.hydroshare.org/>
- <https://irods.org/>
- <https://www.cuahsi.org/data-models/portals/cuahsi-data-services>

YouTube interface showing a video titled "HydroShare: Advancing Collaboration through Hydrologic Data and Model Sharing" by David Tarboton. The video player displays a slide titled "Architecture" with a diagram of the system components.

The diagram illustrates the HydroShare architecture. At the top, "Users" interact with a "Browser" and a "Client (HydroDesktop)". The "Browser" connects to "Web Tools" and "Web Pages", which in turn connect to the "Django web application framework". The "Client (HydroDesktop)" connects to "Web Services (HydroShare REST API)". To the right, "3rd party clients, interoperable systems and web tools" connect to the "Web Services (HydroShare REST API)". Below the Django framework is the "iRODS 'Network File System'", which contains "Content" (Resource Files, Resource Science Metadata, User Accounts) and "Functionality" (Authentication and Authorization Infrastructure (AAI) for Access Control, Discovery (using e.g. ElasticSearch via MSVC plugin)). The "iRODS Native REST API" is also shown, connecting to the "Web Services (HydroShare REST API)". A list of connection points for partners is provided: BIG CZ SS, C-WATER, SEAD, CyberGIS, and SESYNC GI Venture.

HydroShare: Advancing Collaboration through Hydrologic Data and Model Sharing -- David Tarboton

299 views

1 like, 0 dislikes, SHARE, ...

Thanks to the HydroShare team!

- USU
- RENCI / UNC
- CUAHSI
- NCSA / UIUC
- BYU
- Tufts
- UVA
- Univ of Washington



<http://www.hydroshare.org>



ACI 1148453, 1148090, 1664018,
1664061, and 1664119.
2012-2021