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

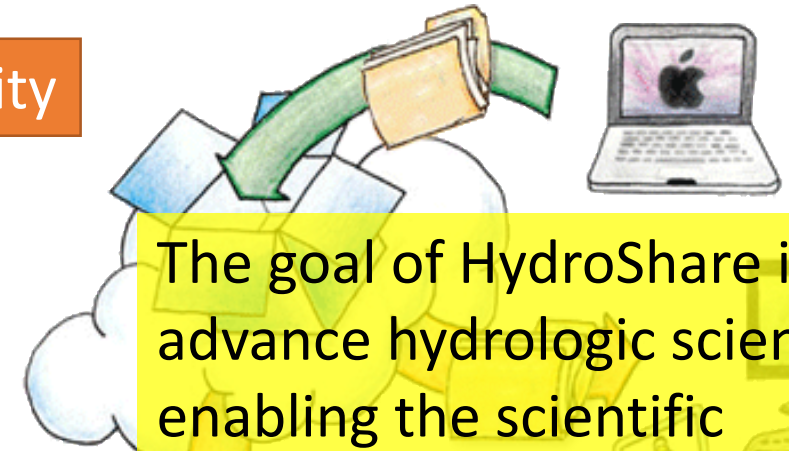Henderson, Phuong Doan *(underlined names @ iRODS 2018 UGM in-person)*

http://www.hydroshare.org

HydroShare is a platform for sharing Hydrologic Resources and Collaborating
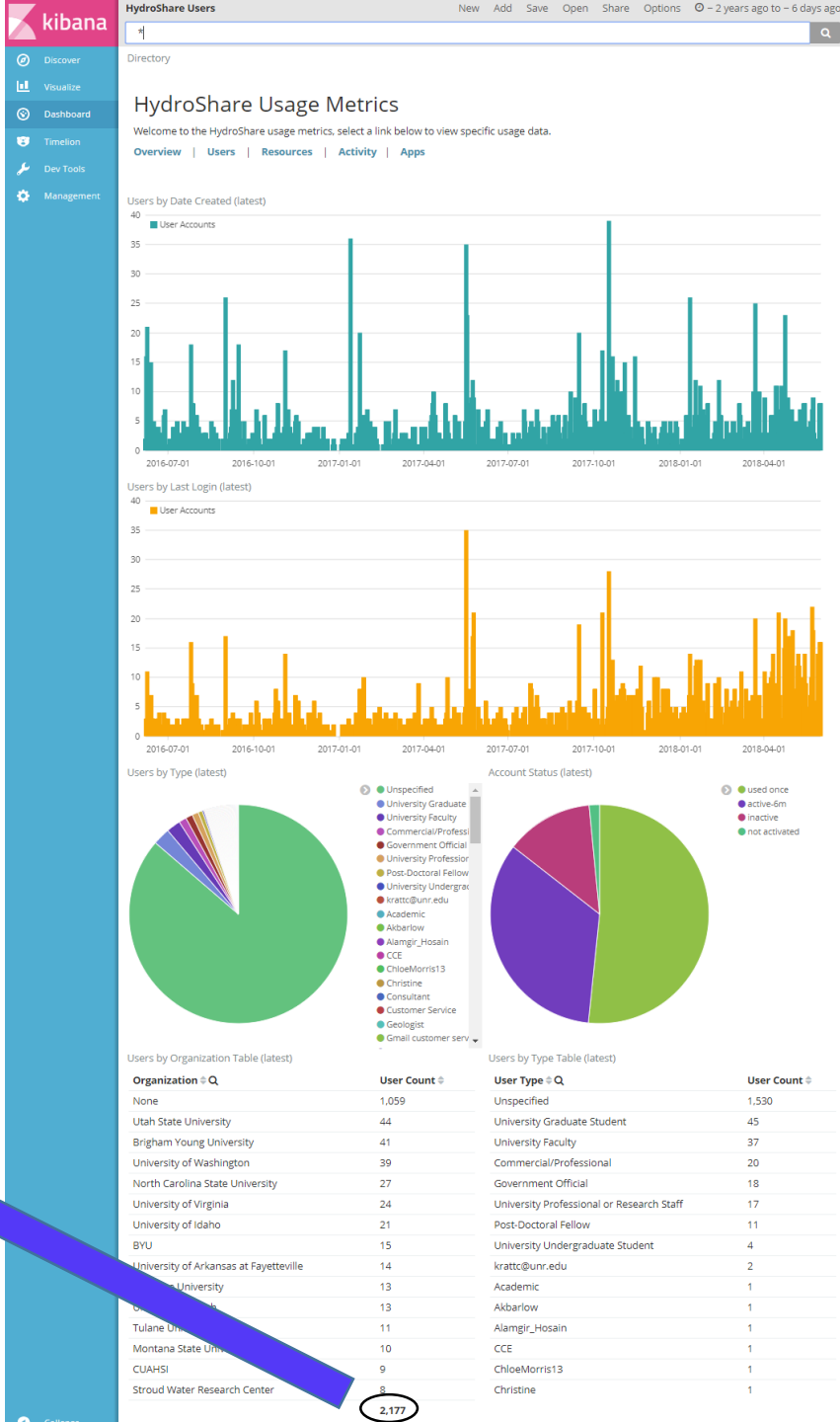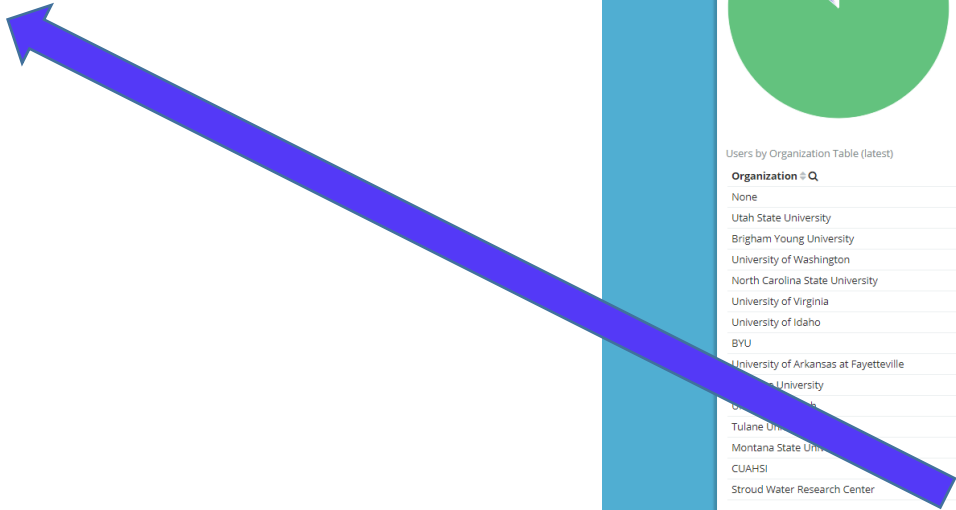
- File Storage

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.

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/

From Dan Ames

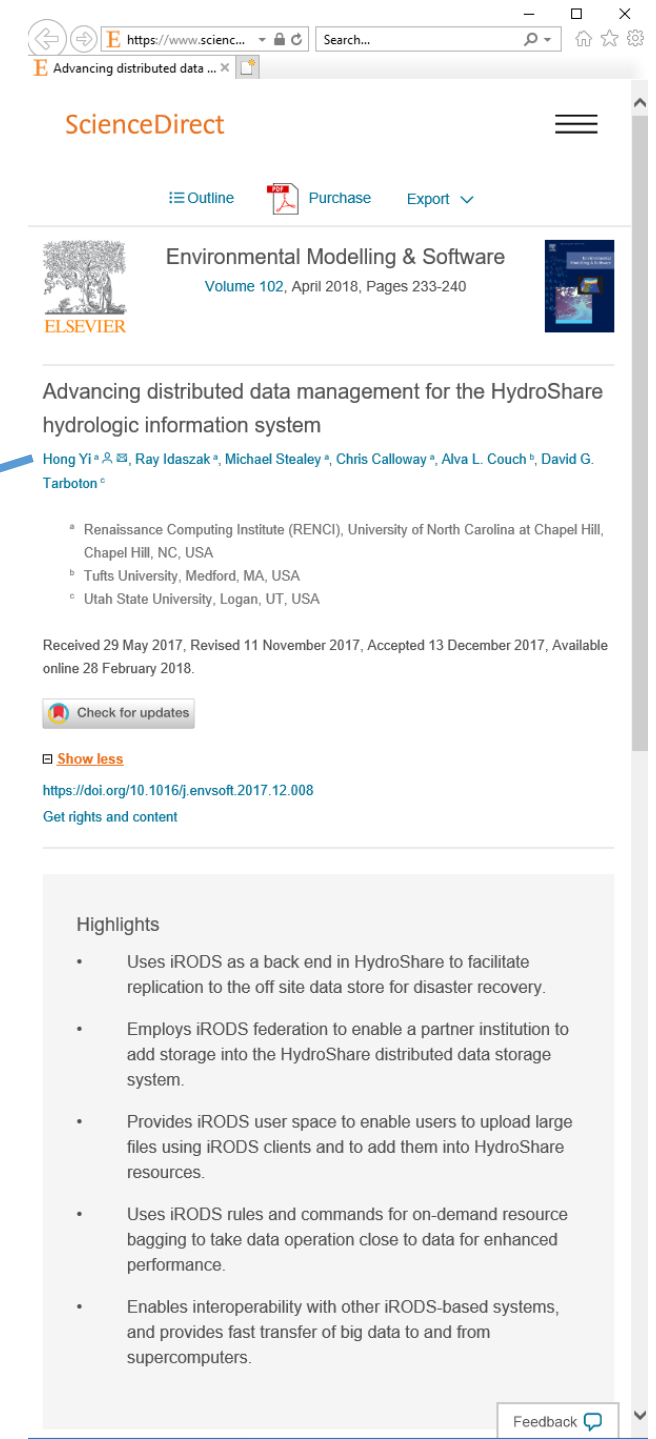# HydroShare Usage Metrics as of June 2018: > 2,000 users

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/hydroshareandirods
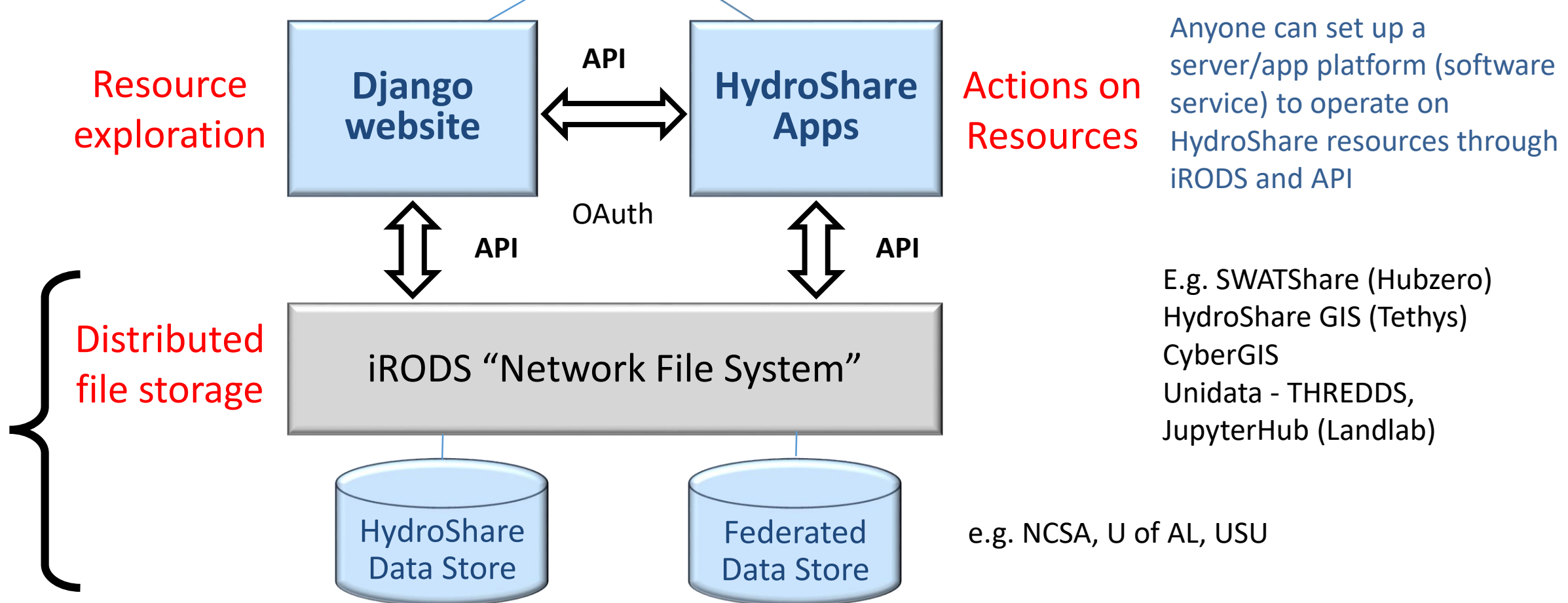
In HydroShare you can:

- 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

- Organize and annotate your content
- Manage access

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

Resource exploration

Django website ⟷ **API** ⟷ **HydroShare Apps**

Actions on Resources

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

OAuth

API   **API**

Distributed file storage

iRODS "Network File System"

E.g. SWATShare (Hubzero)
HydroShare GIS (Tethys)
CyberGIS
Unidata - THREDDS,
JupyterHub (Landlab)

HydroShare Data Store   Federated Data Store
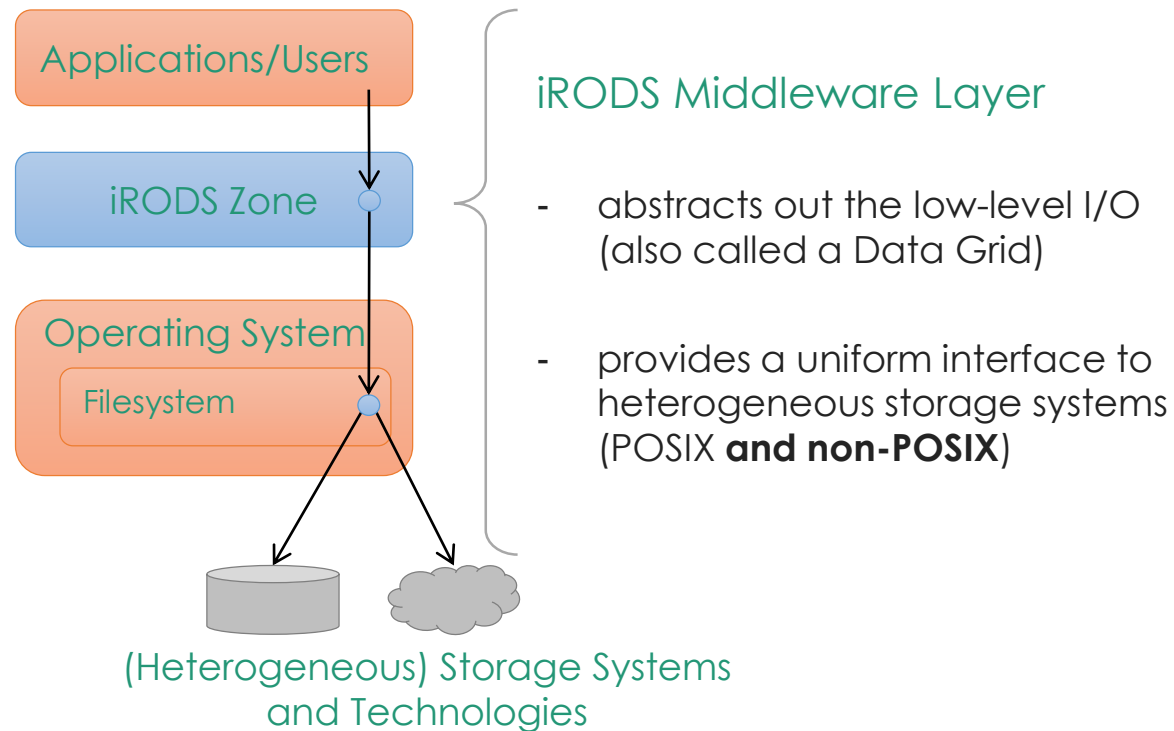
e.g. NCSA, U of AL, USU

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 Middleware Layer

- abstracts out the low-level I/O (also called a Data Grid)

- provides a uniform interface to heterogeneous storage systems (POSIX **and non-POSIX**)

https://irods.org/

IRODS Zone

IRODS provides a virtual system: logical representation of file hierarchies (called Collections) stored in distributed physical storage locations
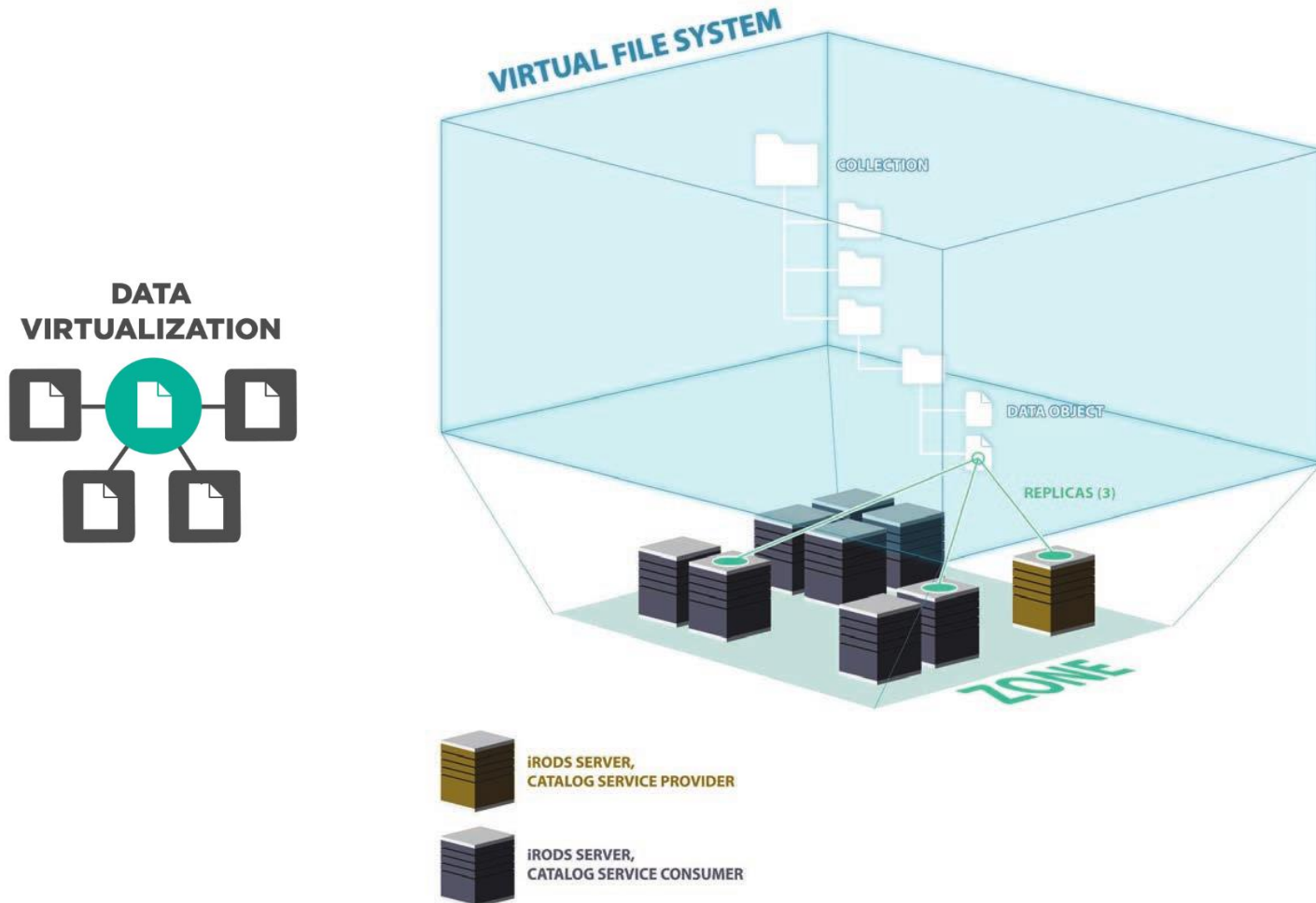
iRODS presents centralizes distributed storage systems under a unified namespace.

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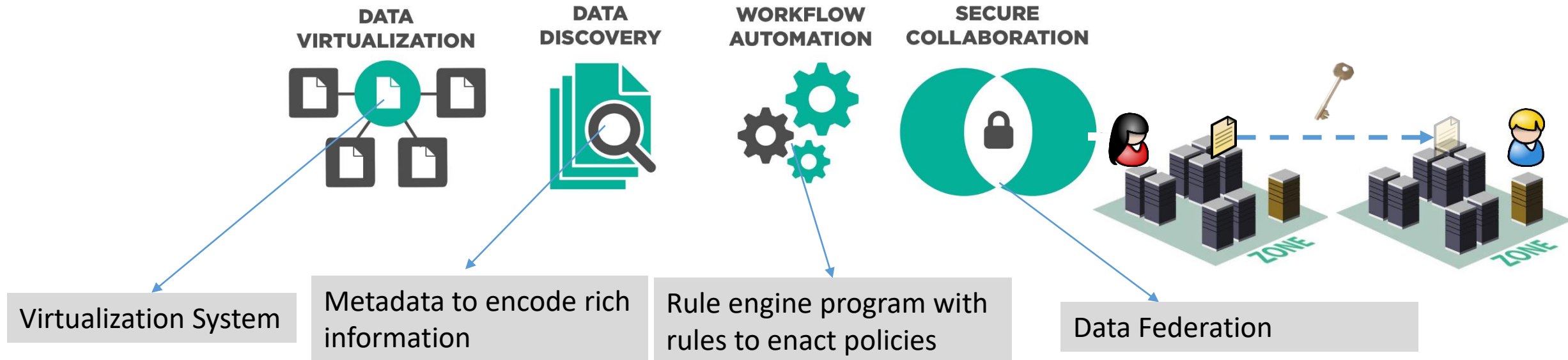
Independent zone can be federated with one another to allow controlled access to remote zones or zones operated by separate workgroups.

iRODS Key Features

The Integrated Rule-Oriented Data System:

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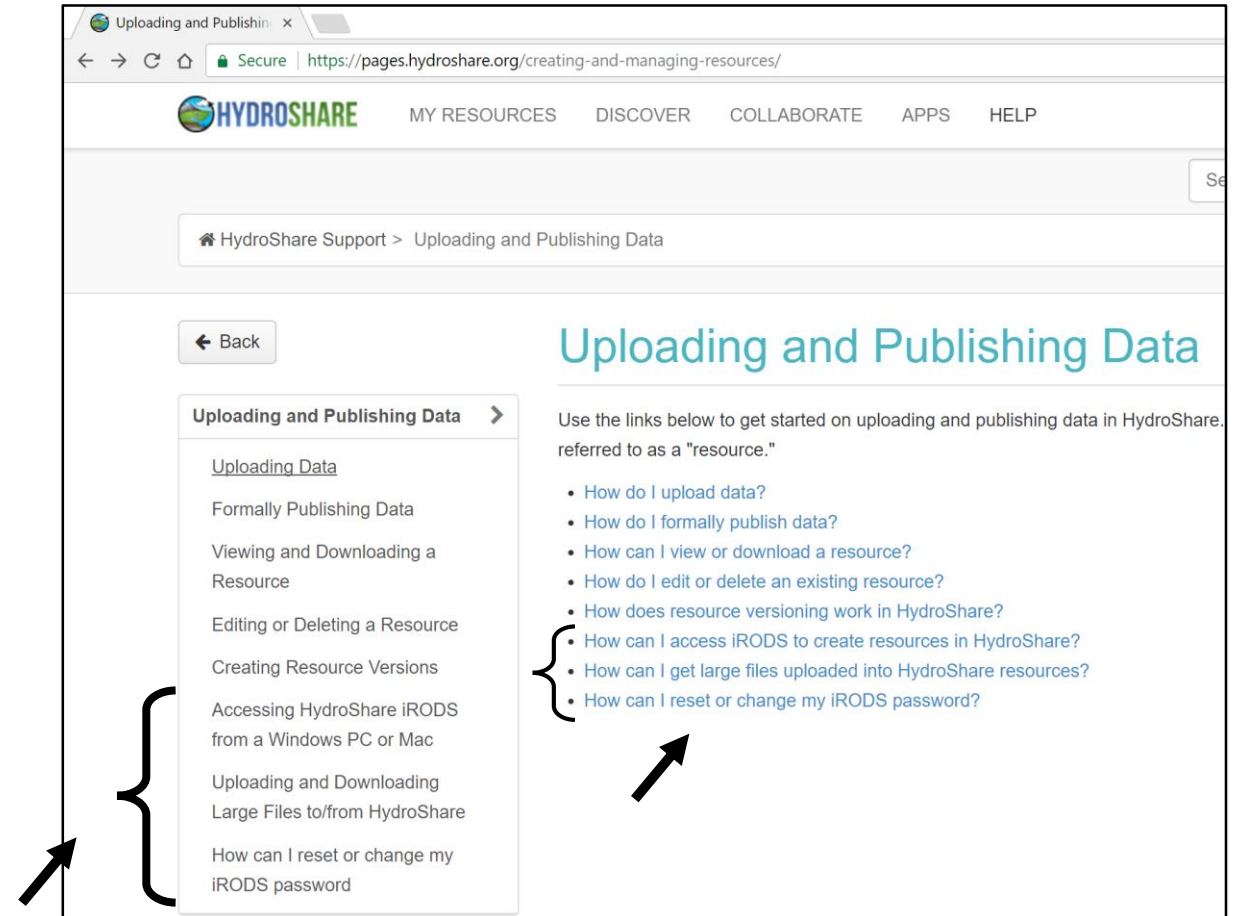- Organizing, securing, preserving, and sharing data

DATA VIRTUALIZATION

DATA DISCOVERY

WORKFLOW AUTOMATION

SECURE COLLABORATION

Virtualization System

Metadata to encode rich information

Rule engine program with rules to enact policies

Data Federation

iRODS in the current HydroShare

- https://help.hydroshare.org/creating-and-managing-resources/
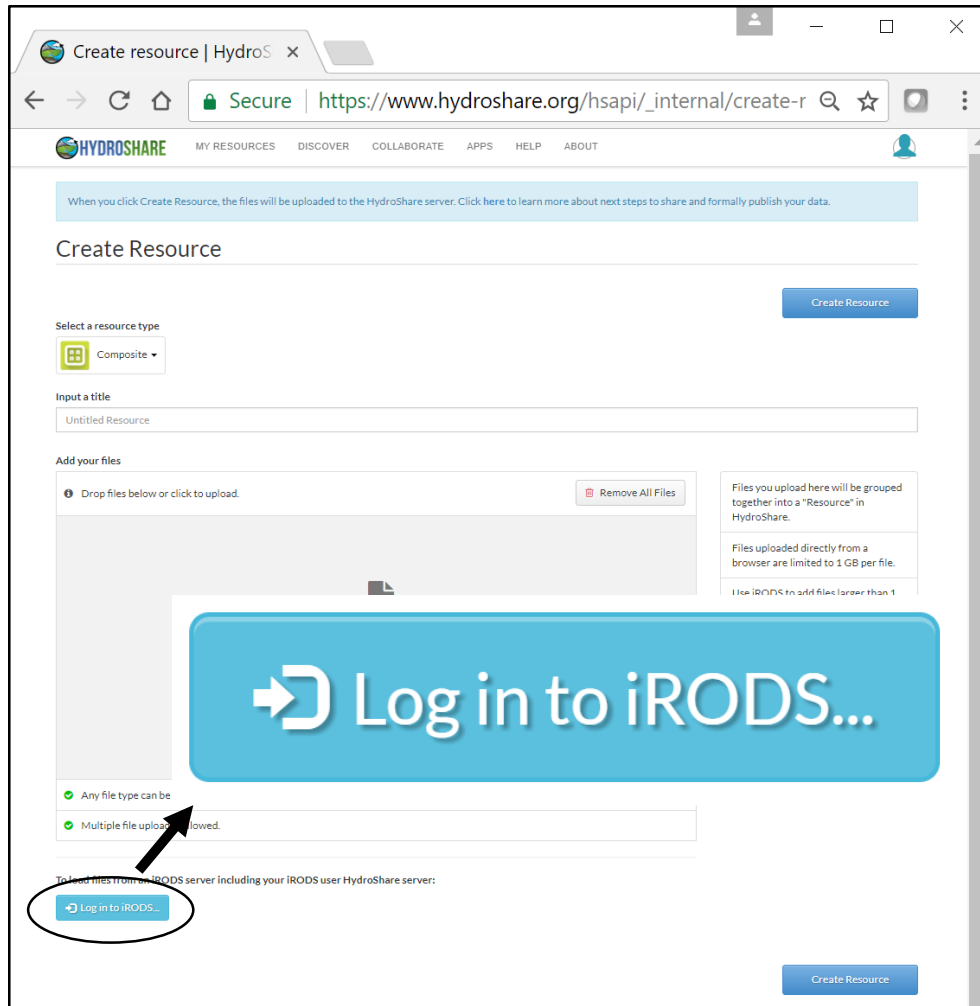
iRODS on the HydroShare resource landing page.

iRODS how-to discussed on the HydroShare Support pages.

NWM Forecast Viewer App

HydroShare: National Water Model Community Data Access Architecture
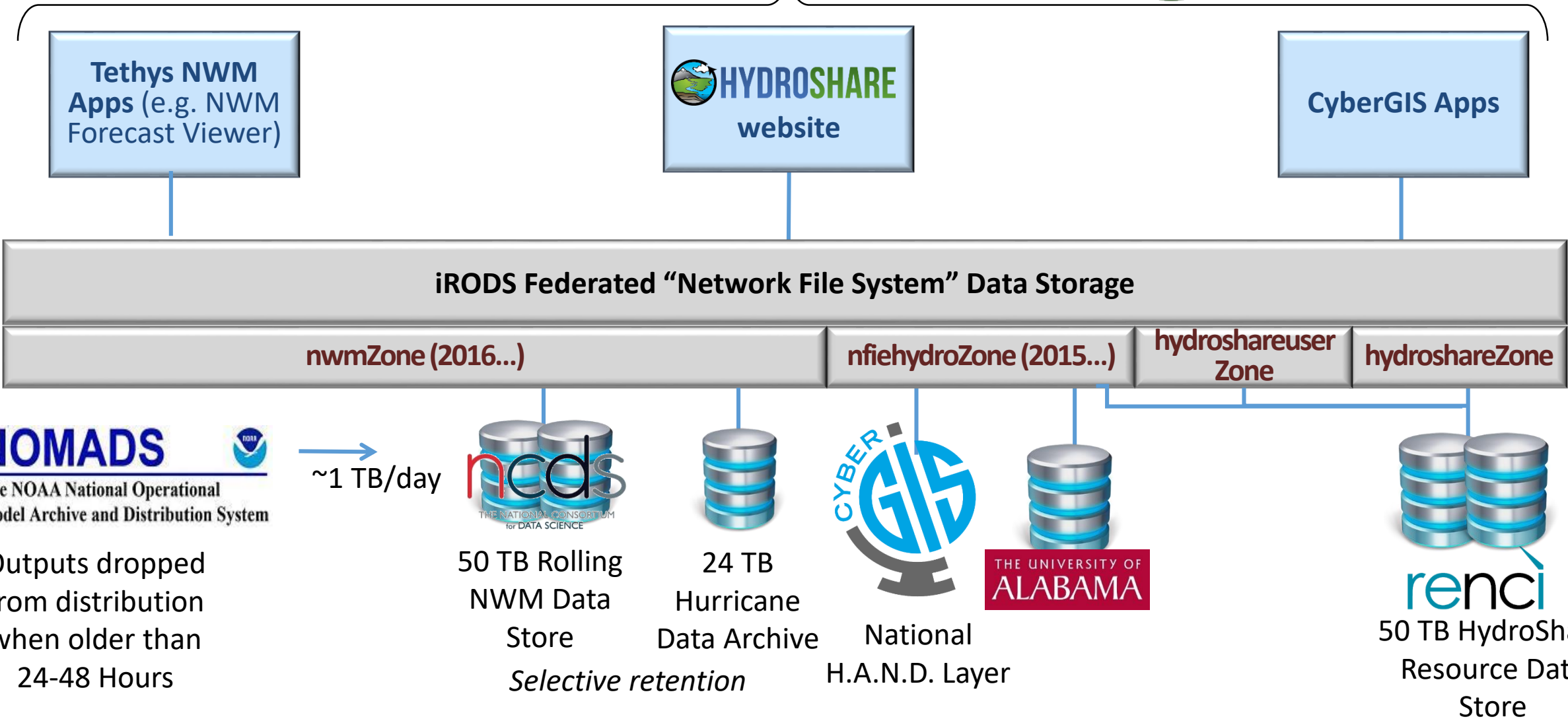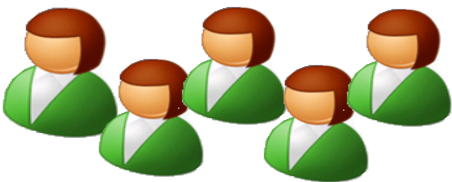
Exploring: HydroShare Extended Storage Ecosystem

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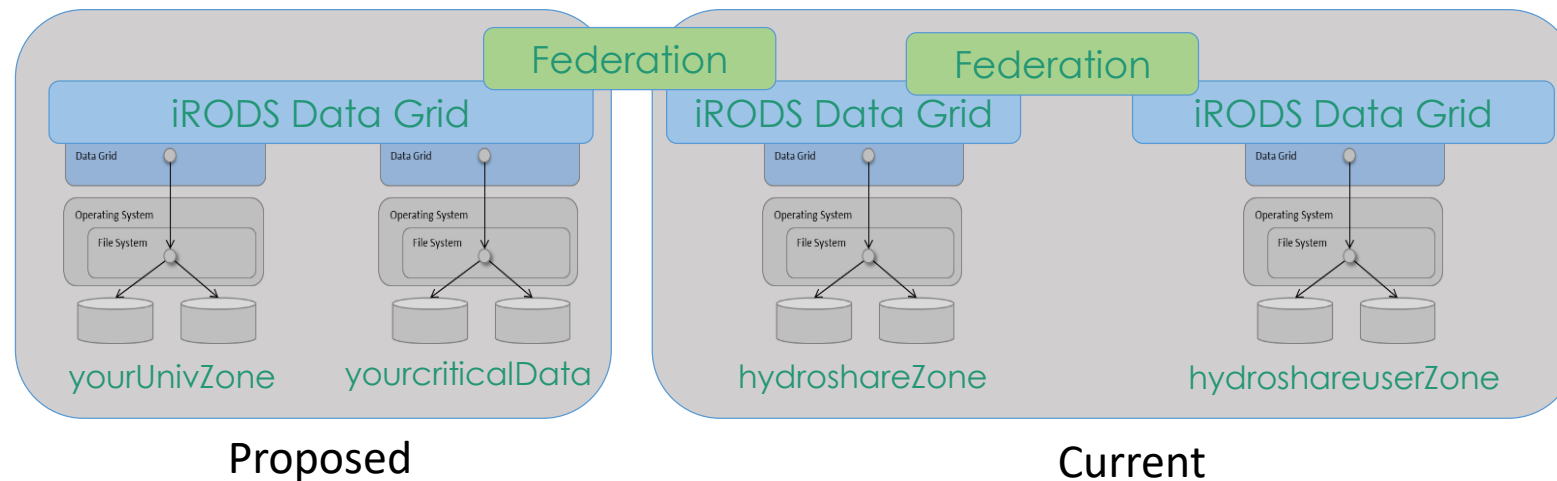
HydroShare: Advancing Collaboration through Hydrologic Data and Model Sharing -- David Tarboton

299 views

Thanks to the HydroShare team!

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

HYDROSHARE

http://www.hydroshare.org