

iRODS User Training Introduction



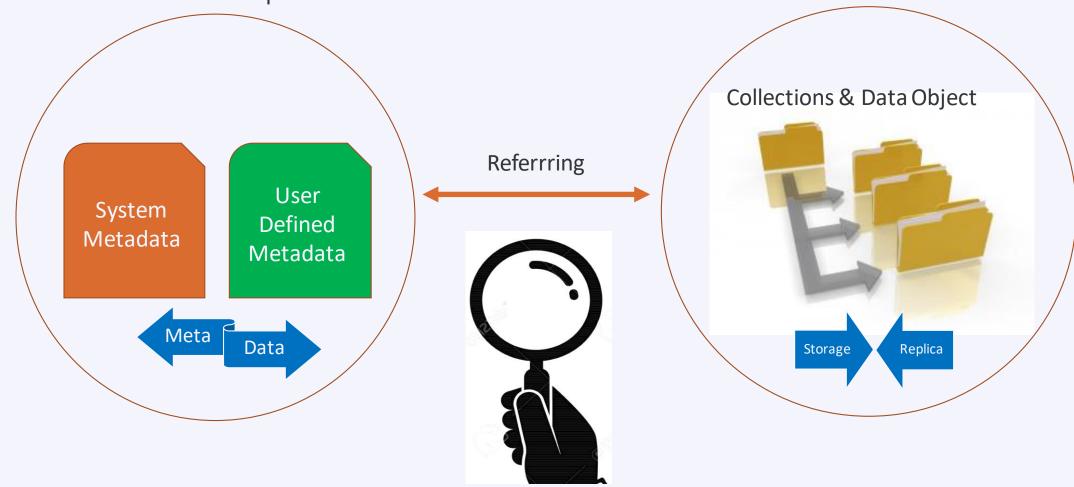
- Introduction
- Data and RDM
- Tier-1 Data platform Architecture
- What is iRODS?
- Functionalities

Introduction

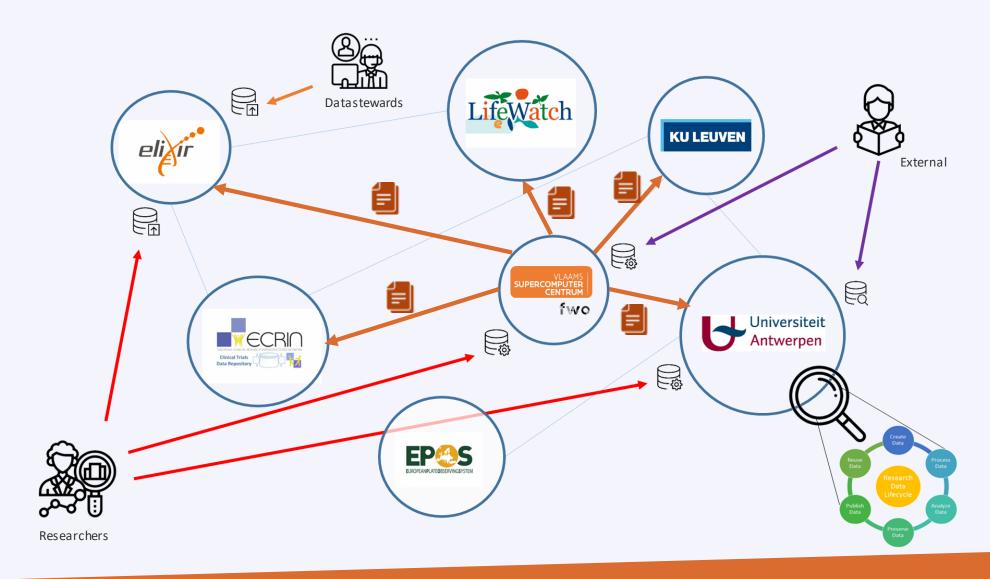
- The aim of the training is to explain basics of Tier-1 Data Service.
- The iRODS training consists of a general introduction, iCommands, VSC-PRC, basic irules and portal clients (yoda, metalnx).
- This training is planned for VSC users.
- It includes hands-on sessions.
- Any questions or feedback will help us improve the quality of the training.

Data and RDM

Data; what is in it? What is the problem with it?



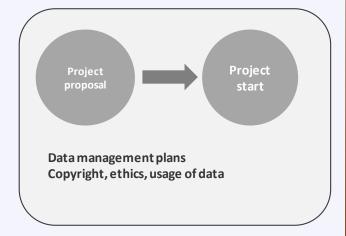
The RDM landscape



Tier-1 Data in the research Data Lifecycle

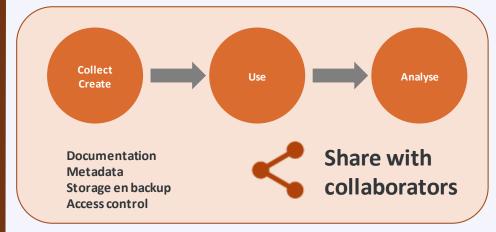


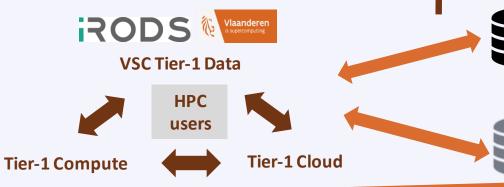
Planning





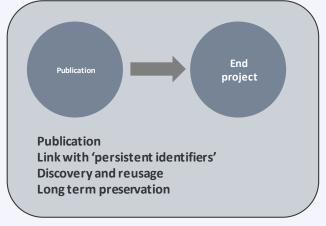
Active Research





Post-publication/Inactive Data

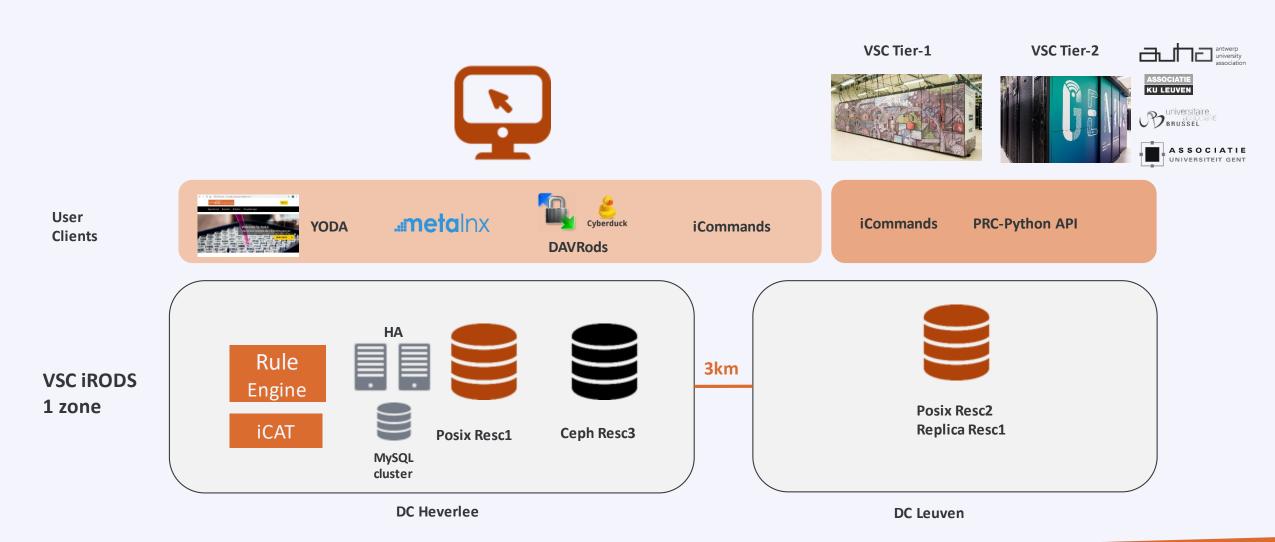
Sharing/Reuse



Regional / domain specific repositories

Institutional repositories

Tier-1 Data architecture



What is iRODS?

- iRODS (integrated Rule-Oriented Data System)
- Open Source distributed data and storage management system
- Configurable data management policies and workflows
- Scalable
- iRODS consortium ensures sustainability by:
 - Guiding further development of the software;
 - Growing the user and developer communities; and
 - Facilitating iRODS support, education, and collaboration opportunities.



iRODS Core competencies



Unified Storage Namespace

Data virtualization of distributed storage systems



Automation

Rule Engine to enforce data polices



Data Discovery

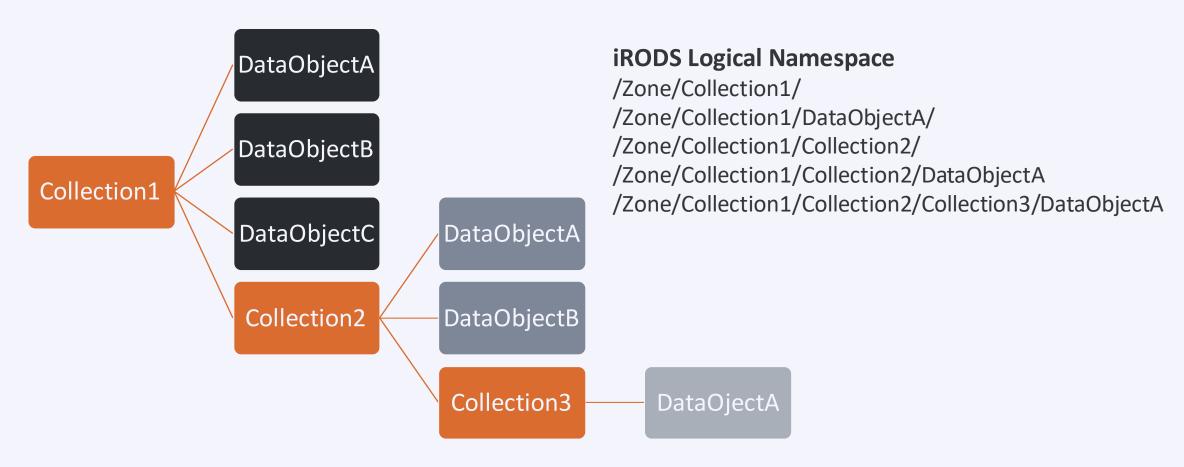
Rich Metadata for collections and data objects (System metadata and user-defined metadata)



Secure collaboration

Three mechanisms: Permissions, Tickets and Federation.

Data organization in iRODS



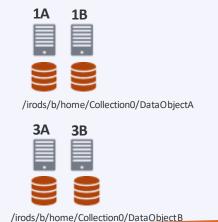
Collections ~ Directories
DataObjects ~ Files

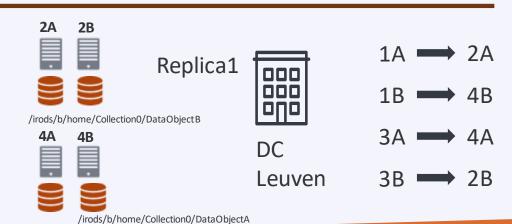
Data virtualization in iRODS

iRODS Logical Namespace DataObjectA /Zone/Collection1/ /Zone/Collection1/DataObjectA/ **DataObjectB** Logical /Zone/Collection1/Collection2/ Collection1 /Zone/Collection1/Collection2/DataObjectA representation DataObjectA DataObjectC /Zone/Collection1/Collection2/Collection3/DataObjectA DataObjectB Collection2 Collection3 **DataOjectA**

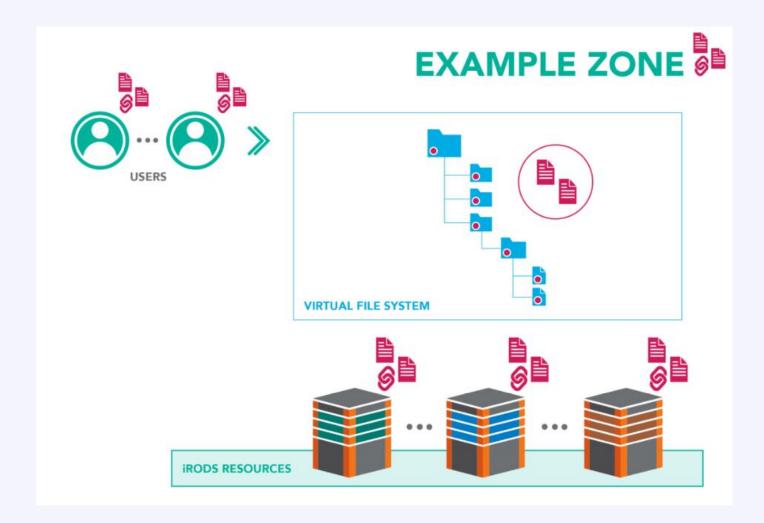
Physical representation







Metadata in iRODS



System Metadata:

• filename, file size, creation date ...

User Metadata:

- Manual introduction
- Metadata templates
- Automation (rules/microservices)



Interaction with iRODS

icommands:

```
iput - iget- irsync -imeta...
```

- uploading/downloading data
- adding metadata to data objects/collections
- querying based on metadata
- deleting data objects/collections
- synchronization of data
- ACLs to data objects/collections

VSC-PRC:

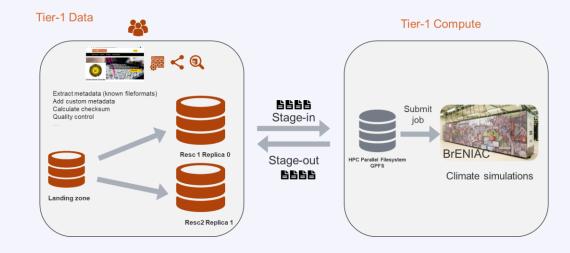
Python3, python-irodsclient

- working with data objects/collections
- adding metadata to data objects/collections
- querying based on metadata
- deleting data objects/collections
- listing the disk usage
- ACLs to data objects/collections

HPC_to_Data:

icommands:

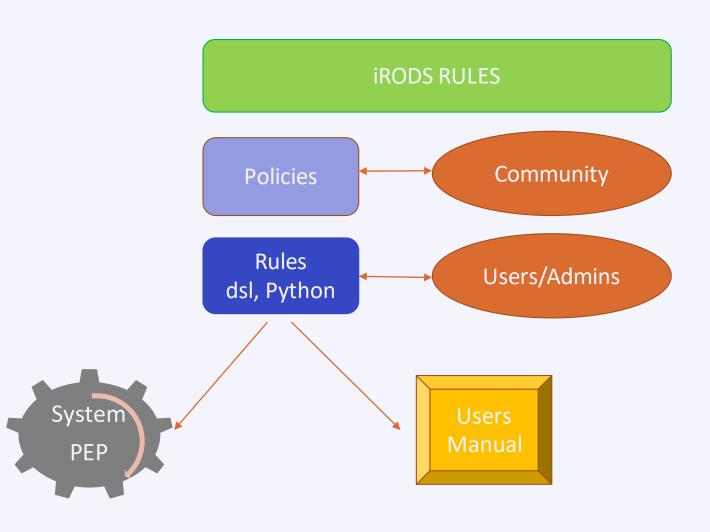
VSC Python client:



- Yoda:
- graphical user interface easiness
- working with data objects/collections
- adding metadata to data collections
- querying visually
- downloading data objects
- group management
- RDM workflow UU

- Metalnx:
- graphical user interface easiness
- working with data objects/collections
- adding metadata to data objects/collections
- downloading data objects
- permission
- iRODS design

- irule:
- user level rules
- written into a local file
- execute it when you need



Documentation and support

Documentation

https://vlaams-supercomputing-centrum-vscdocumentation.readthedocs-hosted.com//sharing/cxplsgyxzaizmf4xg7wl5jj8

NOTE: After you click the special link above, you will reach the latest version of the VSC Documentation which doesn't include "Tier-1 Data Service". Hence you should click the version arrow on the right below side of the incoming page and chose the "data_M" version.

Support

data@vscentrum.be



Questions

