iRODS for Users

ICOMMANDS AND IRODS CONCEPTS – S4R WORKSHOP

305

Christine Staiger





iRODS

- integrated Rule-Oriented Data Systems
- Open source data (and storage) management software
- iRODS consortium:

standardise development and releases

operated at Univ. of North Carolina,

by **RENCI** (Renaissance Computing Institute) and **DICE** (Data Intensive Cyber Environments Centre)



What is iRODS?

You, Researchers, Students, etc.

Want to easily Find, Access, Use, Move, Share Data, and more...
With your Interfaces, your Applications, your Workflows



iRODS Data System - "Middleware"

A "layer" that "connects the dots" while masking and automating your interactions with diverse infrastructure.



The "World of Infrastructure"

Your and other's Storage, Networks, Admin. Domains, Computing Services, Web Services, etc.







THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL











The iRODS system

iRODS Data System Components

User Interface

Web or GUI Client to Access and Manage Data & Metadata*

iRODS Rule
Engine
Implements
Policies

iRODS Metadata
Catalog
Database
Tracks state of data

Data management & requirements

User wants to store data regardless of:

- Size, data sensitivity
- Easy access, data availability
- Not bother too much about details

 technology has to do most

Data manager needs to make sure policies are met:

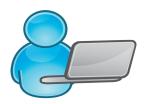
- Data security, safe storage, long-term storage
- Keep costs low → when can data be moved to cheaper media?
- Data is stored on appropriate media (size, availability)

Storage System admins want users to

- Behave! E.g. no small files on my system please.
- Use the systems efficiently!



The iRODS solution



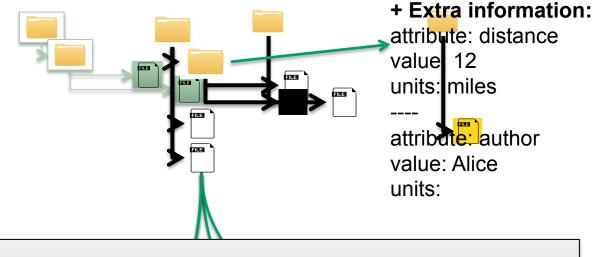
Abstraction layer:

Mapping from logical to physical namespace

Storage layer:

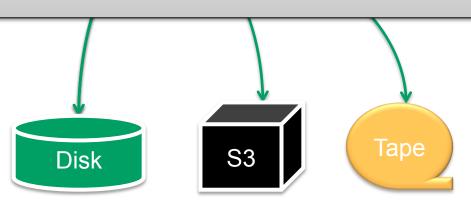
Different storage media

Different protocols to steer data



iCAT (iRODS metadata catalogue):

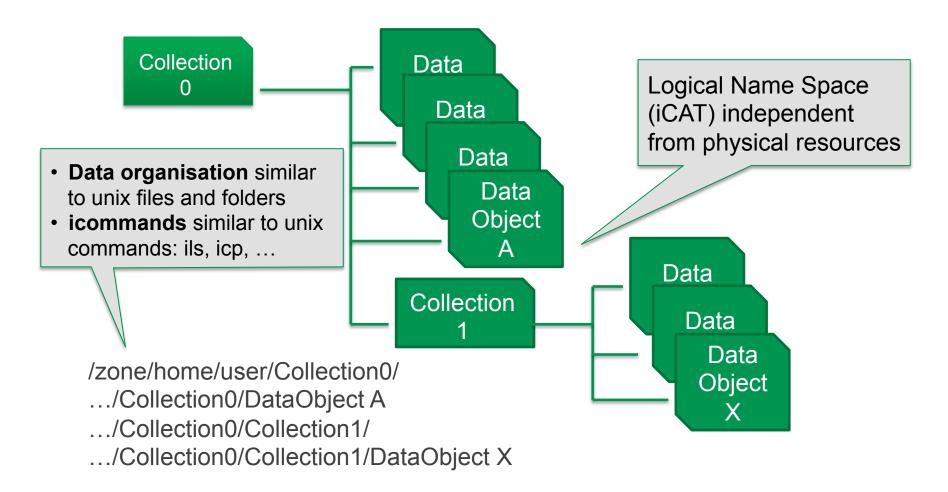
/irodsZone/home/user/Collection0/testfile.txt → /irodsVault/home/rods/testfile.txt







The users view: iRODS Data Collections

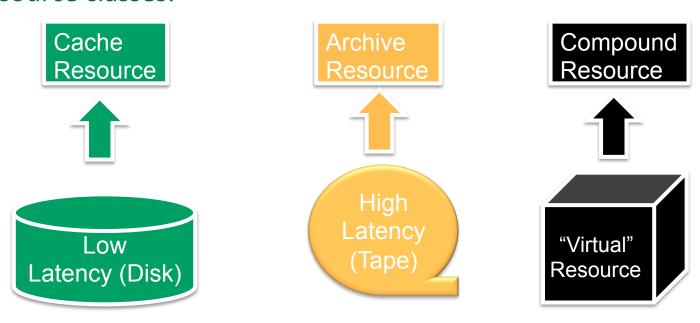


What does iRODS provide?

- Storage virtualization of different disk and tape storage systems
- A logical namespace across storage locations
- A rule engine to automate data management according to defined policies
- Federations between iRODS instances
- Various client interfaces

In the Background: iRODS resources

- (Storage) Resource is a Software or Hardware system that stores data
- 3 Resource classes:



• Resource types: unix/spss file system, s3, MSS opendap, tds (THREDDS), ...



Resource groups

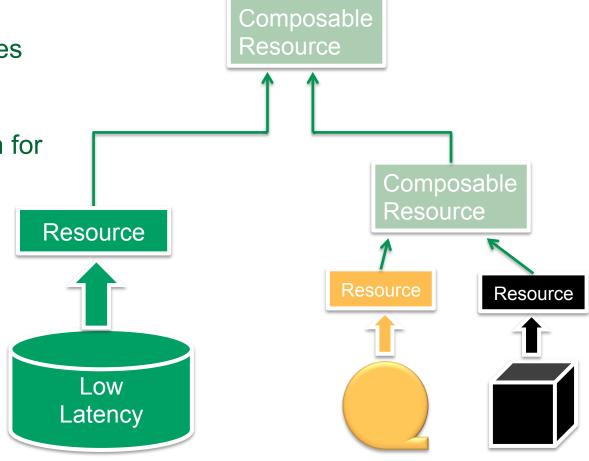
Composable resource Types:

Replication
 synchronise resources

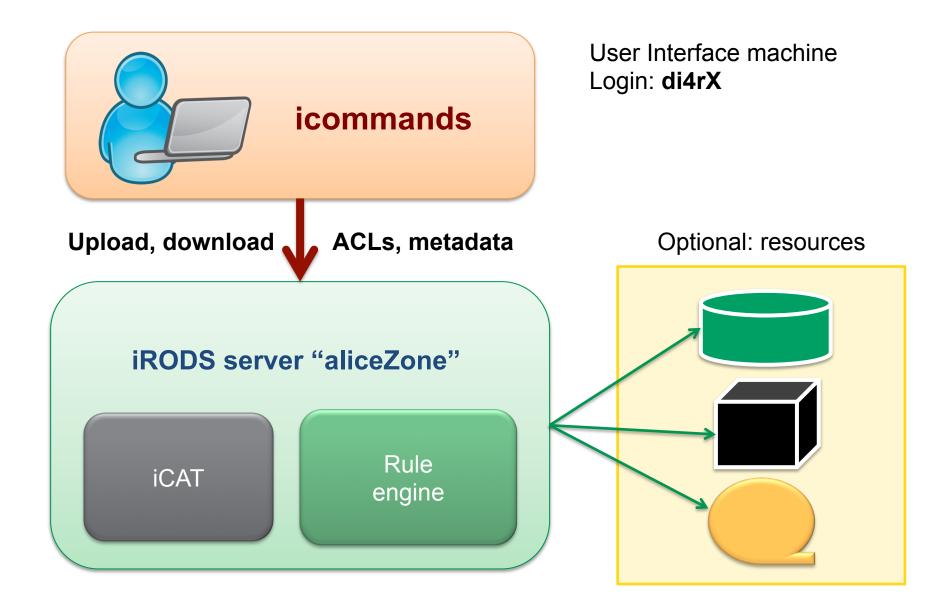
Round Robin
 rotate through children for uploading

Load balance

 Compound resource cache resource and archive resource



Training Setup



Thank you!

Robert Verkerk, SURFsara, NL Claudio Cacciari, CINECA, IT Lykle Voort, SURFsara, NL

Jeroen Schot, SURFsara, NL

HANDS-ON!



