iRODS: integrated Rule-Oriented Data System

Data grid middleware supporting interoperability Software infrastructure for implementation of data policy

Data Grid Middleware

iRODS Unified Virtual Collection iRODS View of Distributed Data User Client User sees a single collection My Data: disk, filesystem, WOS storage unit... My Data: tape, database, filesystem... Peartner's Data remote disk, tape, filesystem...

- iRODS installs over heterogeneous data resources
- Users share & manage distributed data as a single collection

Data Management

Issues in Data Management

- · Data organization and Usage
- Interfaces to Data
- · Audit Tracking/Accounting
- · Data Services

Managing distributed collections, distributed access, controlled sharing, metadata collection and tagging, back-ups and replicas, data integrity & authenticity verification, provenance, discoverability (metadata management and query support), virus checks, format conversion, evolution of usage across data life cycle, long-term preservation, etc.

Extensible Data Management Infrastructure

- Abstract out the data management: independence from storage device, location, administrative domain
- Unit of work: microservice (C code)
- · Rules: workflows of microservices
- Modules: libraries of microservices (added, sometimes by users, for custom services)
- Policy-driven management: event-triggered execution of rules

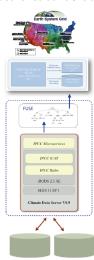
Access Interface Access Interface Policy Enforcement Points Standard Micro-services Map from policy to standard m Map from micro-services to sta



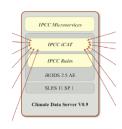
Standard I/O Operations

Interoperability: A NASA Case Study*

iRODS FUSE client to expose repositories as ESG data



NetCDF with iRODS

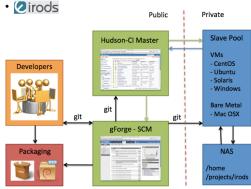


With an iRODS-based Climate Data Server, all of the IPCC NetCDF metadata is externalized through the iCAT database in accordance with the OAIS Reference Model ...

Header metadata extracted into iRODS DB to support search and discoverability

Enterprise iRODS (E-iRODS) at RENCI**

- Hardened binary release of community iRODS (with a delay) – RedHat model.
- Testing, packaging, documentation, support
- Beta release available: http://e-irods.com/



RENCI E-iRODS Test Environment



^{*}Special thanks to John Schnase's group at NCCS for permission to cite work and reproduce images.

^{**}Special thanks to Jason Coposky's group at RENCI for permission to reproduce images.