Policy Based Data Management

iRODS

Reagan W. Moore (DICE-UNC)
Arcot Rajasekar (DICE-UNC)
Mike Wan (DICE-UCSD)
Wayne Schroeder (DICE-UCSD)
Mike Conway (DICE-UNC)
Antoine de Torcy (DICE-UNC)
Hao Xu (UNC)
Sheau-Yen Chen (DICE-UCSD)
Jason Coposky (RENCI)
Lisa Stillwell (RENCI)
Leesa Brieger (RENCI)

http://irods.diceresearch.org
Agenda

• 09:00 Introduction to iRODS technology – (Moore, UNC-CH)
• 10:00 Open discussion on data management needs – (Moore, UNC-CH)
• 10:30 Break
• 11:00 Installation of iRODS clients (Windows, Mac, Linux) – (Moore UNC-CH)
• 11:45 Initial demonstration of iRODS client access – (Moore, UNC-CH)
• 12:00 Lunch
• 14:00 Introduction to iRODS rules and micro-services – (Moore, UNC-CH)
• 14:45 Demonstration of federation of data grids - (Yutaka Kawai, KEK, Japan, Moore UNC-CH)
• 15:00 Simple rule examples including database queries (Moore, UNC-CH)
• 15:30 Break
• 16:00 Complex rule examples including scheduling (Moore, UNC-CH)
• 16:30 SRM development status– (Wei-Long UENG, ASGC)
• 17:00 Open discussion of iRODS applications in digital libraries, data grids, and preservation environments – (Moore, UNC-CH)
• 17:30 End of tutorial
Policy-based Data Sharing

Consensus on Policies and Procedures controls the shared data.

Client

Provider
iRODS controlled procedures
Storage

Provider
iRODS controlled procedures
Storage

Logical Shared Collection
Applications

• International projects
  – Cyber Square Kilometer Array (radio astronomy), Cinegrid (movies)

• National data grids
  – Australia, New Zealand, Portugal, UK, France

• Federal agency archives
  – NASA Center for Climate Simulation, National Optical Astronomy Observatories

• Institutional repositories
  – French National Library, Carolina digital repository, Broad Institute genomics data, Sanger Institute
iRODS Version 3.0

• Released Sept. 30, 2011
• Next release scheduled for March, 2012
• New features
  – New rule engine
    • Strong parameter typing
    • Optimized performance (thousands of rules)
    • Expanded rule programming language
  – Rule versioning
  – Distributed rule base management
  – Soft links to external resources
  – New transport management for large files
  – Improved Java interface
  – Windows native port (C++ compilation)
  – Dropbox interface (iDrop)
Under Development

• Tickets (3.1)
  – Ability to give access to specified file for specified time duration

• netCDF library support (3.1)

• Pluggable Authentication Modules (3.2)
  – Use external identity resource

• Registration of workflows (3.2)
  – Micro-service structured objects
  – Re-execute workflows for derived data products
  – Policy-encoded objects
Policy-based Data Environments

- **Purpose** - reason a collection is assembled
- **Properties** - attributes needed to ensure the purpose
- **Policies** - controls for enforcing desired properties,
  - mapped to computer actionable rules
- **Procedures** - functions that implement the policies
  - mapped to computer actionable workflows
- **Persistent state information** - results of applying the procedures
  - mapped to system metadata
- **Assessment criteria** - validation that state information conforms to the desired purpose
  - mapped to periodically executed policies
Overview of iRODS Architecture

User w/Client
Can Search, Access, Add and Manage Data & Metadata

Access distributed data with Web-based Browser or iRODS GUI or Command Line clients.

iRODS Middleware

iRODS Data Server
Disk, Tape, etc.

iRODS Rule Engine
Track Policies

iRODS Metadata Catalog
Track information
Data Management Applications

Generic Infrastructure

- Data grids – PB-size distributed collections
  - Astronomy – NOAO, CyberSKA, LSST
  - High Energy Physics – BaBar, KEK
  - Earth Systems – NASA (MODIS data set)
  - Australian Research Collaboration Service
  - Genomics – UNC-CH/RENCI
- Institutional repositories
  - Carolina Digital Repository
- Libraries
  - Texas Digital Libraries
  - Seismology - Southern California Earthquake Center
- Archives
  - Ocean Observatories Initiative
Data Virtualization

Access Interface

Policy Enforcement Points

Standard Micro-services

Standard I/O Operations

Storage Protocol

Storage System

Map from the actions requested by the client to multiple policy enforcement points.

Map from policy to standard micro-services.

Map from micro-services to standard Posix I/O operations.

Map standard I/O operations to the protocol supported by the storage system.
## Data Grid

### Clients (48)

<table>
<thead>
<tr>
<th>API</th>
<th>Client</th>
<th>Developer</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Browser</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DCAPe</td>
<td>UNC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iExplore</td>
<td>RENCI-Oleg</td>
<td>C++</td>
</tr>
<tr>
<td></td>
<td>JUX</td>
<td>IN2P3</td>
<td>Jargon</td>
</tr>
<tr>
<td></td>
<td>Peta Web browser</td>
<td>PetaShare</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iDrop web browser</td>
<td>Mike Conway</td>
<td>Java</td>
</tr>
<tr>
<td></td>
<td>Davis web interface</td>
<td>ARCS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rich web client</td>
<td>Lisa Stillwell - RENCI</td>
<td></td>
</tr>
<tr>
<td><strong>Digital Library</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Akubra/iRODS</td>
<td>DICE</td>
<td>Jargon</td>
</tr>
<tr>
<td></td>
<td>Dspace</td>
<td>MIT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fedora on Fuse</td>
<td>IN2P3</td>
<td>FUSE</td>
</tr>
<tr>
<td></td>
<td>Fedora/iRODS module</td>
<td>DICE</td>
<td>Jargon</td>
</tr>
<tr>
<td></td>
<td>Islandora</td>
<td>DICE</td>
<td>Jargon</td>
</tr>
<tr>
<td></td>
<td>Curators Workbench</td>
<td>CDR-UNC-CH</td>
<td>Jargon</td>
</tr>
<tr>
<td><strong>File System</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Davis - Webdav</td>
<td>ARCS</td>
<td>Jargon</td>
</tr>
<tr>
<td></td>
<td>Dropbox / iDrop</td>
<td>DICE-Mike Conway</td>
<td>Jargon</td>
</tr>
<tr>
<td></td>
<td>FUSE</td>
<td>IN2P3, DICE,</td>
<td>FUSE</td>
</tr>
<tr>
<td></td>
<td>FUSE optimization</td>
<td>PetaShare</td>
<td>FUSE</td>
</tr>
<tr>
<td></td>
<td>OpenDAP</td>
<td>ARCS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PetaFS (Fuse)</td>
<td>Petashare - LSU</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Petashell (Parrot)</td>
<td>PetaShare</td>
<td></td>
</tr>
<tr>
<td>Grid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>GridFTP - Griffin</td>
<td>ARCS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jsaga</td>
<td>IN2P3</td>
<td></td>
<td>Jargon</td>
</tr>
<tr>
<td>Parrot</td>
<td>UND - Doug Thain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRM</td>
<td>Academia Sinica</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saga</td>
<td>KEK</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I/O Libraries</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PRODS - PHP</td>
<td>Renci - Lisa Stillwell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C API</td>
<td>DICE-Mike Wan</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>C I/O library</td>
<td>DICE-Wayne Schroeder</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Fortran</td>
<td>Schroeder</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Eclipse file system</td>
<td>CDR - UNC-CH</td>
<td>Jargon</td>
<td></td>
</tr>
<tr>
<td>Jargon</td>
<td>DICE-Mike Conway</td>
<td>Jargon</td>
<td></td>
</tr>
<tr>
<td>Pyrods - Python</td>
<td>SHAMAN-Jerome Fusillier</td>
<td>Python</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Portal</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>EnginFrame</td>
<td>NICE / RENCI</td>
<td>Jargon</td>
<td></td>
</tr>
<tr>
<td>Petashare Portal</td>
<td>LSU</td>
<td>Jargon</td>
<td></td>
</tr>
</tbody>
</table>
## iRODS Clients (Cont.)

<table>
<thead>
<tr>
<th>Tools</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Archive tools - NOAO</td>
<td>NOAO</td>
<td></td>
</tr>
<tr>
<td>Big Board visualization</td>
<td>RENCI</td>
<td></td>
</tr>
<tr>
<td>iFile</td>
<td>GA Tech</td>
<td></td>
</tr>
<tr>
<td>i-commands</td>
<td>DICE</td>
<td></td>
</tr>
<tr>
<td>Pcommands</td>
<td>PetaShare</td>
<td></td>
</tr>
<tr>
<td>Resource Monitoring</td>
<td>IN2P3</td>
<td></td>
</tr>
<tr>
<td>Sync-package</td>
<td>Academica Sinica</td>
<td></td>
</tr>
<tr>
<td>URSpace</td>
<td>Teldap - Academica Sinica</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Web Service</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VOSpace</td>
<td>IVOA</td>
<td></td>
</tr>
<tr>
<td>Shibboleth</td>
<td>King's College</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Workflows</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kepler - actor</td>
<td>DICE</td>
<td>Jargon</td>
</tr>
<tr>
<td>Stork - interoperability</td>
<td>LSU</td>
<td></td>
</tr>
<tr>
<td>Workflow Virtualization</td>
<td>LSU</td>
<td></td>
</tr>
<tr>
<td>Taverna - actor</td>
<td>RENCI</td>
<td></td>
</tr>
</tbody>
</table>
Policy Enforcement Points

• Currently have 74 locations within iRODS framework where policies are checked.
  – Each action may involve multiple policy enforcements points

• Policy enforcement points
  – Pre-action policy  (selection of storage location)
  – Policy execution  (file deletion control)
  – Post-action policy  (derived data products)
| icommands | none | acChkHostAccessControl | acSetUserPolicy | acSetRescSchemeForCreate | acSetRescQuotaPolicy | acSetVaultPathPolicy | acPreProcForModifyDataObjMeta | acPostProcForModifyDataObjMeta | acPostProcForDataObjOpen | acPostProcForMultiReplResc | acPostProcForCreate | acPostProcForPut | acPostProcForCopy | acPostProcForRepl | acPostProcForPhymv | acPreProcForObjRename | acPostProcForObjRename | acPreProcForRmColl | acTrashPolicy | acDataDeletePolicy |
|-----------|------|------------------------|----------------|--------------------------|----------------------|---------------------|-----------------------------|------------------------------|--------------------------|------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| icp       | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| icp -N 2  | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| iphybun   | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| irepl     | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| ibun -cD  | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| iput      | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| iphymv    | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| imv       | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| irm       | X    | X                      | X             | X                        | X                    | X                   | X                           | X                           | X                        | X                      | X              | X              | X              | X              | X              | X              | X              | X              | X              |
| irm -r collection | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
Policy Enforcement Points (74)

**ACTION**
- acCreateUser
- acDeleteUser
- acGetUserbyDN
- acTrashPolicy
- acAclPolicy
- acSetCreateConditions
- acDataDeletePolicy
- acRenameLocalZone
- acSetRescSchemeForCreate
- acRescQuotaPolicy
- acSetMultiReplPerResc
- acSetNumThreads
- acVacuum
- acSetResourceList
- acSetCopyNumber
- acVerifyChecksum
- acCreateUserZoneCollections
- acDeleteUserZoneCollections
- acPurgeFiles
- acRegisterData
- acGetIcatResults
- acSetPublicUserPolicy
- acCreateDefaultCollections
- acDeleteDefaultCollections

**PRE-ACTION POLICY**
- acPreProcForCreateUser
- acPreProcForDeleteUser
- acPreProcForModifyUser
- acPreProcForModifyUserGroup
- acChkHostAccessControl
- acPreProcForCollCreate
- acPreProcForRmColl
- acPreProcForModifyAVUMetadata
- acPreProcForModifyCollMeta
- acPreProcForModifyDataObjMeta
- acPreProcForModifyAccessControl
- acPreProcForDataObjOpen
- acPreProcForObjRename
- acPreProcForCreateResource
- acPreProcForDeleteResource
- acPreProcForModifyResource
- acPreProcForModifyResourceGroup
- acPreProcForModifyToken
- acPreProcForDeleteToken
- acNoChkFilePathPerm
- acPreProcForGenQuery
- acSetReServerNumProc
- acSetVaultPathPolicy

**POST-ACTION POLICY**
- acPostProcForCreateUser
- acPostProcForDeleteUser
- acPostProcForModifyUser
- acPostProcForModifyUserGroup
- acPostProcForDelete
- acPostProcForCollCreate
- acPostProcForRmColl
- acPostProcForModifyAVUMetadata
- acPostProcForModifyCollMeta
- acPostProcForModifyDataObjMeta
- acPostProcForModifyAccessControl
- acPostProcForOpen
- acPostProcForObjRename
- acPostProcForCreateResource
- acPostProcForDeleteResource
- acPostProcForModifyResource
- acPostProcForModifyResourceGroup
- acPostProcForModifyToken
- acPostProcForDeleteToken
- acPostProcForFilePathReg
- acPostProcForGenQuery
- acPostProcForPut
- acPostProcForCopy
- acPostProcForCreate
iput ../src/irm.c checks 10 policy hooks

srbbrick14:10900:ApplyRule#116:: acChkHostAccessControl
srbbrick14:10900:GotRule#117:: acChkHostAccessControl
srbbrick14:10900:ApplyRule#118:: acSetPublicUserPolicy
srbbrick14:10900:GotRule#119:: acSetPublicUserPolicy
srbbrick14:10900:ApplyRule#120:: acAclPolicy
srbbrick14:10900:GotRule#121:: acAclPolicy
srbbrick14:10900:ApplyRule#122:: acSetRescSchemeForCreate
srbbrick14:10900:GotRule#123:: acSetRescSchemeForCreate
srbbrick14:10900:execMicroSrvc#124:: msiSetDefaultResc(demoResc,null)
srbbrick14:10900:ApplyRule#125:: acRescQuotaPolicy
srbbrick14:10900:GotRule#126:: acRescQuotaPolicy
srbbrick14:10900:execMicroSrvc#127:: msiSetRescQuotaPolicy(off)
srbbrick14:10900:ApplyRule#128:: acSetVaultPathPolicy
srbbrick14:10900:GotRule#129:: acSetVaultPathPolicy
srbbrick14:10900:execMicroSrvc#130:: msiSetGraftPathScheme(no,1)
srbbrick14:10900:ApplyRule#131:: acPreProcForModifyDataObjMeta
srbbrick14:10900:GotRule#132:: acPreProcForModifyDataObjMeta
srbbrick14:10900:ApplyRule#133:: acPostProcForModifyDataObjMeta
srbbrick14:10900:GotRule#134:: acPostProcForModifyDataObjMeta
srbbrick14:10900:ApplyRule#135:: acPostProcForCreate
srbbrick14:10900:GotRule#136:: acPostProcForCreate
srbbrick14:10900:ApplyRule#137:: acPostProcForPut
srbbrick14:10900:GotRule#138:: acPostProcForPut
srbbrick14:10900:GotRule#139:: acPostProcForPut
srbbrick14:10900:GotRule#140:: acPostProcForPut
Actionable Rule

• Name | Constraint | Workflow | Recovery

• Associate name with
  – Policy enforcement point
  – User defined rule

• Constraint is a test on any
  – Session variable
  – Persistent state attribute

• Workflow composed by chaining micro-services

• Recovery workflow composed from micro-services
Hello World

myTestRule {
    # Rule to print out "Hello World"
    writeLine("stdout", "Hello World");
}

INPUT null
OUTPUT ruleExecOut
Policies – actionable rules

- Retention, disposition, distribution, arrangement
- Authenticity, provenance, description
- Integrity, replication, synchronization
- Deletion, trash cans, versioning
- Archiving, staging, caching
- Authentication, authorization, redaction
- Access, approval, IRB, audit trails, report generation
- Assessment criteria, validation
- Derived data product generation, format parsing
- Federation of independent data grids
Highly Controlled Environment

• All accesses are authenticated
  – GSI / Kerberos / Challenge-response / Shibboleth

• All operations are authorized
  – ACLs on files, storage
  – User groups, storage groups

• All policies evaluate a constraint
  – Constraints based on persistent state information and session information
iRODS Extensible Infrastructure

- **Clients** – specific to discipline and life cycle state
- **Policies** – specific to discipline
- **Procedures** – specific to discipline
- **Remaining infrastructure is generic**
  - Network transport
  - Authentication / Authorization
  - Distributed storage access
  - Remote execution
  - Metadata management
  - Message passing / distributed debugging
  - Rule engine
Extended Capabilities

- Replication
- Registration of files into the data grid
- Synchronization of remote directory
- Managed file transport (iDrop)
- Automated metadata extraction
- Queries on metadata, tags
- Server-side workflows (loop over result sets)
- Parallel I/O streams & RBUDP transport
Cloud Integration

• iRODS drivers for
  – EC2
  – S3

• Applications include
Integration of Data and Workflows

• Client-side workflows
  – Presentation / display
  – Client managed synchronization / transport

• Policy-driven workflows
  – Automation of administrative functions
  – Processing at the storage system
  – Data-intensive computations

• Traditional computer workflows
  – CPU-intensive computations
Expectations

• Data collection sizes will increase
  – Now petabytes, soon exabytes
    • 1 PB/year = 32 MB/sec
    • 1 PB/day = 11.6 GB/sec
  – Indexing done in the storage system

• Integration of data manipulation with storage controllers - DDN SFA10KE
  – Analyses done within the storage system
Future Applications

• Digital libraries
  – Continuous indexing of contents

• Scientific data collections
  – Extraction of features from data sets
  – Creation of derived data products

• Archives
  – Transformative migrations
  – Validation of assessment criteria
iRODS Development

• Realized Objects
  – Micro-service Structured Object (MSSO)
  – Register into iRODS the workflow needed to create a derived data product and the input parameter / files
  – Accessing the link causes the derived data to be generated and registered as a replica
  – Implemented using compound resource concept
  – The realized object workflow can reference other realized objects, and dynamically re-create an object when any underlying input resource is altered.
  – Effectively manages provenance for workflows
Automated Re-processing Design

Object A workflow depends on
Object B workflow depends on
Object C workflow depends on
Object D

Change object D, and system automatically re-computes objects C, B, and A when object A is clicked
DataNet Federation Consortium

Data Driven Science

- Implement national data grid
  - Federate existing discipline-specific data management systems to enable national research collaborations
- Enable collaborative research on shared data collections
  - Manage collection life cycle as the user community broadens
- Integrate “live” research data into education initiatives
  - Enable student research participation through control policies

Cyber-infrastructure Partners:
- Univ. of North Carolina, Chapel Hill
- Univ. of California, San Diego
- Arizona State University
- Drexel University
- Duke University
- University of Arizona
- University of South Carolina

Science and Engineering Initiatives:
- Ocean Observatories Initiative
- the iPlant Collaborative
- CUAHSI
- CIBER-U
- Odum Social Science Institute
- Temporal Dynamics of Learning Center

National Science Foundation Cooperative Agreement: OCI-0940841
Replication Validation Rule

• Loop over collection in sets of 256 files
• Restart capability for long-running session
  – Tracks identity of the files previously verified
  – Sets TEST_DATA_ID attribute on the collection
• Control execution rate
  – Slow down rate when executing too fast
• Audit trail of all changes to files
• Identify and replace corrupted files
• Create missing replicas
  – Distribute new replicas across storage vaults
iRODS - Open Source Software

Reagan W. Moore

rwmoore@renci.org

http://irods.diceresearch.org

NSF OCI-0848296 “NARA Transcontinental Persistent Archives Prototype”
NSF SDCI-0721400 “Data Grids for Community Driven Applications”