

Prioritization of Tasks

integrated Rule Oriented Data System

Reagan Moore

[moore, sekar, mwan, schroeder, bzhu, ptooby, antoine, sheauc}@diceresearch.org](mailto:{moore, sekar, mwan, schroeder, bzhu, ptooby, antoine, sheauc}@diceresearch.org)

[chienyi, marciano, michael_conway}@email.unc.edu](mailto:{chienyi, marciano, michael_conway}@email.unc.edu)



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



1

THE
NATIONAL
ARCHIVES
ARCHIVES.GOV

Future Development

- Development of simple preservation environment interfaces
 - Template based presentation as in Islandora
- Preservation management features
 - Format parsing routines
 - Representation metadata
- Automated creation of assessment policies
 - Given a template, create rule to validate use
- Development of standard preservation policy sets
 - Starter policy kits for communities



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



THE
NATIONAL
ARCHIVES
ARCHIVES.GOV

New requests - Vote for items of interest

- (7) Virtual ibun that concatenates files for transmission
- (10) Add -P to irsync
- (bug) Issues with symbolic links on FUSE mounted directory - copy symbolic links when directory structure is ignored
- (5-bug) Add object type to audit trail
- (9) irsync across physical resources to repair corruption - similar to irepl
- (8) SHA-256 checksum, CRC checksum
- (9) Export audit trails from iCAT and truncate icp between data grids (bug fixed)
- (3-Mark) List system generated metadata in

iEXP



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Client Requests

- (5) Logging integration into Jargon management
- (8) Multi-thread connection.
- (6) Manage list of connecting IP addresses to track denial of service
- (6) Manage queue of requests. Want to control maximum number of executing requests.
- (6) Implement queuing/client-backoff within Jargon



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



THE
NATIONAL
ARCHIVES
ARCHIVES.GOV

Requests

- (9) Have second notification list about upgrades that address vulnerabilities. Set up an iRODS admin mailing list.
 - (6) Active directory integration with Kerberos, documentation - Chris Jordan
 - (11) Request for Shibboleth based authentication for new users (UNC) credential mapping / assign privilege according to Shibboleth role - TUCASI
 - (1) Automate creation of an account for new users. Example is GSI based access. Extend to LDAP/Kerberos
 - (5) Want to disable triggering of rules from a client by selected users
 - (16) Set access controls on rules and micro-services. Currently on users, collections, resources, metadata.
 - (11-needed) Management of rules within iCAT to enable versioning. Distribution of rules to core.irb file at each server.
-



(12) Use external identity management and external authentication



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Database interaction

- (4) Federation of independent databases: NC B-prepared / EPA / structured information resource drivers
- Ensure end-to-end audit of data sharing policy for pulling data into a secure environment
- Support local policy enforcement on database access (de-identification)
- Support adaptable security policies (change level of access based on situation).
- Use defined SQL commands that can be invoked through a micro-service.
- Need drivers for a wide variety of databases for resources
- (2) Driver for other types of databases for use by iCAT (DB2, SQLServer)
- Integration with VCL policies - policy engine controlled by metadata, which constructs environment
- Which types of databases need to be supported?
- Are there CIM common information models for accessing databases?
- Integration with local metadata (schema), for remote database access
- (7) Support read-only view on selected columns in database



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



THE
NATIONAL
ARCHIVES
ARCHIVES.GOV

Community analysis requests

- (6) Metadata analysis tools for federated data grids
- (6) Collaborator analysis tool for iRODS community
- (6) Project analysis tools for domains using iRODS
- (8) Policy consensus development



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



THE
NATIONAL
ARCHIVES
ARCHIVES.GOV

New requests

- (0) Port SRB scommands onto iRODS (remap options)
- (10) Support for accented letters (metadata and filenames) - Taiwan
- (7) Control number of connections on servers (batch jobs) - connection pool
- (1) iRODS server for memory cache, RAM disk



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List - Dicersearch.org

- (A) Recursive upload of directories - icommands
- (outside) Improve webDAV caching of a working copy of a file correctly
- (6) iDisk for iRODS - iDrop. The iDisk area is on your computer, and reliably synchs to a remote storage system periodically.
- (6) National dropbox. This is similar to iDisk, and provides a way to back up or share your files through a drag and drop interface.
- (2) Native encryption of files, probably done through remote execution. Transforms on data.
- (18) Admin interface to list all current connections and the associated irods users. (keep with record of IP addresses)



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

- (A) Option to restrict irods admin access to specific hostnames/IP addresses.
- (2) Admin interface to rename a resource.
- (6) Add ways to handle the comment fields for data, resources, users, metadata (icommand and micro-service)
- (5) Add AVU metadata to resource groups.
- (bug) Allow GSI authentication to non-ICAT-Enabled-Server without irodsUserName being set.
- A more general mechanism to access external databases. Admin will define location and specify SQL, client will be able to provide arguments. Independent of ICAT. Need a way for the information provided, or saved, to be integrated into the iRODS framework.



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

Rule base Management

- (4) Provide mechanism to synchronize rule bases across servers within a data grid.
- (4) Provide versioning support for rules.
- (A) Provide mechanism to add rule base extensions to a remote rule base.

Installation

- Create a VM build for use of iRODS in tutorials.
- Create a VDT (Virtual Data Toolkit) version of iRODS.
- (A) When compile, verify that only changed files are recompiled.



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

File Manipulation

- (A) Support Soft Links in iRODS (supported in 2.3)
- (0) Manage locks for collaborative editing (either on storage system, in metadata, or portal) - expensive
- (1) Add core.irb rule to check policy on iget. May want to return alternate version (redacted copy) of file.
acPreProcPolicyForOpen
- (1) Support compressed files on disk for minimizing space. Could be done through a preProcForPut core.irb hook.
- (A) Hook for preProcForPut for medical applications.
- (1) Decompress on the fly.
- When fail over to an alternate server, verify that the Rule engine will not have a conflict for a queued Micro-service. Related to versioning of rules and micro-services.



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

Network Transfer

- (2) Configure port numbers for data transfer. Support alternate IP address.
- (A) Support sequential transfers as one-hop, similar to parallel. Done for iput.
- (A) Jargon parallel I/O
- (0) Extensions to Jargon for UDP transport.
- (8) Support checksums in Jargon (before transfer), iput -k
- (bug) Fix UDP on Solaris.
- Support transfer of multiple files using multiple I/O streams. Virtual ibun

Windows

- (3) Create a Windows only environment, using Postgres and GSI (both now run on Windows)
- ~~• (5) Create a Windows only environment, using SqlServer~~



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

Micro-services

- (A) Support Python-based Micro-services by including Python interpreter in the Micro-service. Documentation
- (4) Support Perl-based Micro-services by including Perl interpreter in the Micro-service.
- (9) Support for Java-based micro-services.
- (6) Create script for automating module creation. Provide default template for creating new Micro-service. Documentation
- Provide a library of script-based micro-services
- (7) Provide transaction based interaction with external workflow



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

Security / Authentication

- (4) Support a session shell in iRODS for iRODS commands, isession. Authenticate once.
- (1) Support ticket-based access to iRODS.
- Support token-based identification such as SecureID. One-time passwords - offload to external authentication management



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

Administration

- (7) Add ability for a project PI to create rodsuser accounts for project members. Create accounts under groupadmin
- (7) Restrict ability of ASSIGN to change user name, user role, user permissions.

Drivers and Access

- (A) Mount a flash drive. Treat as a mounted collection.
- (A) Support mounting of a Webdav directory into iRODS. (done through DAVIS)



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

iRODS Metadata Catalog (iCAT)

- (0) Port iCAT to Sybase.
- Support logical registration into iRODS. Ability to associate metadata with a name without requiring a file. Add an info file.
- (bug) Automate Dicom metadata extraction from binary files. Find misextractdicom
- (bug) Implement FITS header parsing and metadata loading. Find msiextractfits
- Create digital signing registration to be able to track origin of files. Given signature, find original copy. Related to definition of an AIP
- (0) Create RDA interface to Sybase.
- (A) Save RDA request results for use in a session, want to pass result list to another Micro-service. Put in a file.



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

Interfaces

- (Stork) Create a Condor interface / port for iRODS
- (0) Develop Perl API.
- (0) Develop Fortran API.
- (12) Add regular expressions to i-commands (wild cards).

Development

- (A) Create a collaboration environment for promoting development of iRODS. (done through DICE)
- (A) Consider use of Google Code repository and other shareware sites. Automate. GitHub (Russell)



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Wish List (Cont.)

SRB

- For SRB to iRODS metadata migration, handle migration of SRB zones. How can multiple SRB zones be re-federated within iRODS easily? Limited scripts.
- (0) Port SRB APIs on top of iRODS, will avoid having to rewrite many application scripts.

Information

- (15) Editorially review of documentation
- Publish the complete iRODS network protocol. (in progress)
- (expand) Post on Wiki a list of the planned developments for Jargon.
- (bug) Add a Doxygen indexed version of iRODS source on the wiki for micro-services and helper functions.
- (A) Add links to WebDav environment on iRODS wiki, provides drag and drop access to files across windows.
- (bug) Post link to VBrower, links to Taverna and EGEE grid.



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Requests

- (15) Links to Petashare, Urspace, Jux, 43 clients
- (1) Modify icommands to use the zoneID:internal ID instead of the file name or add GUID
- (15) Installation - check for firewall interaction (notification of presence, like a ping on each port).
- (15) icommand for checking ports (health check). Jargon admin will do this
- Script for checking/updating IP address (5 locations).
- Support for dynamic IP address.
- (A) Command to force execution on another server. msiremoteexec



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



THE
NATIONAL
ARCHIVES
ARCHIVES.GOV