White Paper: National Data Infrastructure for Earth System Science

Reagan W. Moore
Arcot Rajasekar
Mike Conway
University of North Carolina at Chapel Hill

Wayne Schroeder
Mike Wan
University of California, San Diego

A national data infrastructure is emerging through the federation of data management infrastructure used by National Science Foundation research projects, the National Oceanic and Atmospheric Agency archive, and NASA data grids. The approach is based on the use of collection virtualization technology to enable interoperability between existing data management systems. The technology that enables federation is the iRODS integrated Rule-Oriented Data System [1,2,3]. The capabilities provided by the iRODS data grid are being used today to implement national data infrastructure in the United States, Europe, the Far East, and Australia.

The central concept is the realization that collaboration requires the formation of a shared collection. Researchers rely upon a common name space for identifying files, common semantic terms to understand the context associated with each file, shared procedures for manipulating and analyzing the data, and a shared consensus on the policies for managing the data. The shared context enables a designated community to understand and use the shared collection.

A significant challenge is that the shared context depends upon the knowledge available to the designated community. The original creators of a data set have a highly sophisticated understanding of the methods that generated the data and only need a simple context based on the name of the application and input parameters for simulation data. When the data are shared with a broader community, the context needed to describe the data now should include information about the generating application, the coordinate system in which the data are embedded, the units used to define physical quantities, and the relationships between the physical quantities. When the data are published in a digital library, the context associated with the data must meet the semantic standards of the discipline, including a standard format (data structure), ontologies that map relationships between semantic terms, and access policies for allowed use of the data. When the data are preserved, the required context must meet the needs of future researchers for correct interpretation and processing of the data. The shared context evolves to provide a more detailed description of the data as the user community broadens. The evolution of the shared context constitutes a community-driven data life cycle. An example life cycle is the migration of data from a project collection, to a collection shared with other researchers, to a digital library for formal publication of vetted results, to a reference collection for use by future researchers.

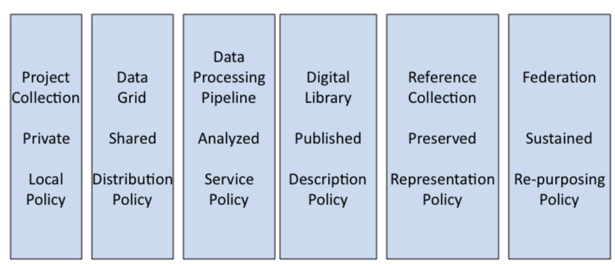

Figure 1. Community-based collection life cycle stages

In the iRODS data grid, each stage of the community-driven data life cycle is governed by an explicit set of policies that represent the community consensus on data sharing. The policies enforce assertions about collection properties, such as internal consistency, completeness, authoritativeness (data source), authenticity, coverage, and reliability. The policies appropriate for a local project may allow the local collection to contain intermediate results, output from trial analyses, and un-vetted results. The policies for a digital library may require that all data be calibrated, evaluated for internal consistency, and explicitly approved for publication. Thus the policies for each stage of the data life cycle are different. The procedures used to apply the policy will also differ, along with the state information that records the outcome of policy enforcement. A community-driven data life cycle requires a highly extensible architecture in which policies, procedures, and state information evolve as the user community broadens.

A policy-based data management environment has the ability to validate assessment criteria and verify that an intended use of a collection is consistent with the published context. An assessment policy can be periodically applied to verify that a desired collection property has been conserved. An example is the validation of an integrity policy that requires that each file be check-summed and replicated. The assessment policy periodically calculates each file checksum, compares with the original value, and replaces a bad copy with a good copy. There is a very strong link between automation of administrative functions and periodic execution of assessment criteria. It is possible to design policies that enable properties of a collection to be verified with automatic correction of detected errors. As collections grow to the multi-petabyte size and are distributed across multiple locations, such automation is essential to minimize administrative labor costs.

Each user community is accustomed to a preferred user interface. As the user community broadens, the set of user interfaces to the collection should also broaden. Thus the original project might prefer a file-system interface and directly manipulate data within the shared collection through unix tools. The preferred access mechanism for the shared collection might be a data processing pipeline based on a standard workflow system, while the preferred access to the digital library may be a web browser. In common practice, there are roughly eleven types of access methods used within research disciplines: File systems, Unix tools, Scripting language load libraries, Workflow systems, Web services, Portals, Grid tools, Digital libraries, Dropbox, Web browsers, and I/O libraries. Some users may prefer using a Fortran I/O library for direct access from an executing application, while other users may prefer a dropbox interface to drag and drop a folder holding 300,000 files into the collection, and have the system monitor the transfer of each file. More than 50 access mechanisms have been ported onto the iRODS data grid.

Each user community also has access to different types of storage environments, ranging from laptop file systems, to high-performance supercomputer disk systems, to tape archives, to cloud storage. The iRODS collection virtualization environment maps from the actions specified by the preferred access mechanism, to the I/O protocol required by the selected storage environment. The iRODS data grid technology consists of software middleware that manages collections that are distributed across multiple storage systems. Policy enforcement points within the iRODS framework ensure that policies are applied independently of the choice of access mechanism. The iRODS data grid enables use of modern access methods to retrieve data from legacy storage systems, while enforcing the community-chosen data management policies.

A national data infrastructure for Earth System Science must also manage interactions between communities that have chosen different management policies. The concept of federation is based on the idea that a new virtual shared collection can be created that represents the data-sharing consensus between the communities. Each community can choose to enforce local management policies within their data grid, and then establish the additional policies that govern the shared collection. Each community can retain local control while enabling researchers to collaborate through shared collections.

Federation of data collections can be managed at multiple levels: through the policies governing direct data sharing between two data grids and through soft links that register a data set from a remote data grid into the local data grid. The concept of soft links has been generalized to support access to an external information resource. The desired resource interaction is encapsulated as an object within the data grid, and can be listed as an object in a logical collection. Clicking on the information object causes the external request to be instantiated, and a local copy of the information or file is registered as a replica of the original information object. Information objects can be created for access to any type of information resource that has a published protocol specification. Within the iRODS data grid, examples include

Z39.50, web pages, anonymous ftp, SRB data grids, and even un-federated iRODS data grids. The shared collection can contain data and information that are linked across heterogeneous data management environments.

Policy-based data management systems are in production use across multiple disciplines, from earth systems science to astronomy, biology, cognitive science, engineering, hydrology, oceanography, seismology, and social science. At the Renaissance Computing Institute (RENCI), iRODS data grids have been established that process NEXRAD data from the National Climatic Data Center (NCDC) on RENCI computers. The RENCI data grid has been federated with the NSF Teragrid data grid (for data replication), and the NSF Ocean Observatories Data grid (for archiving of real-time sensor data) [4]. Additional data grids at the University of North Carolina at Chapel Hill support the Carolina Digital Repository (institutional repository), the TUCASI collaboration environment (regional data grid), the NARA Transcontinental Persistent Archive Prototype (persistent archive), and the LifeTime Library (School of Information and Library Science student digital libraries). The NASA Center for Climate Simulation (NCCS) has imported the Moderate Resolution Imaging Spectro-radiometer (MODIS) atmosphere data set into an iRODS data grid, and federated observational data with simulation data [5]. International user communities include the CyberSKA (radio-astronomy Square Kilometer Array) [6], the French National Institute for Nuclear Physics and Particle Physics (high energy physics, biology, astrophysics, humanities), the Australian Research Collaboration Service, and the Sustainable Heritage Access through Multi-valent ArchiviNg preservation project at the University of Liverpool. Additional communities include environmental science [7], astronomy [8], plant biology [9], and genomics [10,11].

The technology exists today to implement federation of data management systems across NSF research projects, across federal agencies, and across international collaborations. In particular, the NSF DataNet Federation Consortium cooperative agreement is building a federation hub at RENCI to integrate federal repositories (NOAA NCDC, NASA NCCS [12]), national data grids (Teragrid / XSEDE), regional data grids (RENCI data grid), and institutional repositories (Carolina Digital Repository, LifeTime Library). The approach relies on iRODS release 3.0 to federate independent data management systems into a common data name space. The architecture is shown in Figure 2. Peer-to-peer servers implement the major components of the system: storage servers, message server, catalog enabled server, scheduling server, and a workflow server. Each storage server manages conversion to the protocol required by the local storage system, and uses a local rule engine and rule base to apply policies at that server location. The message server is used to track progress and support distributed debugging of the distributed rule engine. The catalog-enabled server manages all state information and audit trails tracking use of the environment. The scheduling server manages execution of deferred and periodic policies, and the workflow server manages interactions with external workflow systems.

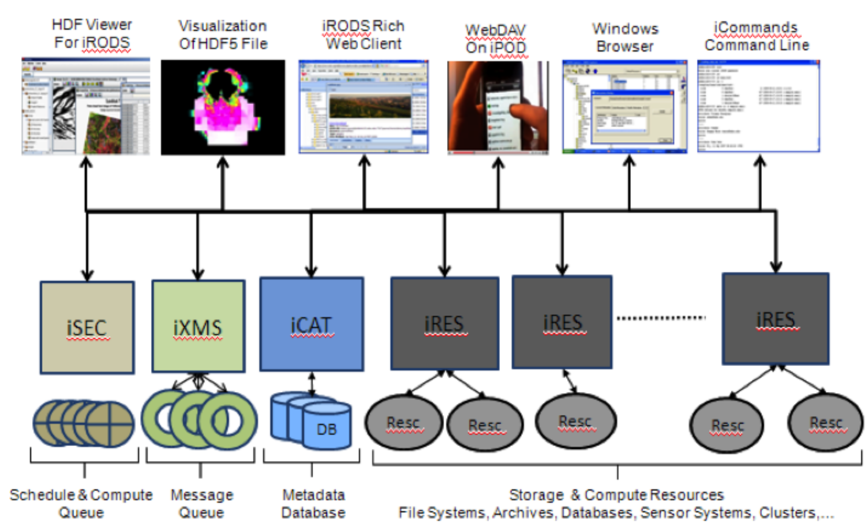

Figure 2. Peer-to-peer Server Architecture

The levels of virtualization needed to manage data distributed across multiple administrative domains and across multiple types of storage systems are shown in Figure 3. The actions requested by each type of access client are trapped by policy enforcement points within the data grid middleware. For actions such as ingestion of a file into the data grid, ten policy enforcement points are traversed. Policies are checked for allowed access by a remote host, access permissions for public files, access permissions for private files, the server to use for storage, whether the storage quota is exceeded, the physical path name to use when writing the file, whether metadata should be modified before or after storage, and whether post processing is required after file creation or after the file write. The policies are applied by the rule engine and read from the local rule base. Each policy controls the execution of a procedure that is composed through the chaining of functional units called micro-services. The micro-services are operating system independent, and issue I/O operations based on an extension to Posix I/O. This implements infrastructure independence, with the same micro-services used on Unix and Windows operating system. The standard I/O protocol is then mapped to the protocol required by the choice of storage system. This approach makes it possible to add new clients, add new policy enforcement points, add new policies, add new procedures, and add new storage systems without modifying other components of the system.

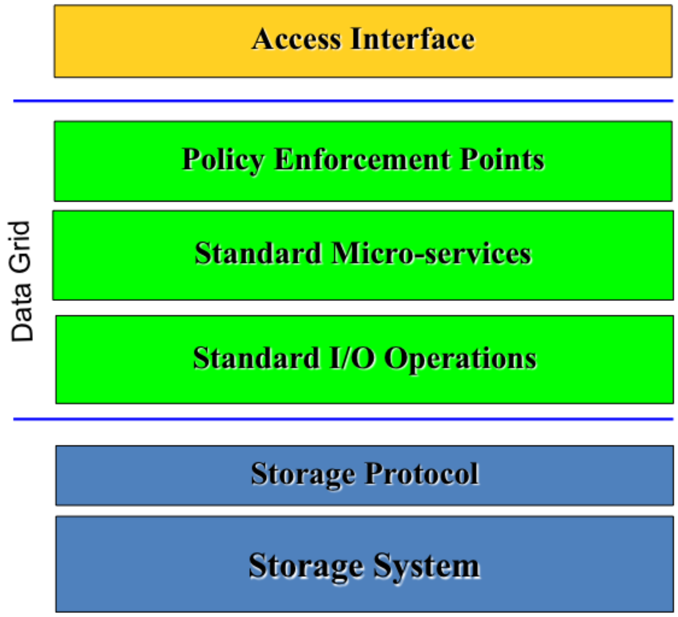

Figure 3. Virtualization Levels

The procedures can be thought of as storage server-side workflows that are intended to implement low complexity operations (small number of operations compared to the size of the file in bytes). These workflows can be used to implement administrative tasks (file migration). They can also enforce data management policies (retention, disposition, distribution, time-dependent access controls) and validate assessment criteria (integrity, authenticity, chain of custody). Each community chooses a preferred client and selects the set of policies and procedures that are appropriate for their data management application.

The forms of federation that can be supported include collections that contain soft links to remote information resources, or the mounting of remote directories into the data grid, or true federation in which two independent data grids establish a trust relationship, and users are cross-registered between the data grids. A user in Data Grid A can be given access to data within Data Grid B, while the policies of Data Grid B control what can be done. Each user is authenticated by their home data grid, while data management policies are enforced by the data grid in which the files are located. This makes it possible to apply multiple types of federation, from creation of central archives into which data grids deposit data for common management; master-slave data grids in which the slave data grids receive data from the master data grid; deep archives in which data are pulled from a staging data grid into the archive; and chained data grids through which data are replicated to ensure multiple copies. The integration of multiple, heterogeneous data management environments is now possible. The creation of national data infrastructure that links academic and federal repositories will accelerate research on societally important questions in earth science.

Acknowledgements: The development of the iRODS policy-based data management system was funded by the NARA supplement to NSF SCI 0438741, (2005-2008) and the NSF OCI-0848296 grant (2008-2012). Funding was also provided by NSF ITR 0427196 (2004-2007) and NSF SDCI 0721400 (2007-2010).

References:

1. iRODS integrated Rule-Oriented Data System, http://irods.diceresearch.org.
2. Rajasekar, A., M. Wan, R. Moore, and W. Schroeder, "A Prototype Rule-based Distributed Data Management System," HPDC workshop on Next Generation Distributed Data Management, Paris, France, 2006.
3. Moore, R., A. Rajasekar, "iRODS: Data Sharing Technology Integrating Communities of Practice", Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, July 25, 2010.
4. Farcas, C., P. Fox, M. Arrott, E. Farcas, I. Klacansky, I. Krueger, M. Meisinger, J. Orcutt, "OOI: OOI CyberInfrastructure - Next Generation Oceanographic Research", American Geophysical Union, Fall Meeting 2008, abstract #IN53A-1182, http://adsabs.harvard.edu/abs/2008AGUFMIN53A1182F.
5. Schnase, J., G. Tamkin, D. Fladung, S. Sinno, R. Gill, "Federated Observational and Simulation Data in the NASA Center for Climate Simulation Data Management System Project," Proceedings of the iRODS User group meeting, Chapel Hill North Carolina, February 2011, http://iren-web.renci.org/irods-meeting/nasa.pdf.
6. Mahadevan, V., E. Rosolowsky, "CyberSKA: A Distributed Datacube Analysis Service for Radio Telescopes", Astronomical Data Analysis Software and Systems XX. ASP Conference Proceedings, Vol. 442, Astronomical Society of the Pacific, 2011, p. 219, http://adsabs.harvard.edu/abs/2011ASPC..442..219M
7. Demontis, R., E. Lorrai, V. A. Marrone, L. Muscas, V. Spanu, A. Vacca, P. Valera, "GIS & iRODS: Environmental applications based on GIS and GRID technologies", EGU General Assembly 2009, held 19-24 April, 2009 in Vienna, Austria, http://meetings.copernicus.org/egu2009, p.11471, http://adsabs.harvard.edu/abs/2009EGUGA..1111471D
8. Barg, I., D. Scott, E. Timmermann, "NOAO: NOAO E2E Integrated Data Cache Initiative Using iRODS", Astronomical Data Analysis Software and Systems XX. ASP Conference Proceedings, Vol. 442, Astronomical Society of the Pacific, 2011, p.497, http://adsabs.harvard.edu/abs/2011ASPC..442..497B
9. Jordan, C., et.al., "iPlant: Comprehensive Data Infrastructure for Plant Bioinformatics", Workshop on Interfaces and Abstraction for Scientific Data Storage, 2010.
10. Chiang, Gen-Tao, P. Clapham, G. Qi, K. Sale, G. Coates, "Sanger Institute: Implementing a genomic data management system using iRODS in Wellcome Trust Sanger Institute", Journal, BMC Bioinformatics, Vol 12, 2011, http://www.biomedsearch.com/nih/Implementing-genomic-data-management-system/21906284.html
11. Smith, C., "Broad Institute: iRODS at the Broad Institute: Implementing a Large Scale Archive", Bio-IT World Conference, 2011.
12. Schnase J., et.al., "NASA: iRODS-Based Climate Data Services and Virtualization-as-a-Service in the NASA Center for Climate Simulation", 2011 AGU Fall Meeting, December 2011.