Control Your Data

iRODS, the integrated Rule-Oriented Data System

A RENCI WHITE PAPER IN
COLLABORATION WITH THE IRODS CONSORTIUM

renci iRODS CONSORTIUM

CONTACT INFORMATION

Brand Fortner, PhD
Telephone: 919.445.9692
Email: bfortne@renci.org

Summary

Data management represents a complex challenge in today's environment of large and often distributed data sets, rapidly evolving technologies, increased reliance on outsourced resources, and developing regulations regarding data access and use (Ahalt et al. 2014). Specific challenges include a need to manage data across multiple, geographically separated users, while ensuring enforcement of data-related policies; automation of tasks; data replication, back-up, and archiving; quality control; metadata capture and catalog; and access control. This white paper provides an overview of a proven technology that overcomes these challenges to enable policy-based, distributed data management: iRODS, the integrated Rule-Oriented Data System.

The Challenge: Data Management Gone Awry

Mr. John Doe just lost his job as Chief Information Officer (CIO) of a successful company that provides data management services to multiple clients. On top of that, the company is now facing a multimillion dollar lawsuit that likely will discredit its reputation and lead to its collapse.

All of these events transpired over the course of one week—when Mr. Doe lost 10 years of clinical trial data that belonged to one of his company's clients, a small but respected Clinical Research Organization (CRO) that was running a multicenter clinical trial on behalf of a large pharmaceutical company.

Mr. Doe's company had been providing data management services to capture data across clinical trial sites, integrate the data within a centralized company database, and provision the integrated data to study personnel upon request via a Web-based user interface. The company's trouble began when its primary database crashed due to a faulty update in the BIOS software that was managing the company's data storage devices.

iRODS At A Glance

- iRODS technology provides a comprehensive solution to policy-based data management across multiple, geographically separated users, with features to enable data and policy virtualization; automation of tasks such as data replication, back-up, and archiving; quality control; metadata capture and catalog; and secure data access.
- iRODS has been deployed in at least 39 countries from around the globe, with user groups drawn from U.S. federal and international agencies, international consortia, commercial entities, non-profit organizations, and academic institutions. iRODS installations currently manage billions of files and >30 petabytes of data.
- The iRODS Consortium has been established to ensure long-term sustainability of the open source iRODS through ongoing development and innovation.

The Team

Brand Fortner*, PhD, Executive Director, iRODS Consortium, and Adjunct Professor, Department of Physics, University of North Carolina at Chapel Hill (UNC); **Stanley Ahalt**, PhD, RENCI Director, Professor, UNC Department of Computer Science, and iRODS Consortium Executive Board member; **Jason Coposky**, Chief Technologist, iRODS Consortium; **Karamarie Fecho**, PhD, Medical and Scientific Writer, RENCI; **Stefan Heinzel**, PhD, Director, Rechenzentrum Garching; **Ashok Krishnamurthy**, PhD, RENCI Deputy Director; **Reagan Moore**, PhD, RENCI Chief Scientist, Professor, UNC School of Information and Library Science, Director Data Intensive Cyber Environments (DICE) Center at UNC, and iRODS Consortium Executive Board member; **Arcot Rajasekar**, PhD, Co-Director, DICE Center, Professor, UNC School of Information and Library Science and RENCI Chief Domain Scientist for Data Grid Technologies; **Charles P. Schmitt**, PhD, RENCI Chief Technical Officer and Director of Informatics and Data Science; and **Wayne Schroeder,** DICE iRODS co-Principal Developer/Product Manager and DICE Principal Investigator at University of California, San Diego.

iRODS Capabilities at a Glance

- Data Virtualization
- Policy Virtualization
- Automation
- Replication
- Back-up
- Archiving
- Quality Control
- Metadata Capture and Catalog
- Access Control

*Brand Fortner is Executive Director of the iRODS Consortium; all other team members are listed alphabetically.

Information Technology (IT) staff investigated but concluded that they could not recover the data. They also informed Mr. Doe that the data were not replicated in a secondary database or at center sites, which meant that the data were no longer being pulled from sites or accessible to the client, leaving the trial essentially suspended. At this point, Mr. Doe was upset but not terribly concerned; after all, he knew that his IT staff backed up company data nightly using a secure database. His IT staff reminded him, however, that the company's administration, in an effort to constrain costs and reduce back-up time, had decided not to back-up the raw data files, just the processed data files.

The raw data files from the first five years of the 10-year period in question were archived using tape drives, but the company stopped archiving data because the system was not automated and the IT staff member who was responsible for archiving left the company to pursue a new opportunity. In addition, the system used to read the tape drives was an older, legacy system that was no longer in general use, and even if current IT staff could find a replacement system, insufficient metadata existed on the data and software to interpret the data and verify data quality. To make matters worse, they discovered that a few of the archived tape drives were missing, and they feared that the tapes were leaked to the former IT staff member's new employer, a competitor and unauthorized user.

Mr. Doe was forced to reveal his company's data management problem to the client CRO and the large pharmaceutical funder. The administration at the pharmaceutical company was, for obvious reasons, outraged. After all, they lost millions of dollars and many years' time collecting data that were now lost forever. They feared that the processed data that were still available would not be sufficient to create a Clinical Study Report and address any concerns that the US Food & Drug Administration might have when evaluating the Investigational New Drug that they were developing.

The pharmaceutical funding agency responded by removing the CRO from their contractor list and filing a lawsuit against Mr. Doe and his company. Mr. Doe was promptly fired.

The Solution: iRODS, the integrated Rule-Oriented Data System

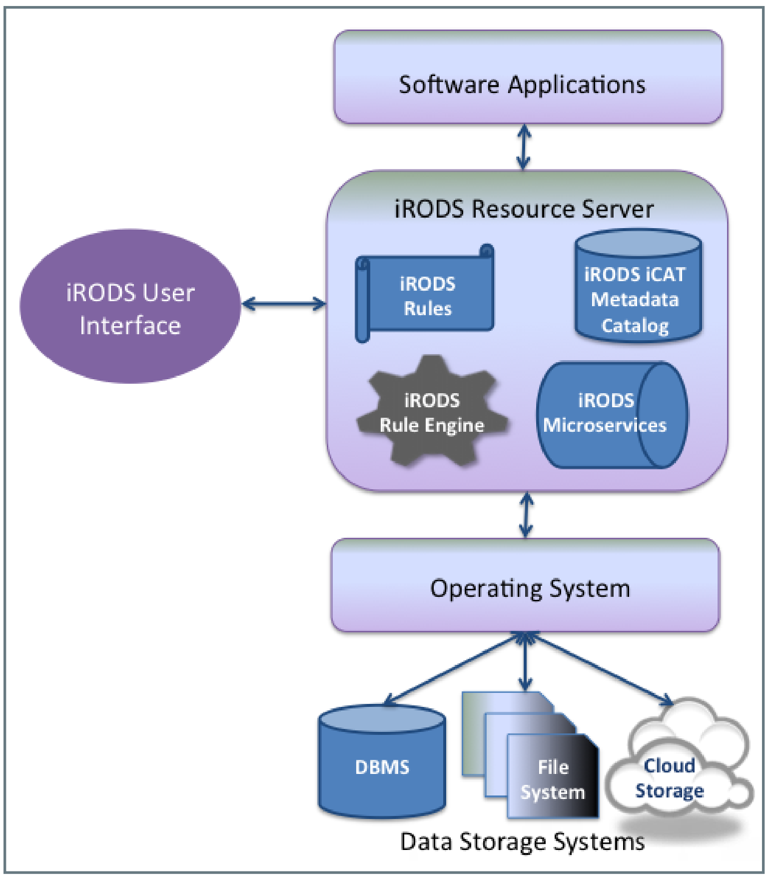

Figure 1. An overview of several key components of iRODS, including: the iRODS Resource Server(s); iRODS User Interface; iRODS Rule Engine; iRODS Rules; iRODS Microservices; and iRODS iCAT Metadata Catalog. DBMS = DataBase Management System. Note that the iRODS Resource Server can access network-based storage systems through plugins such as Amazon's S3 API.

Our protagonist's story could have had a very different outcome had he adopted the open source iRODS for his data management needs.

iRODS was architected by the Data Intensive Cyber Environments Center in collaboration with RENCI, and with contributions from multiple groups throughout the world, as a comprehensive solution to the complex challenges involved in managing data across multiple, geographically separated users. iRODS is a file system that functions as middleware in the sense that it is runs on top of other software programs and operating systems and can be called by higher-level software. iRODS has several key components (see Figure 1): the iRODS Resource Server(s); iRODS User Interface;

iRODS Rule Engine; iRODS Rules; iRODS Microservices; and iRODS iCAT Metadata Catalog. An iRODS Data Grid can be created by linking multiple iRODS Resource Servers and related components Each component provides key functionalities that address specific challenges in distributed data management.

The **iRODS Resource Server** manages access to a defined collection of digital objects; the digital objects may map to a traditional file storage system, a cloud-based storage system such as Amazon S3, an Object Store such as the Data Direct™ Networks WOS, non-traditional file systems such as Apache® HDFS, or even dynamically instantiated data via a web service or SQL query against a RDBMS. Access to the iRODS Resource Server is achieved via the **iRODS User Interface**, of which there are several existing types, including command line–driven, web-based, and client-based GUIs. In addition, domain-specific interfaces can be created to meet individual user needs. The iRODS User Interface is used to securely visualize, access, analyze, upload/download, and share all digital objects that are managed by the iRODS Resource Server. iRODS supports multiple methods of user authentication (e.g., secure passwords, GSI, Kerberos, LDAP/PAM) and data encryption, both at rest and in motion, to provision secure data access. As middleware, iRODS can also be accessed directly via user interfaces associated with software applications running on top of it, thus providing integration within existing systems and reducing deployment costs by eliminating the need to adapt existing infrastructure.

Multiple iRODS Resource Servers can be supported by a single instance of iRODS through a distributed **iRODS Data Grid**, which is accessed through the iRODS User Interface and enables data virtualization and access to all iRODS Resource Servers in the data grid. In the example shown in Figure 2, five different iRODS Resource Servers are supported across five geographically separated users. The iRODS Resource Servers can be located anywhere in the world, and iRODS does not limit the number of servers that can be supported.

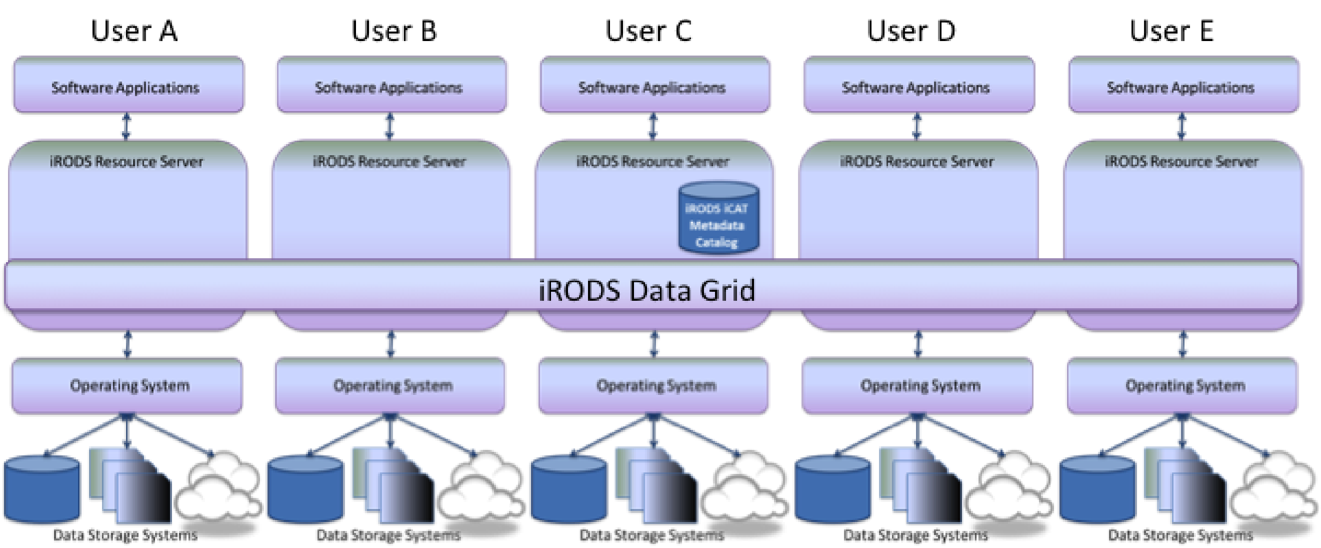

Figure 2. The iRODS Data Grid. In the example shown, the iRODS Data Grid is comprised of five distributed iRODS Resource Servers across five geographically separated users. Note that the iRODS iCAT Metadata Catalog typically resides as a centralized relational database in a specified server in the data grid (User C in this example). Any commercial database can be used to host the iRODS iCAT Metadata Catalog, and most support replication of the database to either a secondary database or to distributed databases through a "master-slave" configuration. Not depicted are the iRODS Rule Engine, Rules, and Microservices, which are stored redundantly within each server in the iRODS Data Grid, or the iRODS User Interface.

The i**RODS Rule Engine** resides within each iRODS Resource Server and implements user-defined policies by invoking executable **iRODS Rules**. The policies are customizable and can address all aspects of the data management lifecycle, including: data flow (e.g., data collection, storage, movement, back-up, replication); data curation (e.g., data provenance, annotation, metadata, archiving, publication); data analytics (e.g., data cleaning, transformation, filtering, visualization, and quality control); and data security (e.g., data access, user authentication, data encryption, audit trails). For example, the policies may specify that the data should be replicated to one or more secondary iRODS Resource Servers in the iRODS Data Grid to ensure that the data remain available in the event of a failure of one or more data systems.

Other policies may address the frequency and methods for data back-up to enable disaster recovery or data archiving to provide long-term, high-quality preservation of data. Policies also may define the analytical steps required to conduct quality control checks on the data in order to verify data integrity. Yet other policies may specify authorized users and conditions related to data access.

Metadata Example

Logical Name: /SiteX/RDDept/Flow/Study101

Physical Name: /London/var1/proj/labx/flowdata

Laboratory PI: Jane Doe

Date: 01/01/2014

Time: 01:45:12

Study: Clinical Trial on the Safety and Efficacy of DrugX in Stimulating Immune Function among Patients with Chemotherapy-induced Immunodeficiency

Data Source: Peripheral blood

Antibodies: PE-conjugated anti-CD3 antibody, FITC-conjugated anti-CD4 antibody

Assay Conditions: Data captured under conditions x, y, z

Data Capture Technology: Brand X, v 2.0, Flow Cytometer

iRODS Rules implement these policies through **iRODS Microservices** that define specific policy-related tasks. iRODS Microservices are, as the name implies, small, user-defined operations (e.g., conducting a specific search of an external database, performing an analysis of the data, zooming in on an image) or system-defined operations (e.g., deleting backed-up data older than one month, checking for corrupt data, verifying a user password). **iRODS Microservices** automate complex workflows and are executed by iRODS through the **iRODS Rule Engine**, thus providing policy virtualization. Updates to the iRODS Rules and Microservices can be made independently of modifications to core software code, thereby providing system extensibility. iRODS also accepts numerous plugins to provide even greater extensibility.

iRODS implements a single logical namespace to organize and catalog digital objects across data systems in the iRODS Data Grid; this approach ensures standardization across users and provides data virtualization across the distributed grid. iRODS captures and catalogs this information with additional metadata to maintain an up-to-date record on the state of the digital objects (past and current). The metadata are stored in the **iRODS iCAT Metadata Catalog**. The iCAT Metadata Catalog is typically configured as a centralized relational database that resides within a specified iRODS Resource Server; however, any commercial database can be used to host the catalog, and most commercial databases have features to enable replication of the database to a secondary database or to distributed databases through a “master-slave” configuration. Any commercial database can be used to support the iRODS iCAT Metadata Catalog, and most have features to enable replication of the database to a secondary database or to distributed databases through a “master-slave” configuration. The metadata include: user domain–related metadata, such as data annotations, data origination steps, and domain-specific data schemas; and system-related metadata, such as user namespace (e.g., contact information, title, role on project), file namespace (e.g., creation date, location, access controls), and storage namespace (e.g., capacity, quotas, type [archive, disk, fast cache]), (see box “Metadata Example” at left). Thus, the iRODS iCAT Metadata Catalog manages user-defined metadata, maintains consistent administrative information about the distributed environment, and tracks data provenance.”

A final feature of iRODS that is worth emphasizing is that iRODS was developed as open source software, with continued development and updates provided by the iRODS Consortium of users and developers (**irods-consortium.org**). Thus, long-term sustainability of iRODS is secured through proven open source community development (Hammond 2009).

Our Protagonist's Story: A Better Outcome

As discussed, Mr. Doe's company was hired to provide data management services on behalf of a CRO and pharmaceutical funder for a 10-year clinical trial involving multiple centers. Suppose that iRODS had been installed to manage the study data. iRODS can provide data virtualization across the center sites by implementing a unified name space and creating a distributed data grid, thus eliminating the need for data to be centralized at a single site.

iRODS also can be used to provide policy virtualization and the automation of tasks. For example, iRODS policies can be defined and deployed to ensure that study data are automatically replicated to multiple, geographically separated iRODS Resource Servers, thereby guaranteeing that they remain accessible in the event of a failure of one or more data systems. iRODS policies can also be defined to automatically back-up processed data every 24 hours and archive raw data once per week. This approach would allow Mr. Doe's company to keep costs and back-up time to a minimum, while ensuring immediate access to nearly all processed data and eventual recovery of most raw data in the event of a catastrophic failure of a storage system. The iRODS iCAT Metadata Catalog can be used to archive data with the necessary metadata to enable interpretation of the data and validation of data quality over time and across evolving storage systems, thus minimizing downstream issues related to legacy storage systems and time-related data degradation. Additionally, iRODS can secure data access by implementing policies requiring, for example, two-factor user authentication and encryption of data at rest and in motion, which would minimize the risk of data leakage.

Had Mr. Doe relied on iRODS for his company's data management needs, he likely would still be employed as CIO, he and his company would not be facing a large lawsuit, and his client (the small CRO) would not be facing an uncertain future.

The Upshot

iRODS provides a comprehensive solution to the many challenges involved in distributed data management.

Key user features:

1. Data virtualization is provided through a unified logical namespace for digital objects to provide standardization and shared access to data across a distributed data grid.
2. Policy virtualization is provided through the automation and execution of all user-defined data management policies by a powerful, distributed rule engine.
3. Data management tasks, such as data replication, back-up, archiving, and quality control, are automated and can be defined based on user needs and available resources.
4. A centralized iCAT Metadata Catalog is used to maintain consistent administrative information about the distributed environment, manage user-defined metadata on digital objects, and track data provenance.
5. Data access is secured through administrator-defined policies related to authentication of each user, authorization of each operation, management of audit trails, and transport-level encryption.

Key technical features:

1. The middleware technology is directly accessible by software applications through a user-level file system and supports access to multiple types of data storage resources, including databases, file systems, Network File Systems, cloud storage systems (e.g. Data Direct™ Network WOS, Amazon's S3 storage), Apache® HDFS, and tape archive systems (e. g. Sun Sam™ FS, Quantum® Scalar).
2. The technology is extensible, via plugins and updates to rules and microservices that can be made independent of modifications to core software code.
3. The technology is adaptable for a variety of users and applications (e.g. data sharing in research projects, pipeline processing via workflows, data publication in digital libraries, and data preservation in archives).
4. The technology is scalable, from personal use to multi-user, distributed collaborations involving petabyte-sized data sets.
5. The technology is open source, with ongoing sustainability ensured via the iRODS Consortium.

The Big Picture

iRODS represents a novel and versatile technology designed to enable policy-based, distributed data management across the data lifecycle. The iRODS Consortium supports ongoing development and innovation of iRODS, thus ensuring long-term sustainability. iRODS is currently used by numerous groups and for a variety of applications (e.g., Hedges et al. 2007; Rajasekar et al. 2006; 2010a,b; Barg et al. 2011; Chiang et al. 2011; Schnase et al. 2011; Ward et al. 2011; Schmitt et al. 2013).

User groups include the National Aeronautics and Space Administration, National Oceanic and Atmospheric Administration, National Archives and Records Administration, National Optical Astronomy Observatory, United States Geological Survey, Broad Institute, International Neuroinformatics Coordinating Facilities, Wellcome Trust Sanger Institute, Computing Center of the French National Institute for Nuclear Physics and Particle Physics (CCIN2P3), the French National Library, the French Data Grid, CineGrid, the iPlant Collaborative, BaBar High Energy Physics Experiment, the Australian Research Collaboration Service, DOW Chemical, Bejing Genome Institute, Distributed Bio, Computer Sciences Corporation (CSC), Cleversafe, and numerous academic research units from around the world.

iRODS has demonstrated success at scale and currently supports more than 20 petabytes at the Wellcome Trust Sanger Institute, over 6 petabytes of data at the French CCIN2P3, about 2 petabytes of data in the BaBar High Energy Physics Experiment, approximately 15,000 users in the iPlant Collaborative, more than 500 storage resources in the Australian Research Collaboration Service, about 300 million attributes in NASA's Center for Climate Stimulations, and dozens of sites located across the United States, Europe, and Japan through the CineGrid organization. In all, iRODS installations have been used to manage billions of files and over 30 petabytes of data.

Acknowledgments

iRODS core technology was developed with principal funding from the National Archives and Records Administration and the National Science Foundation. Additional support was provided by numerous user and developer groups worldwide.

iRODS has received numerous awards, including the *Internet2 Driving Exemplary Applications IDEA Award* for innovation enabling transformational progress in research, teaching, and learning and the *Society of American Archivists Jameson Award* for addressing major challenges of managing, preserving, and providing access to electronic records.

Karen Green provided editorial and design support for the preparation of this white paper.

About RENCI

RENCI, an institute of UNC Chapel Hill, develops and deploys advanced technologies to enable research discoveries and practical innovations. RENCI partners with scientists, policy makers, and industry to engage and solve the problems that affect North Carolina, the U.S., and the world. RENCI is a collaboration involving UNC Chapel Hill, Duke University and North Carolina State University. For more information, see **www.renci.org**.

Vol. 2, No. 2 in the RENCI White Paper Series, June 2014. Created in collaboration with the **iRODS Consortium.**

About the iRODS Consortium

The iRODS Consortium works with universities, research organizations, businesses, and government agencies to guide the continued development of iRODS, obtain funding to support that development, and broaden the iRODS user community. Consortium members receive a variety of benefits, including prioritized access to support, training, and consulting, and the opportunity to influence the developmental roadmap of future software releases. For more information, see **www.irods-consortium.org.**

How to reference this paper:
Fortner, B., Ahalt, S., Coposky, J., Fecho, K., Krishnamurthy, A., Moore, R., Rajasekar, A., Schmitt, C., and Schroeder, W. (2014). Control Your Data: iRODS, the integrated Rule-Oriented Data System. RENCI, University of North Carolina at Chapel Hill. Text. **http://dx.doi.org/10.7921/G0BG2KWZ**.

References

Ahalt, S., Bizon, C., Evans, J., Erlich, Y., Ginsberg, G., Krishnamurthy, A., Lange, L., Maltbie, D., Masys, D., Schmitt, C., & Wilhelmsen, K. (2014). Data to Discovery: Genomes to Health. A White Paper from the National Consortium for Data Science. RENCI, University of North Carolina at Chapel Hill. doi: 10.7921/G03X84K4. **www.renci.org/news/ncds-white-paper**. [Accessed February 27, 2014]

Amazon S3. **aws.amazon.com/s3**. [Accessed March 5, 2014]

Apache® HDFS (Hadoop® Distributed File System). **hadoop.apache.org/docs/r1.2.1/hdfs_design.html**. [Accessed March 5, 2014]

API (Application Programming Interface). API – application program interface. **www.webopedia.com/TERM/A/API.html**. [Accessed April 15, 2014]

Barg, I., Scott, D., & Timmermann, E. (2011). NOAO E2E integrated data cache initiative using iRODS. In I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots (Eds.), Astronomical Data Analysis Software and Systems XX. ASP Conference Proceedings, Vol. 442, pp. 497-500. San Francisco, CA, USA: Astronomical Society of the Pacific.

BIOS (Basic Input/Output System) Software. **www.webopedia.com/TERM/B/BIOS.html**. [Accessed April 15, 2014]

Chiang, G.T., Clapham, P., Qi, G., Sale, K., & Coates, G. (2011). Implementing a genomic data management system using iRODS in the Wellcome Trust Sanger Institute. BMC Bioinformatics, 12 (1), 361. **www.biomedcentral.com/1471-2105/12/361**. [Accessed July 31, 2013]

Data DirectTM Networks WOS (Web Object Scaler). **www.ddn.com/products/object-storage-web-object-scaler-wos**. [Accessed March 5, 2014]

GSI (Grid Security Infrastructure). See Murri R. (2011). Grid Security Infrastructure. Grid Computing Competence Center. **www.gc3.uzh.ch/edu/lsci2011/gsi.pdf**. [Accessed April 15, 2014]

GUI (Graphical User Interface). See O'Neill S. (undated). The evolution of the desktop GUI. CIO; article 690686. **www.cio.com/article/690686/The_Evolution_of_the_Desktop_GUI**. [Accessed April 15, 2014]

Hammond, J.S. (2009). Best practices: improve development effectiveness through strategic adoption of open source. Cambridge, MA: Forrester Research, Inc. **http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CCoQFjAA&url=http%3A%2F%2Fdownload.microsoft.com%2Fdownload%2F7%2F1%2FB%2F71BEC711-FBB4-400A-984C-0DCCC36248E0%2FBP-OpenSource.pdf&ei=1pUPU_DDJOTS2gWhlIG4BA&usg=AFQjCNEHbpfzP3MeToVZg0P4o_6JnRR6EQ&bvm=bv.61965928,d.b2I**. [Accessed February 27, 2014]

Hedges, M., Hasan, A., & Blanke, T. (2007). Management and preservation of research data with iRODS. Proceedings of the ACM first workshop on CyberInfrastructure: information management in eScience, pp. 17-22. New York, NY, USA: ACM.

iRODS (integrated Rule-Oriented Data System) Consortium. **irods-consortium.org**. [Accessed March 5, 2014]

Kerberos. **web.mit.edu/kerberos/**. [Accessed March 5, 2014]

LDAP/PAM (Lightweight Directory Access Protocol Server/Pluggable Authentication Module). **wiki.debian.org/LDAP/PAM**. [Accessed March 5, 2014]

Liu, S., Schulze, J. P., Herr, L., Weekley, J. D., Zhu, B., VanOsdol, N., Plepys, D.M., & Wan, M. (2011). CineGrid Exchange: A workflow-based peta-scale distributed storage platform on a high-speed network. Future Generation Computer Systems, 27 (7), 966-976.

Object Stores. See Dineley D. (2013). What is object storage? Highly scalable, simple, cheap, distributed storage for the cloud. InfoWorld, April 5, 2013. **www.infoworld.com/t/data-center/what-object-storage-215778**. [Accessed April 15, 2014]

Quantum® Scalar. **www.quantum.com/products/tapelibraries/index.aspx**. [Accessed March 5, 2014]

Rajasekar, A., Moore, R., Hou, C. Y., Lee, C. A., Marciano, R., De Torcy, A., Wan, M., Schroeder, W., Chen, S. Y., Gilbert, L., Tooby, P., & Zhu, B. (2010a). iRODS Primer: integrated Rule-Oriented Data System. Synthesis Lectures on Information Concepts, Retrieval, and Services, 2 (1), 1-143.

Rajasekar, A., Moore, R., Wan, M., Schroeder, W., & Hasan, A. (2010b). Applying rules as policies for large-scale data sharing. Proceedings of the UKSim/AM SS First International Conference on Intelligent Systems, Modelling and Simulation (ISMS), pp. 322-327. Washington, DC, USA: IEEE Computer Society Washington.

Rajasekar, A., Wan, M., Moore, R., & Schroeder, W. (2006). A Prototype Rule-based Distributed Data Management System. HPDC workshop on Next Generation Distributed Data Management, May 2006, Paris, France.

RDBMS (Relational DataBase Management System). RDBMS – relational database management system. **www.webopedia.com/TERM/R/RDBMS.html**. [Accessed April 15, 2014]

Schnase, J. L., Webster, W. P., Parnell, L. A., & Duffy, D. Q. (2011). The NASA Center for Climate Simulation Data Management System. 2011 IEEE 27th Symposium on Mass Storage Systems and Technologies MSST, pp. 1-6. Washington, DC, USA: IEEE Computer Society Washington.

Schmitt, C., Wilhelmsen, K., Krishnamurthy, A., Ahalt, S., & Fecho, K. (2013). Security and Privacy in the Era of Big Data: iRODS, a Technological Solution to the Challenge of Implementing Security and Privacy Policies and Procedures. RENCI, University of North Carolina at Chapel Hill. doi: 10.7921/GOH41PBR. **http://renci.org/wp-content/uploads/2014/02/0313WhitePaper-iRODS.pdf**. [Accessed November 27, 2013].

SQL (Structured Query Language). SQL – structured query language. **www.webopedia.com/TERM/S/SQL.html**. [Accessed April 15, 2014]

SunTM SAM-FS. **www.sam-fs.com**. [Accessed March 5, 2014]

Ward, J., Wan, M., Schroeder, W., Rajasekar, A., de Torcy, A., Russell, T., Xu, H., & Moore, R. (2011). The integrated Rule-Oriented Data System (iRODS 3.0) Micro-service Workbook. DICE Foundation: ISBN: 9781466469129.