MAGIC Meeting Minutes
March 6, 2013

Attendees

Rich Carlson	DOE/SC
Yong Chen	
Chris Jordon	Texas Computing Lab
Dan Katz	NSF
Dries Kimpe	ANL
Miron Livny	OSG
David Martin	Northwestern U.
Shawn McKee	U. Mich.
Grant Miller	NCO
Reagan Moore	RENCI
JP Navarro	
Lavanya Ramakrishnan	LBNL
Don Riley	U. Maryland
Martin Swany	U. Indiana
Kevin Thompson	NSF
Von Welch	Indiana U.

Action Items

Proceedings

This MAGIC Meeting was chaired by Rich Carlson of DOE/SC and Dan Katz of the NSF.
Ruth Pordes organized a session on Storage management with Reagan Moore, Dries Kimpe, and Shawn McKee as presenters.

The briefings are summarized below. For the complete briefings, please see the MAGIC Wiki site under the March 6, 2013 Meeting at:
http://www.nitrd.gov/nitrdgroups/index.php?title=Middleware_And_Grid_Interagency_Coordination_(MAGIC)#title

Managing Distributed Data Resources with Middleware: Reagan Moore

Use cases for managing distributed data are illustrated by several use cases, particularly genomics and physics uses:

- Broad Institute organizes genomics data from multiple projects, with automated processing into standard data forms and storage of results
- Wellcome Trust Sanger Institute organizes genomics data from multiple projects, with automated processing into standard data forms and storage of results
- UNC-CH Genomics Data Grid organizes genomics data from multiple projects, with automated processing into standard data forms and storage of results

c/o National Coordination Office for Networking and Information Technology Research and Development
Suite II-405 · 4201 Wilson Boulevard · Arlington, Virginia 22230
Phone: (703) 292-4873 · Fax: (703) 292-9097 · Email: nco@nitrd.gov · Web site: www.nitrd.gov

- NSF-iPlant Collaborative accesses a wide variety of distributed data depositories
- T2K QMUL neutrino experiment aggregates small data files before archiving and distributes data from Japan to London
- BaBar High Energy Physics replicates 2 Petabytes of data between archival storage systems at Lyons, France and SLAC

Observatory and Library use cases include:

- NSF Ocean Observatories which manage real-time sensor data streams and automatically archive a copy at the NCDC
- National Optical Astronomy Observatory which archives images from telescopes in Chile on storage systems in Arizona and Illinois.
- National Climate Data Center (NCDC) which manages data ingestion from multiple external projects and stores it across multiple types of storage systems
- Texas Digital Libraries which federates storage systems across university libraries in Texas.
- French National Library which migrates digital holdings across storage system technologies

Simulation and Medical use cases include:

- NASA Center for Climate Simulations which provides access to major digital holdings (MODIS)
- XSEDE which requires high performance read/write to local disk for PB of simulation information per day
- Australian Research Collaboration Service distributes observational data sets to the university with associated experts
- UK e-science data grid which decouples storage resources from compute resources
- Sickkids Hospital manages HIPAA data across multiple storage systems

Generic operations of the storage systems include:

- High performance access to disk caches while computing
- Management of storage space
- Archiving data to alternate storage systems
- Distribution to multiple sites
- Access to data at multiple sites
- Organization of data into collections that span storage systems
- And others

Generic approaches for managing storage systems include:

- Defining the global logical name spaces
- Defining the operations applied on each global logical name space
- Defining the virtualization mechanisms
- Defining the management policies
- Defining the procedures
- Defining federation mechanisms

Middleware systems address a wide range of users, objects,…, operations, and virtualization mechanisms. Interoperability mechanisms are applied to data, information and knowledge subject to existing policies for control and execution. Specific protocols and standards moderate authentication, data access, data manipulation, workflows, networks, clients, storage access, messaging, vocabulary, and management.

Future architectures will address:

- Managing workflows including sharing the process instead of the data
- Storage controllers with embedded processing
- Future Internet Architecture

Storage Management Discussion: Dries Kimpe

Data management is a complex dynamic process involving storage, retrieval, protection, provenance, and efficiency (faster and cheaper). Increased data output and increased network capacities have made storage a major component of overall data management costs. This led to creation of multiple QoS levels such as Microsoft Azure, Amazon glacier, and Burst buffers. The storage system needs to be responsive to performance, provenance, indexing, and archiving. This requires modernizing I/O APIs which have been static for 25 years. Once we have data we need the ability to:

- Find data: traditional directory organization does not scale, databases do not suffice for everyone. The storage system needs to be involved with the indexing and search
- Analyzing data: improved efficiency/cost may require moving the computation to the data or recomputed may be cheaper than fetch. This creates new scientific workflows.

Distributed Storage from the WLCG/ATLAS perspective: Shawn McKee

The LHC experiments have each developed their own distributed data management tools and infrastructure. Atlas currently uses a system called DQ2 to provide meta-data functions. DQ2 provides bookkeeping, managing data movement between end-points, and enforcing data-access and user quotas. DQ2 relies on other tools including FTS and PANDA (a workload management system). ATLAS maintains a Logical File Name (LFN) Catalog for each storage site that maps LFNs to physical instances.

Datasets are an aggregation of data typically spanning more than one physical file, typically 1-10000 files. Files are the basic data unit and they are immutable.

The data management service includes the client API, site services, central catalogs, and tools such as OSG, LHC computing Grid and NorduGrid. DQ2 manages dataset details: which sites have which data sets, dataset status, composition, ACLs, and quotas. The File Transfer Service (FTS) typically interacts with end-site storage through an abstraction. End-site data storage has been managed by a concept of space-tokens but ATLAS is moving to a more flexible system.

Atlas will be shut down for a long period to provide upgrades. They will move to:

- DQ2 => Rucio This will provide improvements in policy, accounting, replication, permissions, and scalability
- FTS 2.0 => FTS 3.0 This removes the “overlay” channel model and provides more protocol support
- PANDA will move to incorporate network awareness.

ATLAS is experimenting with moving jobs to the data, Federating access to storage allows better resiliency and resource usage. Xrootd is the protocol to implement this structure.

Some Thoughts:

- Storage needs some attributes exposed to make it usable and manageable.
- All issues are end-to-end
- Storage is not separate from the network for distributed science

Discussion among MAGIC members indicated that:

- To move to next generation storage systems, standards are needed probably several sets of standards to not preclude natural evolution and development of the technologies.
- We have to consider if we want to establish a dictatorship (establish and enforce standards) or create a technical authority with the power to get things done (which implies funding).
- We are trying to move to a storage concept that allows several applications to be layered on it. This will require an improved indexing capability for the storage system.

Upcoming Meetings

March 11-15 : OGF meeting in Charlottesville
March 12-15: OSG All Hands Meeting: Indianapolis
July 22-25 XSEDE Meeting, San Diego

OSG and XSEDE are offering a summer school to provide understanding of the principles, concepts and applications. A link to this meeting is provided on the XSEDE Web page.

Next MAGIC Meetings

- April 3, 2:00-4:00, NSF, Room II-415
- May 1, 2:00-4:00, NSF, Room II-415