Managing Distributed Data Resources with Middleware

Reagan Moore
UNC-CH

iRODS
http://irods.diceresearch.org

Use Cases – Genomics, Physics

- Broad Institute – they organize genomics data from multiple projects, automate the processing of the data into standard forms, and store results in an archive.
- Wellcome Trust Sanger Institute – they organize genomics data across multiple projects, automate the processing of the data into standard forms, and store results in an archive.
- UNC-CH Genomics data grid – they organize genomics data from multiple projects, automate the processing of the data into standard forms, and store results in an archive.
- NSF iPlant Collaborative – they access a wide variety of distributed data repositories (typically databases), extract
- T2K QMUL neutrino experiment – they aggregate small data files before archiving, and distribute data from Japan to London
- BaBar High Energy Physics – they replicated 2 Petabytes of data between archival storage systems at SLAC and Lyon, France

Use Cases – Observatories, Libraries

- NSF Ocean Observatories Initiative – they manage real time sensor data streams, and automate the archiving of a copy of the data at the National Climatic Data Center.
- National Optical Astronomy Observatory – they archive images taken on telescopes in Chile on storage systems in Arizona and Illinois.
- National Climatic Data Center – they manage data ingestion from multiple external projects, replicate their digital holdings, and store the data across multiple types of storage systems.
- Texas Digital Libraries – they federated storage systems across university libraries in Texas, replicating data into the Texas Advanced Computer Center
- French National Library – they periodically migrate their digital holdings across storage system technologies

Use Cases – Simulation, Medical

- NASA Center for Climate Simulations – they provide access to major digital holdings (MODIS Moderate Resolution Imaging Spectroradiometer 650-Terabyte data set) for access from the Earth Systems Grid through a file system interface
- XSEDE – they require high-performance read/write to local disk for generation of petabytes of simulation data per day, and then replicate data across multiple independent data archives
- Australian Research Collaboration Service – they distribute observational data sets to the university that has the associated experts, and then share data across storage systems
- UK e-Science data grid – they decoupled storage resources from compute resources, and distributed data products generated by each computation into a national storage environment.
- Sickkids Hospital (Ontario, Canada) – they manage HIPAA data across multiple storage systems

Generic Operations

1. High performance access to disk caches while computing
2. Management of storage space
3. Archiving of data to alternate storage systems
4. Distribution of data to multiple sites
5. Replication of data sets
6. Access to distributed data sets
7. Enforcement of access controls in the distributed environment
8. Automated application of processing steps to data sets
9. Automated management of caches
10. Organization of data into collections that span storage systems

Generic Approach

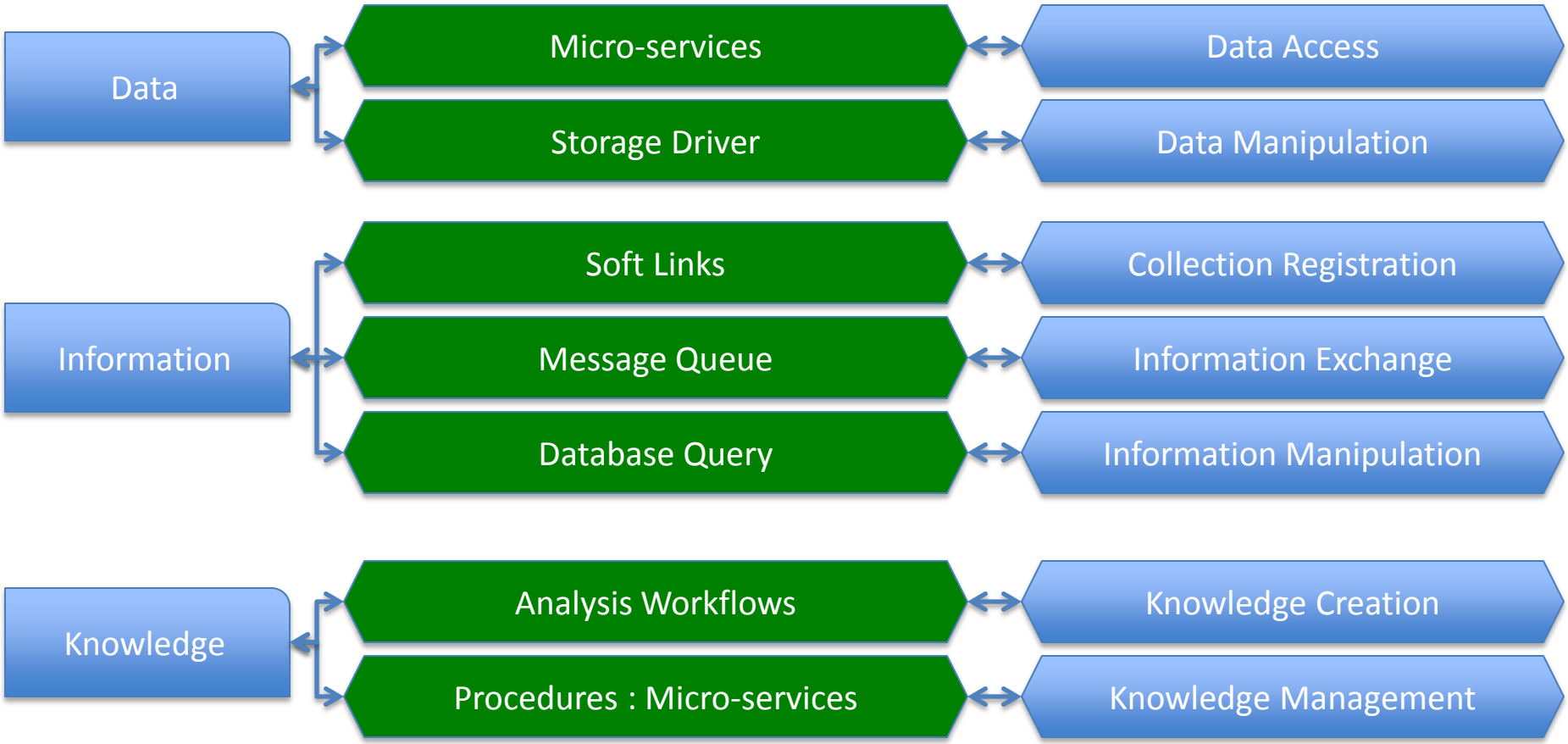
- Define the global logical name spaces that will be used to identify resources. A mapping will be required from the global logical name space to the physical naming conventions within each type of storage system.
- Define the operations that are applied on each global logical name space. Examples include manipulation of entities, aggregation of entities, access controls on entities.
- Define the virtualization mechanisms that will enable application of the operations across the multliple hardware and software systems.
- Define the management policies that will be enforced on the name spaces (quotas, integrity, chain of custody, authenticity, arrangement)
- Define the procedures required to enforce the policies (typically applied across multiple name spaces as a system level operation)
- Define federation mechanisms

Middleware Systems

Name Space	Operations	Virtualization mechanism
Users	Authentication, authorization, groups	GSSAPI / PAM
Objects	Partial I/O, move, copy, replicate, share	Posix I/O & staging
Collections	Organization, Browsing	System metadata
State information	Add, update, delete, query	Catalog interface to DBMS
Resources	Load leveling, fault tolerance, grouping	Storage drivers
Policies	Management, administrative, verification	Policy language
Procedures	Basic functions on each name space	Workflows

Interoperability Mechanisms

Policies control execution of each interoperability mechanism



DFC Interoperability Layers

Authentication	PAM, GSSAPI	InCommon, GSI, Kerberos, Shibboleth, LDAP
Data Access	Micro-Services	DataONE, Data Conservancy, CUAHSI, NCDC
Data Manipulation	Format Drivers	NetCDF, HDF5, THREDDS, ERDDAP
Workflows	Micro-Services	Kepler, NCSA Cyberintegrator, Taverna, NCSA Polyglot
Networks	Network Drivers	HTTPS, TCP/IP, Parallel TCP/IP, RBUDP

DFC Interoperability (2)

Clients	Jargon	Web browsers, Web Services, Workflows, FUSE, Synchronization, Mediawiki
Storage Access	Storage Drivers	File Systems, Tape Archives, Object Stores, Cloud Storage
Messaging	Micro-Services	AMQP, iRODS
Vocabulary	Micro-Services	HIVE, (Cheshire)
Management	Procedures	(RDA Policies), (ISO 16363 Criteria)

State Information

- File identifier
- File creation time
- File modification time
- File size
- File checksum
- File storage location
- File type
- File retention period
- File disposition
- File version
- Replica number
- Replica location
- Replica creation time
- Replica checksum
- Collection identifier (logical groups of files)
- Storage system identifier
- Storage groups
- Storage quota per user or user group
- User identifier
- User groups
- Access controls
- Collection sticky bit (for inheritance of access controls)
- Audit trail

Future Architectures

- Manage workflows
  - Registration, provenance, re-execution, sharing
  - Share the process instead of the data
- Storage controllers with embedded processing
  - Automate extractions of features from data sets
  - Index data based on features that are present
- Future Internet Architecture
  - Embed policies in the network
  - Map virtual collection to virtual network
  - Access by file name, ACLs, caching, debugging