Reagan W. Moore San Diego Supercomputer Center
22-September-2005

Status of This Memo: Informational

This memo provides information to the Grid community specifying the services that are being developed in the International Virtual Observatory Alliance. An attempt is made to identify the interactions with Global Grid Forum standards groups that will aid in the development of grid compliant IVOA services.

Copyright Notice

Building IVOA Services for the Astrophysics Community

Abstract

The integration of grid technology in the development of International Virtual Observatory Alliance (IVOA) compliant services is desired to ensure interoperability across compute and storage resources. We define the current IVOA services, the approaches that are being followed in their development, and the opportunities for interaction with GGF standards groups.

Contents

Abstract 1
1. Virtual Observatory Applications 3
2. Virtual Observatory Services 4
3. Mapping of Requirements onto Global Grid Forum standard groups 6
3.1 Infrastructure Standards Groups 6
3.1.1 Ipv6 6
3.1.2 Network Measurement 7
3.1.3 Data Transport 7
3.1.4 Grid High-Performance Networking 8
3.1.5 Network Measurement for Applications 8
3.2 Data Standards Groups 9
3.2.1 Data Access and Integration Services 9
3.2.2 Grid File Systems 10
3.2.3 Data Format Description Language 11
3.2.4 GridFTP 11
3.2.5 Grid Storage Management 11
3.2.6 Information Dissemination 12
3.2.7 OGSA Data Replication Services 12
3.2.8 Transaction Management 13
3.2.9 OGSA Data 13
3.2.10 Byte IO 13
3.3 Compute Standards Groups 14
3.3.1 Grid Resource Allocation Agreement Protocol 14
3.3.2 Job Submission Description Language 14
3.3.3 Grid Scheduling Architecture 14
3.3.4 OGSA Basic Execution Services 14
3.4 Architecture Standards Groups 15
3.4.1 Open Grid Services Architecture 15
3.4.2 Grid Protocol Architecture 15
3.4.3 OGSA Naming 15
3.5 Applications Standards Groups 15

3.5.1 Grid Remote Procedure Call 15
3.5.2 Grid Information Retrieval 16
3.5.3 Distributed Resource management Application API 16
3.5.4 Simple API for Grid Applications 16
3.5.5 Grid Checkpoint Recovery 16
3.6 Management Standards Groups 16
3.6.1 Application Contents Service 16
3.6.2 Configuration Description, Deployment, and Lifecycle Management 16
3.6.3 Grid Economic Services Architecture 16
3.6.4 OGSA Resource Usage Service 17
3.6.5 Usage Record 17
3.7 Security Standards Groups 17
3.7.1 Open Grid Service Architecture Authorization 17
3.7.2 OGSA-P2P-Security 17
3.7.3 Firewall Issues 17
3.7.4 Trusted Computing 17
4. Summary 17
5. Author Information 17
6. Acknowledgements 18
7. Glossary 18
8. Intellectual Property Statement 18
9. Full Copyright Notice 18
10. References 19

1. Virtual Observatory Applications

The Astronomy community provides digital reference sets for use by researchers and educators. The observational data consist of images taken by many different projects, each of which organizes their images in independent survey archives. Information about star and galaxy location, magnitude, variability, morphology, etc. is organized as records in catalogs. The image surveys typically have on the order of 5 million images and 10 Terabytes of data. The catalogs may have a billion records.

The Virtual Observatory is an attempt to federate the existing catalogs and image archives, in order to improve access to observational data for all astronomers and to facilitate access by educators. The Virtual Observatory is developing standards under the aegis of the International Virtual Observatory Alliance.

The approach that is being followed to provide uniform access to the Astronomy community digital holdings consists of:

- Development of a standard vocabulary (semantics) for describing the physical quantities stored in each catalog. The standard vocabulary terms are called “Uniform Content Descriptors”.
- Development of a standard ontology describing how terms are related. This requires the ability to impose an organization structure across the terms, going from general categories to specific instances.
- Development of a standard data model (encoding format) for each type of measurement. Image data is cast into the FITS format, which encapsulates standard header information about the provenance of the image with each image. Tables are cast into the VOTable format for transport.
- Development of a standard query language for issuing spatial, temporal, and semantic queries across the catalogs.
- Development of standard access services for retrieving catalog records or image cutouts. More sophisticated access mechanisms create mosaics of images, or retrieve spectra.
- Development of standard mechanisms for interacting with storage systems (VOStore).
- Development of standard mechanisms for managing a shared collection (VOSpace).
- Development of standard authentication mechanisms for accessing shared collections.
- Development of standard event notification services.

The Virtual Observatory community has experimented with Grid technology. Services have been built upon the WSDL-based grid architecture. Services have also been built using alternate technology, including SOAP-based systems, Java-based systems, digital-library based systems, and http-based systems. The experiences to date indicate that the following criteria are essential for building viable Virtual Observatory services:

- Robustness. This includes the sustainable number of simultaneous users of the system, the stability against network interruptions, the size of queries that can be processed.
- Ease of implementation. An environment is needed that is simple enough for researchers to apply the Virtual Observatory (VO) services to their local collections. This includes minimizing the number of layers of software required to implement a new service, and minimizing the amount of coding that the researcher must write.
- Scalability. The services need to support the manipulation of entire catalogs and analysis of entire image archives. This includes support for bulk operations in which thousands of files or hundreds of thousands of records area accessed, and support for remote processing of data or distributed joins across collections.
- Standards. For large-scale analyses, Grid technology will be used to support the processing of the data. Thus the Virtual Observatory services need to interact with Grid workflow systems and Grid security mechanisms.
- Timeliness. The Virtual Observatory services are being developed and applied today. This has forced the Virtual Observatory community to develop services before the GGF standards are in place.

2. Virtual Observatory Services

The US National Virtual Observatory held a recent summer school (Sept. 19-23, 2005) funded by the National Science Foundation. The summer school provided an introduction to the use of Virtual Observatory compliant services for Astrophysics researchers. An attempt was made to provide information for all levels of the software infrastructure needed to build Virtual Observatory services, including the web technologies used to implement services, the components needed to support web services, and scripting languages.

The types of software systems that the participants were expected to understand are illustrative of the complexity of services-oriented architectures. The following software infrastructure support environment was used:

- Web technology
  - Tomcat – webserver
  - SQL structured query language
  - XMLSchema / Xpath / Xquery / XML Stylesheet Language
  - WS-Security / SAML environment
  - SOAP messages
  - SOAP attachments / DIME, MTOM
  - WSDL services
  - WSRF statefull and asynchronous web services
- Languages
  - Java
    - AXIS - Java web services
      - wsdl2java command to generate stubs,
      - adminclient,
      - starttomcat/stoptomcat webserver control,
      - bouncetomcat to restart web service
    - ANT – Java-based software build tool
      - Java Server Pages
      - Web Application Archives
    - Java SDK 1.4.2_XX – Java software development kit
    - Java3d – support for visualization tools
    - Java class libraries
  - C# programming language (Microsoft.Net and open source Mono version)
  - IDL Interactive data language – libraries to support reduction, analysis, visualization of astronomical data
  - PHP scripting language and interface to MySQL database
  - Python scripting language
- Data management
  - MySQL – database
  - Teragrid grid technology and data grid technology

The Virtual Observatory services and applications build upon the above software infrastructure to implement Virtual Observatory services. For the services to provide a reasonable set of capabilities, the Virtual Observatory first had to standardize the encoding format for data, and the meaning of the semantic terms used to describe physical quantities that are stored in the catalogs. The services then were designed to operate on the standard data encoding format using the standard semantic terms. The current Virtual Observatory standards for data format and semantics are:

- Data Encodings

  - VOTable – XML standard for the interchange of data represented as a set of tables. The standard separates data and metadata, and uses XML tags to allow validation and transformation of an input document.
  - FITS – Image encoding standard
  - ADQL – Astronomy Data Query Language extensions to SQL for spatial queries

- Virtual Observatory Metadata
  - VOResource description standard for registering catalogs. Components include the access URL for the information contained in the resource.
  - UCD – IVOA controlled formal vocabulary for astronomical data. Uniform Content Descriptors are built by combining words from the controlled vocabulary to define the type of physical quantities. The syntax specifies the general category and specific instance as "category;instance;type". Examples are "phys.temperature" for temperature, "phot.flux;em.radio;arith.ratio' for a ratio of radio fluxes.
  - STC – Space-Time Coordinate descriptors

The development of standard Virtual Observatory services proceeded through two steps: the creation of standard access mechanisms to both catalogs and image archives, and the creation of more sophisticated services that manipulated the data. The standard access mechanisms are in some sense generic Virtual Observatory Protocols, and candidates for expression as Grid services. The more sophisticated services use the Virtual Observatory protocols and support astronomy specific functions. To federate across multiple catalogs, Portals were developed that manage the results of the application of Virtual Observatory services.

- Virtual Observatory Protocols
  - SIAP - Simple Image Access protocol returning pointers to the requested information encapsulated in a VOTable format.
  - SSAP - Simple Spectral Access Protocol for one-dimensional spectra, time series, and Spectral Energy Distributions. Returns pointers to the information encapsulated in a VOTable format.
  - Cone-Search – catalog access
  - VOEvent – event notification service
  - VOStore – access to image archives
  - VOSpace – management of shared data collections

- Virtual Observatory Services
  - Montage – dynamic creation of mosaics of images
  - Mirage
  - Topcat
  - VOPlot - visualization
  - Downloadmanager script to support queries on Cone/SIAP services and file download
  - ReadVotable java program to list the rist line of a VOTable and all its fields
  - Mosaic – creation of mosaics of images retrieved from image archives
  - Voogle – Virtual Observatory interface
  - WESIX – extract source catalogs from stacked Sloan Digital Sky Survey images
  - Object classifiers
  - Outlier-detection

- Portals
  - OpenSkyQuery – correlation service for cross matching records from major catalogs in small areas of the sky. Parameters are selected using criteria based on comparisons between values.
  - SkyNode – standard interface to existing catalogs

- Virtual Observatory Registry – OAI-PMH compliant registry of available Skynodes
- SkyPortal – interface for composing ADQL queries
- IRAF Image Reduction and Analysis Facility – general purpose software system for the reduction and analysis of optical/infra-red astronomical data. Commands include task invocation, parameter specification, graphics/image display, interactive and batch execution, command language scripting

3. Mapping of Requirements onto Global Grid Forum standard groups

The Virtual Observatory Protocols are candidates for implementation through GGF standard grid services. We examine each of the Global Grid Forum Standards groups and assess the relevance of the activity to the Virtual Observatory community. We express this relevance through the expected utility of the proposed standard, how the standard would interact with Virtual Observatory requirements, and whether there are current concerns about the viability of the proposed standard. The generic concerns are:

Utility
- How the proposed GGF standard might be used to implement a Virtual Observatory Protocol.
- What advantages the GGF standard would provide over the current implementation.

Virtual Observatory requirements
- Support for interactions with existing catalogs and image archives.
- Interoperability with Virtual Observatory community standards.

Viability concerns:
- Ease of implementation of the GGF standard.
- Robustness of the standard.
- Scalability of the resulting service to support manipulation of 10-Terabyte collections and catalogs with 1 billion records.
- Timeliness.

For each concern, we outline impacts on the Global Grid Forum Standards Groups. The impacts will evolve based upon the standards created by each group. This will require iteration of requirements between the Astronomy and Applications Astro Research Group and each standards group.

3.1 Infrastructure Standards Groups

The over-riding Virtual Observatory requirement is how to ensure backward compatibility between the new versions of the standards and the prior versions of the standards. If a new standard requires a new protocol, then implementations are need that will be able to work with both the old and new protocols while the infrastructure is upgraded.

3.1.1 Ipv6

This GGF group examines how to create network-neutral grid services. The Ipv6 standard addresses expansion of the network address space and provides additional network management functions. The document GFD.40 on “Guidelines for IP version independence in GGF specifications” provides essential information on how to build services that can handle the new Ipv6 addressing scheme. As such, this enables infrastructure independence across multiple versions of IP networks.

Utility:
- We expect the use of Ipv6 will be mandated in future Virtual Observatory environments to ensure the ability to continue to be able to address remote storage systems over wide area networks. The standard from this group should be used for all Virtual Observatory services.

Virtual Observatory requirements:

- Compatibility of Ipv6 addressing with logical name spaces for resources in the VOStore interface. The IP address forms the physical address for each storage resource. The VOStore logical resource name is the invariant address. Thus it will be possible to map to new forms of the IP address.

Viability:

- Protection against malicious users would be greatly aided by the ability to track the source of each network message. This means added functionality to tie all delivered messages to the origin IP address.

3.1.2 Network Measurement

This group is proposing standards for a common data model and format for reporting network measurements. They propose serializing measurement data in an XML schema, using standard attribute names to characterize the meaning of the data. This approach is compatible with the use of XML schema by the digital library community.

Utility:

- Multiple types of end-to-end measurements will be used by Virtual Observatory environments to optimize use of networks. An example is bulk-access to image surveys. Headroom measurements on networks will be needed to maximize throughput. The VOSpace, SkyPortal, OpenSkyPortal should evaluate the GGF standard.

Virtual Observatory requirements:

- Of interest is the ability to map the Network Measurements schema to XML standards that may be required by the digital library community. An example is the METS Metadata Encoding Transmission Standard. The Network Measurements schema could be cast as a METS profile. It will then be possible to manage event information about network performance in the same structures that are used to manage digital libraries of astrophysics derived data projects. An archive of event information could then be manipulated with existing digital library tools.
- A second interest is the ability to characterize the integrity of network transmissions when moving data for analysis. The goal is to be able to make assertions about the reliability of the data transfers (from disk to disk) and the path over which the data was transmitted. Part of the workflow information for optimizing bulk operations should be the IP address from which the material was sent. This is related to tracking the status of a file. When a file is moved over a less reliable network, additional verification of the integrity of the file will be required. Network measurements need to include estimates of reliability of the transfer.

Viability:

- How will network measurements be incorporated in cost models? We will need the ability to define the cost of bulk analyses, including data movement costs.
- How can the measurements be linked directly to the workflow status? The grid can be viewed as a workflow virtualization mechanism, enabling the execution of workflow processes at multiple sites. The distribution of processing between sites will depend upon the network throughput measurements.

3.1.3 Data Transport

This group promotes the creation of new standards for secure, robust, high-speed transport of data.

Utility:

- Virtual Observatory environments will need the secure, robust, high-speed transport of both data and metadata. The emerging bulk transport mechanisms for the Virtual Observatory should consider this standard.

Virtual Observatory requirements:
- Will the data transport mechanisms include the ability to aggregate metadata before transport? A similar capability is needed to pack small files before transport.
- How will end-to-end performance be assured? The transmission bottleneck may be the number of sources, the number of parallel I/O streams, or the number of receiving nodes?
- How will the data security interact with Virtual Observatory security requirements, namely authentication against the distinguished user name space that the Virtual Observatory environment is using to identify individuals?

Viability:
- Virtual Observatory environments will need the ability to interoperate across multiple data transport protocols. Given that the congestion algorithms used by the transport mechanisms will probably be different across different versions, metrics are needed to evaluate the preferred transport mechanism for access to a given remote storage system. The metrics for evaluating which transport to use will in turn impact the standard schema developed by the network measurement group.
- The VOSPace service will need to base its cost models on the expected data transport performance, and will need assessment tools to do the cost evaluation.

3.1.4 Grid High-Performance Networking
This group serves as a liaison to networking standards bodies.

Utility:
- As a liaison, it could promote interoperability standards between transport protocols. A reasonable goal is to minimize the impact of transport changes on preservation environment software as standards evolve.

Virtual Observatory requirements:
- To interact with legacy systems, a Virtual Observatory environment needs the ability to use any standard network protocol. As protocols evolve, an upgrade path is needed that minimizes the impact on preservation software.

Viability:
- Ease-of-use dictates that the choice of network be automated. What mechanisms will be provided to enable the Virtual Observatory processing pipelines to select the appropriate network, and then use the resulting protocol? Will there be mechanisms to negotiate network protocol between services?

3.1.5 Network Measurement for Applications
This group focuses on the development of network-aware middleware, which is able to exploit knowledge about network parameters including bandwidth, latency, and jitter.

Utility:
- The ability to use the correct protocol is important for delivery of advanced data products such as real-time data from remote telescopes. The Virtual Observatory portals should consider this standard for delivering advanced products.

Virtual Observatory requirements:
- The implication is that the type of data product will influence the choice of transport mechanism. This in turn implies that the Virtual Observatory environment will have to

support multiple transport mechanisms, and that the choice will depend on multiple parameters:

- Network performance
- Size of data to move
- Type of data to move

A generic solution is desired that minimizes the number of transport protocols that must be maintained by the Virtual Observatory environment.

Viability:

- The incorporation of network performance knowledge into Virtual Observatory processing pipelines implies that characterization of the network will use standard attributes. Will these standard attributes be uniform across all transport protocols, or will the processing pipelines require analysis of a different set of attributes for each protocol that is used?

3.2 Data Standards Groups

Virtual Observatory environments employ a radically different perspective on management of state information compared to grid services. In a shared collection, all state information is directly associated with each registered item. Grids organize state information in service specific catalogs. A major challenge is migrating from multiple independent service information catalogs into a VOSpace catalog to describe the properties of shared collections.

Most of the Grid standards relate to data access rather than data management. The challenge of building a shared collection (VOSpace) is the guarantee that descriptive metadata will be correctly associated with each registered file, that operations on each file can be tracked to check status, and that the grid state information is appropriately updated after every operation. At the moment, the management of consistent state information is left up to the VOSpace or processing pipeline implementation. The development of data grids which manage the state information self-consistently will make both VOSpace and processing pipelines easier to implement. An environment that imposes consistent state information management is possible if the following constraints are implemented:

- Data grid ownership of material. Unless the data grid owns the files that are stored at a remote site, it will not be possible to track operations done on the files.
- Data grid virtualization of trust management. The data grid acts as the surrogate for the shared collection manager to ensure that access controls established for the shared collection will be followed no matter where in the data grid the material is stored.
- Data grid control of the name spaces used to identify storage resources, files, users, and metadata. This requirement makes it possible to provide persistent naming for use by the shared collection, even when material is migrated to new storage solutions.
- Decoupling of the presentation interface protocols from the storage repository access protocols. This makes it possible to manage access mechanisms independently of the choice of storage repository. The expectation is that more cost-effective storage solutions will be found as technology evolves. Migrating to the new storage systems should not impact the access mechanisms.

These requirements enable complete data virtualization, the decoupling of the management of data from the storage systems in which they are deposited.

3.2.1 Data Access and Integration Services

This group promotes standards for consistent access to autonomously managed databases.

Utility:

- Virtual Observatory environments will also manipulate databases, and will need standard services for doing joins across multiple databases. The SkyNode, ConeSearch Virtual Observatory services should consider this standard.

Virtual Observatory requirements:

- A standard access mechanism is needed for both legacy and new databases. This requires the ability to write drivers to interact with new database technology, and the ability to maintain the drivers for interacting with prior database technologies. At the moment, separate DAIS access services are used for each type of database. A single access service is desired.
- Stable design parameters are needed for the access mechanisms. As new database features are added (such as descriptions of relationships that must be true for a metadata attribute name to be meaningful), the interface should still support prior access mechanisms.
- Support for bulk operations on databases is needed. Since catalogs will manage hundreds of millions to billions of records, the ability to register new metadata in bulk is needed.
- Mechanisms are needed to assert the validity of the metadata. This includes correspondence to header information encapsulated in FITS images, self-consistency checks on the metadata (presence of required attributes), and integrity checks that the metadata is not corrupted (checksums).
- Access controls on metadata. Roles are needed for creating and updating metadata along with authentication and authorization of use of the roles. Typically, a project defines a standard procedure for creating metadata, and executes the procedures in a standard pipeline processing system.

Viability:

- The preservation environment will manage the lifetime of the archive. The grid services will need to interact with the life-cycle management policies of the shared collection.

3.2.2 Grid File Systems

This group promotes standards for describing and organizing file-based data.

Utility:

- The Grid File System Directory Service could manage the namespace of federated and virtualized data across file system resources. The VOSpace Virtual Observatory service should evaluate this standard.

Virtual Observatory requirements:

- Support is needed for files with header information, such as FITS files. Links between entries in the FITS headers and information stored in the grid file system will need to be consistently managed.
- Support is needed for automating interactions with the data. This implies the ability to track all operations, accesses, updates performed upon both files and the descriptive metadata. It is vital in a shared collection that the link between descriptive metadata and the corresponding files be maintained across all operations performed upon the files.
- Support is needed for chain of custody. The name space identifying researchers should be decoupled from the user name space managed by the storage system. The shared collections (DSpace) should manage the distinguished user name space, and store data under its control on a single account within the storage system. This implies that access permissions are controlled above the level of the grid file system.
- Support is needed for integrity through the ability to create and validate checksums.
- Support is needed for execution of data manipulation processes at the remote storage system, such as checksum validation.
- Support is needed for federation across independent grid file systems. This helps minimize risk of data loss due to operational procedures within a single data grid.

Viability:

- Support is needed for minimizing cost of ownership. A middleware implementation is preferred rather than an operating system kernel mod, as middleware is more easily ported across new types of storage systems.

- Support is needed for scalability. A critical element is the ability to query the directory service without actually accessing the storage system to retrieve information about the stored data.
- Support is needed for multiple types of data copies. At a minimum, the shared collection will need the ability to create replicas (synchronized copies), versions (numbered copies), backups (time-stamped copies), and independent copies of files.

3.2.3 Data Format Description Language
This group promotes the creation of an XML-based language to describe the structure of binary and character encoded files and data streams.

Utility:
- This project can create the standard representation for data encoding formats that allows the construction of generic parsing routines. This is turn makes it possible to decouple data parsing from applications that operate on the structures for display and manipulation. The hope is that a new application can define the set of operations that it will perform, map these operations to a standard set supported by the DFDL libraries, and then perform the operations on the structures that have been characterized by DFDL. In particular, a characterization of both FITS files and VOTables is desired.

Virtual Observatory requirements:
- A generic XML-based description is needed that can describe all types of data formats. Of interest are complex encodings such as databases and published documents.

Viability:
- The tools that process the encoding format need to be implemented in a portable language to ensure that the system will function on future operating systems.

3.2.4 GridFTP
This group promotes improvements to the GridFTP protocol for bulk file transfer, including parallel transfer, GSI authentication, and striped transfers.

Utility:
- Reliable data movement will be needed by Virtual Observatory environments. The emerging Virtual Observatory services for manipulation of entire image archives should consider this standard.

Virtual Observatory requirements:
- Bulk transport of small files is needed. This requires packing small files before movement using parallel I/O streams.
- Bulk registration of files is needed. This requires packing metadata about each file (size, name, time stamps), moving the metadata in bulk, and loading into a metadata catalog. An example is the recursive registration of an existing directory into a shared collection.
- Support for execution of remote procedures. An example is metadata extraction, the packing of the metadata into an XML file, the movement of the file, and the registration of the metadata into the shared collection.

Viability:
- The system needs to operate on the name spaces managed by the shared collection (VOSpace).

3.2.5 Grid Storage Management
This group promotes a standard Storage Resource Manager to support dynamic space allocation and file management of shared storage components on the Grid. Capabilities include storage reservation and information on storage availability.

Utility:

- The management of storage resources, including the staging of files and space reservation, will be needed by Virtual Observatory environments. This is a candidate for a standard VOStore protocol.

Virtual Observatory requirements:

- The system should support integrity functions including checksum validation.
- The system should work with the shared collection logical name spaces.

Viability:

- The system should work off of the same metadata catalogs used by the shared collection.
- The system should be portable onto new operating systems.

3.2.6 Information Dissemination

This group is defining the low-level operations needed to support data and event dissemination, and a high-level interface for information dissemination.

Utility:

- The Virtual Observatory is developing an event-notification service for explicit astronomical events. The GGF standard is a candidate for the implementation of a VOEvent service.

Virtual Observatory requirements:

- The integration of event notification with workflow systems can improve the ability to automate processing of event information.
- Grid services support the virtualization of workflows. The ability to maintain information about workflows is essential for tracking the status of processes that have been applied.

Viability:

- A true workflow virtualization system is needed for workflow processes. For the grid services to be useful, they need to be integrated into a coherent system that tracks the results of the application of the services and associates the types of event processing with each processed data object.
- Of interest is interaction with other metadata transport mechanisms, such as the Open Archives Initiative, Protocol for Metadata Harvesting - OAI-PMH.

3.2.7 OGSA Data Replication Services

This group is refining grid service specifications for data replication services, in particular catalogs about data location.

Utility:

- Virtual Observatory environments need to manage multiple replicas to mitigate risk of data loss. The VOSpace Virtual Observatory implementation should consider this standard.

Virtual Observatory requirements:

- The replicas need to use the logical name spaces managed by the shared collection.
- Multiple types of copies of data are needed, including transformative migrations to alternate encoding formats, backups (time-based snapshots), versions (numbered copies). State information is needed to differentiate between these types of copies.
- Support is needed for replication of containers of images.
- Support is needed for replication between independent data grids.
- Support is needed to ensure that descriptive metadata remain linked to replicas.

Viability:

- Virtual Observatory environments associate state information with each registered item. A mapping will be needed between the replica state information catalog and the preservation metadata catalog.

3.2.8 Transaction Management

This group is applying transaction management techniques to grid systems for updates.

Utility:

- A key component of workflows is tracking status of processing. Transaction management would enable the ability to ensure that each process completes, or is rolled back. The Virtual Observatory portals such as IRAF, Montage, OpenSkyQuery should consider this standard.

Virtual Observatory requirements:

- Transaction management needs to be integrated with workflow environments, with resulting state information saved for each item after the application of each workflow process.

Viability:

- The resulting system needs to support bulk operations on collections of files and metadata. The challenge is how to apply transaction processing to groups of files. One approach is to build re-entrant processes, such that the process can be re-executed to correct partial application. The other approach is to roll back the state of the entire submitted group of files and re-try from scratch.

3.2.9 OGSA Data

This group is designing an overall architecture for virtualization of workflows. This includes the message patterns and interfaces for integrating grid services.

Utility:

- The management of state information for workflows is necessary to manage status information during application of Virtual Observatory pipelines. The IRAF, OpenSkyQuery, and Montage portals should consider this design.

Virtual Observatory requirements:

- For consistent state information, shared collections need to track all operations done on data.
- Integrity requires automation of validation mechanisms for checksums.

Viability:

- The workflows will need to manipulate collections of data, manage error conditions, and maintain consistent state information about operations performed upon data.

3.2.10 Byte IO

This group is designing an interface to read sequences of bytes from multiple types of resources.

Utility:

- Virtual Observatory environments need the ability to read data from multiple types of storage systems. This is a second possibility for implementing the VOStore interface. The expectation is that this would provide a minimal data access interface that could be ported in front of any storage system.

Virtual Observatory requirements:

- Current data grids provide support for POSIX byte I/O.

Viability:

- Additional operations are used to support metadata extraction, bulk operations on remote data, aggregation of data. Virtual Observatory environments use more than simple byte I/O in remote operations. Experience is needed on the sophistication of the interface to meet 80% of the Virtual Observatory processing requirements.

3.3 Compute Standards Groups

Virtual Observatory environments apply pipeline processes when ingesting data. These processes can run in a grid environment. Hence there is strong interest in the ability to apply the processes at the locations where there is sufficient compute power and storage for each record series.

3.3.1 Grid Resource Allocation Agreement Protocol

This group is producing a common resource management protocol for advanced reservation of resources.

Utility:

- Distributed processing of workflows will become important as the volume of material to be processed increases. The VOSpace Virtual Observatory service should consider this standard.

Virtual Observatory requirements:

- Optimization of workflows requires the ability to track the location where pipeline processes are executed.

Viability:

- Standard grid services should be sufficient for allocation of grid resources. What is not obvious, is whether allocation mechanisms will be provided for multiple types of storage systems (disk caches, on-line collections, archives).

3.3.2 Job Submission Description Language

This group is specifying an abstract Job Submission Description Language for interacting with popular batch systems.

Utility:

- Interaction with batch systems will be part of large scale processing demands for Virtual Observatory environments. The Virtual Observatory portals should interact with this standard.

3.3.3 Grid Scheduling Architecture

This group is defining a scheduling architecture that can control use of networks, software, storage, and processing units, and the interactions of these systems with data management.

Utility:

- Again scheduling of resources will be important for processing large amounts of material.

3.3.4 OGSA Basic Execution Services

This group is designing a set of Execution Management Services. This is equivalent to the creation of workflow management.

Utility:

- The management of processes executing in a distributed environment is needed for large scale processing. The Virtual Observatory portals should interact with this standard.

Virtual Observatory requirements:

- Consistent pipelines require the tracking of processes applied to each file, and associating the process execution state with each file.

Viability:

- Consistent management of state information resulting from application of grid services is essential in Virtual Observatory environments.

3.4 Architecture Standards Groups

The current architecture is being devised independently of the standards used for data management. An integration of the architecture used for grid services and the architecture used for digital libraries is needed. This requires a change in perspective towards identification of the name spaces needed to manage data rather than execute services. Integration with data management standards is also needed:

- METS Metadata Encoding Transmission Standard
- Archival Information Package
- Open Archives Initiative Protocol for Metadata Harvesting
- FITS headers and VOTable transmission standard

3.4.1 Open Grid Services Architecture

This group is developing an architecture roadmap for grid services, focused on functionality and interrelationships between OGSA services. A major challenge is support for bulk operations, in which grid services are applied to a collection of files. A desire is the use of re-entrant services, such that partial completion of the processing of a collection can be completed by the re-application of the service.

3.4.2 Grid Protocol Architecture

This group is developing a conceptual framework for grid services that focuses on a minimal set of protocols. Virtual Observatory environments need the simplest possible implementation to improve sustainability and robustness.

At the same time, Virtual Observatory environments want to be able to record the information content that has been created by a service. This corresponds to checkpointing the service, and reliably storing the associated information content. For large scale processing, it may then be desireable to re-instantiate the service, possibly bringing up the service on a new hardware infrastructure. The ability to checkpoint services needs to be designed into the grid services environment.

3.4.3 OGSA Naming

This group is implementing a three-level naming specification for web services (WS-naming, WS-addressing, physical address). The goal is to support interoperability between different name resolution services. Current Virtual Observatory environments rely on a two-level naming specification (logical name, physical address). A second issue is that Virtual Observatory environments use two-level naming for files and resources, and maintain control of the name spaces for users and metadata.

3.5 Applications Standards Groups

The development of higher-level services requires simplified interfaces to grid services. Requiring a larger number of levels of software to implement applications will lead to an environment that is harder to maintain.

3.5.1 Grid Remote Procedure Call

This group is defining a grid remote procedure call. Most grid services will benefit from the use of remote procedure call style invocation of remote operations. This minimizes the number of messages that need to be sent over wide-area-networks. This is particularly important for manipulating large numbers of small files.

3.5.2 Grid Information Retrieval

This group is defining an architecture for information retrieval, including document collection management, indexing, searching, and query processing. The integration of document collection management with the technologies coming from the digital library community is needed to ensure compatibility. In particular, the two communities are pursuing different standards for organizing information (METS) and for managing information (AIPs). The management of astronomy publications will need an information retrieval service.

3.5.3 Distributed Resource management Application API

This group is developing an API specification for controlling jobs submitted to Distributed Resource Management Systems. This includes submitting, terminating, and suspension. These capabilities will be needed for large scale processing.

3.5.4 Simple API for Grid Applications

This group is defining a simple API for remote job submission, file transfer from within a program. The simple API may not provide the set of bulk operations needed to manipulate large numbers of small files.

3.5.5 Grid Checkpoint Recovery

This group is defining a user-level API and associated services to permit check-pointed jobs to be recovered and continued. This will be needed for large-scale processing of data.

3.6 Management Standards Groups

The management of policies is a critical component of a shared collection. When to apply a management policy such as migration to new technology, media refresh, transformative migration, checking of Service Level Agreements for storage, pipeline processing validation, is needed as a general capability.

3.6.1 Application Contents Service

This group is defining central management for descriptions of applications. They are developing an Application Repository Interface, and an Application Archive Format for bundling components of an application. Virtual Observatory environments will need to apply this to both grid services and pipeline processes, while asserting consistency requirements.

3.6.2 Configuration Description, Deployment, and Lifecycle Management

This group is developing mechanisms to describe configuration of services, their deployment, and management of their deployment life cycle (instantiation, initiation, start, stop, restart). Note that data management requires long-running services (the catalog that manages a collection will never be turned off). However the ability to recreate a service is a critical component of the management of technology evolution. Consider the following:

- Create an information management service
- Checkpoint the information management service. This means that all information required to implement the service is archived, along with all of the information content being managed by the service.
- Archive the checkpoint.
- Retrieve the checkpoint
- Re-instantiate the service on new technology, demonstrating that no information content has been lost

The result is the migration of the long-running service onto new technology.

The Virtual Observatory VORegistry should consider interactions with this design.

3.6.3 Grid Economic Services Architecture

This group is defining protocols and service interfaces to support multiple economic models for the charging of Grid services. This includes the provision and consumption of grid services, a grid banking service for financial transactions, and a chargeable Grid service to encapsulate existing services and charge for use.

3.6.4 OGSA Resource Usage Service
This group is defining a Resource Usage Service to track resource utilization. This is a component that is required for building cost models for the shared collection.

3.6.5 Usage Record
This group is defining a common usage record for interchanging accounting and usage information. This component will be needed to support large-scale processing across multiple resources.

3.7 Security Standards Groups
The management of shared collections requires the ability to control access for privileged write operations, for updating metadata, for registering new material.

3.7.1 Open Grid Service Architecture Authorization
This group is defining specifications for interoperability and pluggability of authorization components. They plan to leverage SAML and XACML. This will be needed for integrating the Shibboleth approach to trust used in digital libraries with the Grid Security Infrastructure.

3.7.2 OGSA-P2P-Security
This group is defining how to build grids that access desktop systems to enable distributed computing. The security requirements must handle the situation where the machine user is also the administrator. One approach is to integrate data virtualization with workflow virtualization through installation of virtual Machine environments across the desktop systems. The ability to guarantee that processing has not been compromised and that the resulting data products remain under shared collection control is essential before desktop systems will be used for building Virtual Observatory environments.

3.7.3 Firewall Issues
This group is examining data transport policy enforcement. This is the enforcement of policy decisions on behalf of participating systems used by an application. Network examples include firewalls, network address translators, application level gateways, and VPN style gateways. This technology will be needed for building shared collections that may go across firewalls.

3.7.4 Trusted Computing
This group is evaluating how trusted computing initiative concepts can be applied in grids. This includes using hardware modules to enable and manage data and individual identities. The use of hardware modules may be the ultimate mechanism for establishing trust within a shared collection.

4. Summary

The requirements that must be supported by a shared collection (whether a catalog of astronomy records or an image survey archive) can be cast as constraints on both the choice of system architecture and the management policies required to maintain the shared collection. An examination of the software infrastructure that is being assembled for the Virtual Observatory community has identified opportunities to build specific Virtual Observatory services on proposed Grid service standards. Whether this will happen is strongly dependent upon the timeliness of the GGF service descriptions, the robustness of the implementations, and the scalability of the GGF services to handle millions of images and billions of records.

5. Author Information

Reagan W. Moore
San Diego Supercomputer Center

moore@sdc.edu

6. Acknowledgements

The ideas expressed here were developed under the National Virtual Observatory project with the National Science Foundation.

7. Glossary

The terms used to describe Astronomy services are listed in this section.

ADQL – Astronomy Database Query Language
ConeSearch – simple database interface using VOTable for results
FITS – the standard encoding format for astronomy images.
IRAF – Image Reduction and Analysis Facility, workflow processing for analyzing data
IVOA – International Virtual Observatory Alliance
OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting, used to implement a registry service for SkyNodes
SIAP – Simple Image Access Protocol, based on use of VOTable format.
SkyNode – standard interface to an astronomy catalog
SkyPortal – interface for composing ADQL queries
SSAP – Simple Spectrum Access Protocol, based on use of VOTable format.
STC – Space-time coordinate descriptors
UCD – Uniform Content Descriptor
VO (Virtual Observatory) – the set of semantic terms, encoding formats, and services used to support access to both existing astronomy collections and future astronomy collections. To avoid confusion with Virtual Organization, the term “VO” has been spelled out in all places where it is used, except as part of a longer term.
VOEvent – event notification service
VORegistry – registry for skynodes based on OAI-PMH
VOSpace – a standard shared collection management system for organizing data distributed across multiple storage systems.
VOStore – a standard interface for accessing data within a storage repository.
VOTable – an XML-based encoding of tabular data for transmission over a network.

8. Intellectual Property Statement

9. Full Copyright Notice

10. References

[1] IVOA – International Virtual Observatory Alliance, http://ivoa.org

[2] NVirtual Observatory – National Virtual Observatory, http:/us-nvo.org