NSF Workshop

Storage Resource Broker Data Grid Preservation Assessment

Reagan W. Moore, San Diego Supercomputer Center
Joseph JaJa, University of Maryland
Arcot Rajasekar, San Diego Supercomputer Center

San Diego, California
December 8-9, 2005

SDSC Technical Report TR-2006.3

Table of Contents

Executive Summary:

The National Science Foundation workshop on the assessment of the relevance of the Storage Resource Broker data grid technology for use in preservation environments was held at the San Diego Supercomputer Center on Dec 8-9, 2005. The workshop was attended by over 70 persons, including ten experts who participated on three panels. The experts contributed white papers to the workshop that were discussed in the panel sessions. The workshop held parallel sessions after each panel to enable further discourse on issues related to preservation. A final session summarized recommendations by the participants on the last day.

The principle findings focused on the need to share, publish, and preserve data collections that comprise the intellectual capital on which research is based. The urgency behind this desire was expressed by the participants that attended the workshop from multiple projects funded by NSF, NARA, Library of Congress, NIH, DOE, and NASA. Each community is assembling digital holdings that are used for time periods longer than the lifetime of the supporting hardware and software systems. Thus preservation is an integral component of the management requirements for large collections.

The Storage Resource Broker (SRB) is software middleware that enables the creation of shared collections that span multiple administrative domains and multiple types of storage systems. The SRB implements the principle concepts needed to support both sharing of data and long-term preservation of data: Data virtualization, Trust virtualization, Latency management, Collection management, and Federation. Presentations were given on the mapping of these information technology concepts to the preservation concepts of Authenticity, Integrity, and Infrastructure independence. The panel members included groups using the SRB to build data grids for sharing data, for distributed data management as a component of a digital library, and as a persistent archive for long-term preservation.

The central results of the workshop are that:

- The SRB technology is used as generic middleware infrastructure that manages distributed data for many types of applications, including data grids, preservation environments, digital libraries, and real-time sensor networks. Each application ports their preferred access and management mechanisms on top of the distributed data management infrastructure.
- The existing generic technical infrastructure needed to support distributed shared collections is basically complete, although additional extensions to the SRB technology will be requested as a general consequence of technology evolution.
- The two primary areas that need additional attention are: 1) support for policy management, sustainability, and governance. The development of administrative mechanisms that simplify governance will need to be tied to the particular type of data management application. 2) improved technical education and documentation of the use of data grid technology. Data grids

are sophisticated systems that currently require computer science expertise to run as production systems.

The workshop participants extensively discussed the need to establish governance and sustainability policies for the collections that are being preserved as well as the software infrastructure used to manage the collections. Assessments were also made of the functionality that was provided by the SRB, the extensions needed to the SRB infrastructure, and the development of a standards effort for data grid protocols.

The specific recommendations included:

- Data grids provide the essential mechanisms required for long-term preservation of the bits that comprise data records.
- NSF should promote development of policies for identifying which scientific data should be preserved that are generated using federal funding.
- NSF should promote sustained support for the generic software infrastructure that is used for long-term preservation.
- NSF should promote development of governance policies for preservation repositories that manage scientific data collections.
- The preservation and digital library communities should promote development of Information Technology expertise needed for managing production preservation environments.

The path forward for the SRB technology is being driven by the strong desire to be able to characterize, organize, manage, and apply policies to data collections. The current technology has focused on distributed data management, based on the consistent update of state information generated by operations on the distributed data. All governance policies were managed externally to the SRB software, and applied by invoking SRB operations at periodic intervals or after external events. The next level of sophistication in data management is to be able to express governance policies as rules that are applied by the data management system itself. The characterization of the policies as rules, the expression of the state information that the rules manipulate, and the execution of the rules are the goals of multiple NSF and NARA research initiatives. The next generation of the SRB technology will seek to supply these automated governance capabilities.

1. Introduction

The need for distributed data management is pervasive. Data grids, digital libraries, and preservation environments all rely on the management of distributed collections [44]. Scientific disciplines are assembling shared collections of research results that are distributed across international boundaries [61]. The shared collections represent the intellectual capital of the discipline. They are used to compare observational data with simulation output, analyze effectiveness of new algorithms and theories, and are used as primary resources in education [64]. Digital libraries are now incorporating the ability to manage collections that span multiple institutions [62]. Preservation environments deploy Deep Archives that store replicas at a remote site [36]. The Storage Resource Broker data grid [4] is used in multiple projects from each of these communities for the management of distributed data. The projects rely on federation of data grids [58] to enable aggregation of individual shared collections to create digital holdings that represent the intellectual capital of an entire discipline. In practice we observe convergence between the data management approaches. A common desire is to create a distributed data management system that incorporates the capabilities of data grids for sharing data, digital libraries for publishing data, and persistent archives for preserving data [37,41]. Long-term preservation is an essential component of all data management approaches.

In August 2003, the National Science Foundation and the Library of Congress co-sponsored a compelling report: **It's About Time – Research Challenges in Digital Archiving and Long-term Preservation.** The report describes the gap between the growing body of digital collections and our ability to capture, manage and preserve them: "*… from a long-term preservation perspective, there is a dark side to the rapid growth in digital information. The technologies, strategies, methodologies, and resources needed to manage digital information for the long term have not kept pace with innovations in the creation and capture of digital information [22].*" The report described the importance of working *now* to preserve the digital assets that represent the cultural history and intellectual capital of education, science, and government institutions. These assets are threatened by lack of adequate infrastructure, lack of adequate resources, and technology evolution within access mechanisms, encoding formats, and storage systems.

Increasingly, intellectual content is "born digital," as a consequence of which the digital library and digital archival communities find themselves faced with unique preservation challenges—challenges that call for comprehensive digital preservation lifecycle management processes. The NSF Cyberinfrastructure program, the Library of Congress National Digital Information Infrastructure and Preservation Program (NDIIPP), and the National Archives and Records Administration (NARA) Electronic Records Archive—all are actively concerned with the preservation of data.

The Storage Resource Broker (SRB) Data Grid preservation assessment workshop is an NSF-sponsored attempt at understanding how data grid technology can be used to enable long-term preservation. Data grids provide the data and trust virtualization mechanisms needed both to manage technology evolution and to ensure that the properties of a shared collection can be managed independently of the choice of storage or database technology [42]. An assertion by the developers of the SRB is that the distributed data management capabilities of data grids are sufficient to handle the storage requirements of preservation environments [34]. The concept behind this claim is the recognition that at the point in time when new technology is being incorporated in a preservation environment, both the old and the new technology are present. The data virtualization mechanisms provided by data grids make it possible to interact simultaneously with both versions of the technology. A preservation environment based on data grid technology inherently contains the functionality needed to manage technology evolution. The primary goal of the SRB preservation assessment workshop was to understand this claim and decide whether the SRB data management technology indeed does provide the essential capabilities for long-term data management.

The workshop was organized as a collaborative effort between the University of Maryland, under Dr. Joseph JaJa, and the San Diego Supercomputer Center, under Dr. Arcot Rajasekar. Both had received funding support from the Library of Congress and the National Science Foundation to implement Digital Archiving projects. Since both projects relied upon the use of the Storage Resource Broker for distributed data management, NSF requested an assessment of the SRB technology for use in other preservation environments.

The workshop attendees were selected from three main communities:

1. Archivists and staff from preservation projects
2. Research groups using data grids to manage distributed data
3. Digital librarians, in particular the groups who were integrating digital libraries with data grid technology.

The list of attendees is given in Appendix A. The group includes representatives from projects funded by NSF, NARA, NHPRC, Library of Congress, NASA, DOE, and DARPA. The group also includes representation from international projects based in Australia, Taiwan, Japan, the United Kingdom, and France. A total of 19 participants were from the preservation community.

The preservation projects that were represented included the National Archives and Records Administration research prototype persistent archive [47], the Library of Congress National Digital Information Infrastructure Preservation Program [48], the National Historical Publications and Records Commission Persistent Archive Testbed [56], the California Digital Library digital preservation repository [7], and the Chronopolis initiative [33]. The data grid projects that were represented included the Australian Partnership for Advanced Computation [2], the NSF Teragrid [50], the Academic Sinica data grid [8], the UK e-Science Data Grid [74], the BaBar high-energy physics data grid [3], the KEK high energy accelerator research organization BELLE data grid [27], the NIH Bio-medical Informatics Research Network [5], the NSF Real-time Observatories, Applications, and Data management Network [67], and the NSF ENZO cosmology application [12]. The digital library projects that were represented included the DSpace digital library [11], the Fedora digital library middleware [13], the UCSD Libraries image digital library [73], and the NSF Southern California Earthquake Center digital library [68].

This cross-section of projects from the data grid, digital library, and persistent archive community is similar to the collaborations that drove the development of the Storage Resource Broker technology [32]. Of the 71 persons attending the workshop, 55 were members of institutions using the SRB data grid technology. Nineteen were members of SDSC who provided technical information on the use of the SRB technology. Thirty-two participants applied the SRB in data grids, fourteen participants applied the SRB in digital libraries, and ten applied the SRB in preservation environments. Thus the workshop participants had sufficient expertise to understand the implications of using data grid technology in preservation environments as well as for management of shared collections.

The workshop was organized around three panels of experts, who provided their assessment of the important issues related to preservation. Each panel was given three topics to consider. Each panel member provided a white paper that was distributed at the workshop. After each panel session, three parallel breakout sessions were held to promote discussion of the issues raised during the panel. A summary session was then held to present the findings of the parallel sessions to the entire workshop. A final summary session was held on the second day of the workshop to collectively identify recommendations. The workshop was held at the San Diego Supercomputer Center on December 8-9, 2005.

2. Workshop General Recommendations

The general recommendations from the workshop mainly focused on the need for sustainability and governance in long-term preservation environments. This perspective was consonant with the driving requirement behind the development of data grid technology; the provision of infrastructure independence for shared collections [30]. Data grids enable the management of the properties of a shared collection independently of remote storage systems. In effect, a shared collection is isolated from the local sustainability and governance issues inherent within any single administrative domain. By replicating data across multiple administrative domains, a data grid enables the encapsulation of governance issues and the development of policies that are applied to the shared collection independently of the institutions providing the storage [35].

We can use the concept of replication across governance domains to mitigate risk of data loss against possible failure of a chosen sustainability and governance model. We associate an explicit governance model with an institution that operates a preservation facility. The institution chooses a governance model based on the driving motivation behind the construction of persistent archives. The driving motivations might be one of the following:

- Federal version of a persistent archives of federal records managed under federally mandated governance policies
- Education version of a persistent archives managed by an academic institution
- State version of a persistent archives managed by a state archive
- Commercial version of a shared collection managed by a company

Note that each of these driving motivations may require a different governance and sustainability model. However, there will be common requirements across all of the associated governance models that can be met through use of data grid technology.

We note that data grids mitigate risk of data loss through replication, risk of metadata loss through federation, and risk of technology obsolescence through infrastructure independence [36]. We can use the same mechanisms to mitigate against risk of failure of a governance model by replicating the digital holdings between multiple independent institutions which either collaborate on a chosen sustainability model, or which federate across independent sustainability models.

- Chronopolis model: Create a consortium of collaborating institutions that jointly govern a shared collection under a common sustainability model [33]. Each institution may hold a replica of the data and metadata, ensuring that risk of data loss is minimized. The governance policy of the consortium can require that at least three institutions participate in the governance of the collection. If an institution withdraws from the consortium, a new member must then be sought or an existing member must assume responsibility for the compromised collections.
- Federated Chronopolis model: Establish multiple independent digital repositories, each with a separate sustainability and governance model. Federate the digital repositories to enable each independent repository to pull data and metadata as desired from one of the other independent repositories. This ensures that a particular choice of governance policy will not lead to loss of a valued collection, since there are multiple independently governed copies of the valued collection. If one governance policy fails, another governance policy still preserves the valued collection. An effective approach is to federate repositories that have different motivations for preserving the records. An example is educational use of federally preserved records, in which the educational resource is governed independently of the federal archives.
- Interoperability model: Maintain interoperability mechanisms to ensure that a shared collection may be migrated between institutions with different governance policies. Infrastructure independence asserts that the supporting infrastructure does not introduce any dependencies that prohibit the easy migration of a collection onto newer technology. In this case, the shared collection is migrated from the governance of an original institution into the governance of a new institution.

Data grids provide the mechanisms needed to enable collection migration between independent digital repositories, between different institutions, and between different preservation environments. The fundamental concepts on which data grids are based enable the application of sophisticated governance models.

The general recommendations from the workshop propose approaches that may be followed by the National Science Foundation to ensure the long-term preservation of scientific digital holdings:

- **Data grids provide the essential mechanisms required for long-term preservation of the bits that comprise data records.**
  The concepts of infrastructure independence, data virtualization, and trust virtualization, as implemented in the SRB data grid, are essential for long-term preservation of data. Data virtualization ensures that collection properties can be managed independently of the choice of storage system. Trust virtualization ensures that authentication and authorization can be managed independently of the administrative domains across which the records are replicated. Infrastructure independence ensures that new technology can be incorporated into a preservation environment, making it possible to upgrade an existing persistent archives to avoid obsolescence from outdated technology. Scientific collections have life

times on the order of 10-20 years, implying the need to manage technology evolution that occurs on a 3-year time period. Thus scientific disciplines inherently need preservation mechanisms to ensure continued access to their digital holdings.

- **Policies are needed for identifying which scientific data generated using federal funding should be preserved.**
  NSF funds the development of simulation results and capture of observational data that comprise the intellectual capital to which future science advances will be compared. Policies are needed that describe the retention of these data collections as records of the US government. The policies need to build upon the expertise of each scientific discipline to decide which data collections are relevant and represent the state-of-the-art results. The intellectual capital in turn needs to be shared between all members of the scientific discipline to enable research advances by all academic institutions.
- **Develop sustained support for the generic software infrastructure that is used to provide infrastructure independence for long-term preservation.**
  The mechanisms currently used by the San Diego Supercomputer Center to develop the Storage Resource Broker rely on the aggregated support of about fifteen funded collaborations at any point in time (see Appendix B). Indeed, the amount of support provided by the SDSC Core NSF program in 2004 was only 4% of the total SRB development and application funding. The reliance on multiple funding sources ensures that the loss of a single project will not cause loss of support for SRB software infrastructure development, but does require a strong management team to ensure that new projects are continually being started. The management task is eased considerably through the strong support provided by the National Archives and Records Administration for research and development of the SRB data grid technology. The incorporation of data grid technology as a component of Cyberinfrastructure would ensure long-term accessibility for scientific collections.
- **Develop a consortium to assess the highest priority capabilities that should be developed and incorporated in the generic software infrastructure.**
  The multiple projects currently supporting SRB development effectively comprise such a consortium. The formal organization of the collaborating projects can lead to improved assessment of the highest priority features needed for long-term preservation. Contributors to the SRB software development are listed in Appendix C. It is worth noting that persons or institutions who contribute to the software development are typically applying the SRB technology on local projects [59]. However, there are contributors from the academic community who contribute to data grid development as part of fundamental research on data management. A consortium for prioritizing data grid development should include not only the funding sources, projects using the technology, but also academia interested in data management research.
- **Develop governance policies for preservation repositories that employ data grid technology to manage infrastructure independence.**
  The governance of the institution managing a preservation repository is as important as the governance of the project that is developing preservation

technology. The Chronopolis project is exploring governance policies for a preservation environment that spans multiple institutions (SDSC, the UCSD Libraries, NCAR, and the University of Maryland). Governance policies for long-term preservation are also being explored in the NARA Pledge project and the NSF National Science Digital Library project [49]. The development of governance policies is thus being pursued by a much broader community comprised of multiple independent research efforts.

- **Promote development of Information Technology expertise within the preservation and digital library communities.**
  The expertise needed to understand, validate, update, and apply the software used in preservation environments requires a strong background in computer science. This expertise has traditionally resided in the companies providing the software, and access to the expertise is ensured through maintenance contracts. With the emergence of open-source software, the expertise is more frequently provided within the institution using the software. The preservation and digital library communities now need training courses in maintenance of software infrastructure as they seek to apply open-source software. The SRB is distributed as source to academic institutions, non-profit organizations, and US federal agencies. The impact of a source distribution is that the receiving institution must build the software, integrate the software into their data management infrastructure, manage upgrades to the software, and identify bug fixes. The SRB group at SDSC manages a SRB-chat e-mail discussion list to promote communication between groups using the software. The hope is that the SRB user community can develop sufficient expertise to manage their own environments and develop shared solutions to problems that are identified.
- **Develop standard interfaces to ensure interoperability between preservation environments.**
  The goal of infrastructure independence, as supported within the SRB, ensures that no dependencies on choice of software and hardware implementation will prohibit the migration of records onto new or alternate preservation environments. The development of standard interface mechanisms simplifies this process. Note that data virtualization supports the ability to port new standard interfaces onto the SRB environment. The separation of client access protocols from storage access protocols ensures that new interfaces can be ported onto existing collections, enabling migration of records into new preservation environments. Finally, even if a specific standard interface is defined today, a new version will be developed in the future. All components of a preservation environment will change over time, including the standard access interfaces. Thus the ability to port new access mechanisms onto an existing preservation environment is essential for long-term preservation.

3. Pervasive Data

The pervasive need for distributed data management is well demonstrated by the list of workshop attendees (Appendix A), the list of current projects that are supporting the development of the Storage Resource Broker data grid technology

(Appendix B), the list of collaborators who have contributed to the development of the Storage Resource Broker (Appendix C) and the list of sites that have downloaded the Storage Resource Broker technology (Appendix D). In total they represent over 170 projects that are either exploring the management of distributed data collections, or that are now running production data grids. The projects include collaborations that are local to a single institution (such as the UCSD digital image archive), or that span multiple institutions within a single nation (such as the NIH BIRN project), or that span institutions in multiple countries (such as the BaBar high energy physics project).

The collaborations range from principle investigator-driven research projects to consortia that are assembling the digital holdings of a discipline (National Virtual Observatory [51]). This latter project is an example of the desire to manage and preserve the intellectual capital on which future astronomy research will be based. A concept expressed in the workshop by Carl Lagoze is that a preservation repository enables the preservation of knowledge. The goal is to facilitate future research by enabling comparisons with the intellectual capital derived in the past. The ability to access, manipulate, and apply information and knowledge extracted from a preservation repository is the real motivation for preserving data.

4. Assessment Criteria

The assessment criteria for appraising the capabilities of the Storage Resource Broker are based on practical examples of persistent archives. The simplest example is to consider the functionality that is needed to implement a "Deep Archive". This is a preservation repository that implements:

- Logical air gap between the preservation environment and the external world. There is no direct access from any external site to the data grid managing the deep archive. Instead, all data movement is staged through an intermediate "staging data grid".
- Non-disclosure of all administrative names. The identity of the storage resources and the names of the archivists who manage the deep archive are not visible from the external world.
- Ability to pull records into the deep archive through the "staging data grid" and push records through the "staging data grid" back to the external world. The records within the deep archive can be accessioned and disseminated under archivist control.

A traditional approach for constructing a deep archive is to require that all accessioning and dissemination be done using removable media. The physical act of transporting the media then corresponds to the "air gap" that minimizes risk of attack. Figure 1 show a deep archive that is constructed through controlled federation of three data grids. Each data grid or zone manages an independent set of name spaces for identifying resources, users, files, and metadata. The SRB data grid allows controls to be enacted on the sharing of name spaces between zones. Thus it is possible for a name corresponding to an archivist (U2) who operates the staging data grid to be registered into a remote zone without sharing any other name spaces. This person can issue commands from the staging data grid to pull records onto the staging data grid. A protocol that supports server-initiated data transfer is used to load the data onto the staging data grid through a single communication port. No person from the external world has access to the staging data grid, because they have no registered identity in the staging data grid.

The archivist who manages the deep archive (U3) similarly can register her name into the staging data grid. The archivist can then pull data from the staging data grid into the deep archive using a private virtual network. This is a two-stage process. The records are first loaded onto the staging data grid, and then loaded into the deep archive. The identity of the deep archive archivist (U3) is not seen by the external world. Nor are the names of the storage resources or locations of the metadata catalogs within the deep archive.

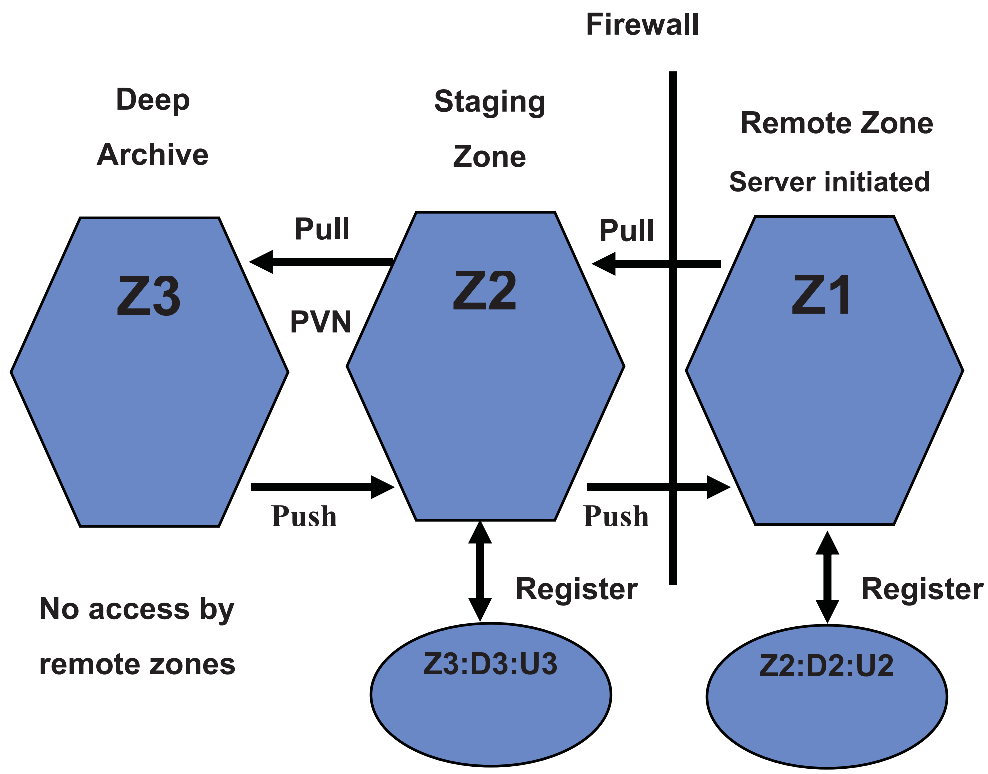

Figure 1. Deep Archive assembled through federation of multiple data grids

The capabilities required by such an environment constitute the minimal set of capabilities that a data grid should provide for supporting preservation environments. These capabilities include [38,40]:

- Logical name space for identifying users, independently of the storage repositories
- Logical name space for identifying files, independently of the storage repositories
- Logical name space for managing metadata, independently of the database (schema indirection)
- Logical name space for identifying storage resources

- Management of access controls as relationships between the four logical name spaces, ensuring that access controls remain invariant as data is migrated to new storage systems
- Standard operations for managing wide area network communications (parallel I/O, packing of small files in containers, remote processes for filtering data or extracting metadata, protocol support for firewalls, third-party data movement, bulk operations such as registration of files, metadata insertion, loading of files, metadata export)
- Standard operations for interacting with storage systems (file read and write, metadata creation and update, automated update of state information on completion of standard operations)
- Standard administrative attributes for properties of records (replicas, versions, annotations, audit trails, descriptive metadata, aggregation in containers)
- Trust virtualization to support authentication and authorization independently of the storage system.
- Standard operations for supporting client interfaces (standard interoperability mechanisms). Most clients now build upon either a C library interface, a Unix shell command interface, or a Java class library.
- Shibboleth style authentication between data grids [69]. An individual is always authenticated by her "home data grid".
- Support for name space federation, the controlled sharing of a selected portion of a name space with another data grid.

5. SRB Preservation Capabilities

The preservation concepts of authenticity, integrity and infrastructure independence define the capabilities needed for long-term preservation [45]. Authenticity consists of assertions made at the time of accessioning, typically by the creator of the record. Depending upon the preservation model, the authenticity assertions may be handled as static descriptive metadata. Integrity consists of assertions made by the archivist that the records have not been corrupted, the chain of custody has been tracked while the records have been under archives control, access controls have been maintained, sufficient replicas exist to minimize risk of data loss, modern access methods can be used to access the records, etc. Infrastructure independence is an assertion that the preservation environment has no dependencies upon any choice of hardware or software system, protocol, or network that prohibits or impedes the migration of the records to alternate technology without loss of authenticity or integrity.

The preservation requirements are effectively handled by the principle concepts underlying data grid technology [34,63]: Trust virtualization; Data virtualization; Latency management; Collection management; and Federation. Data grids insulate the governance of records from the governance properties of a particular administrative domain associated with a storage repository. Data grids can impose governance policies that span all of the administrative domains where records are stored. Data grids implement a shared collection and manage the properties of the shared collection independently of the storage repositories.

The data grid concepts map to the following capabilities:

- Trust virtualization
  - ⇒ Management of authentication, authorization, and audit trails
- Data virtualization
  - ⇒ Management of logical name spaces for resources, users, records, metadata
  - ⇒ Standard operations for interacting with storage systems
  - ⇒ Standard operations for supporting clients
- Latency management
  - ⇒ Mechanisms required for scalable data and metadata transport
- Collection management
  - ⇒ Mechanisms for managing a catalog that resides in a database
- Federation
  - ⇒ Management of controlled sharing of logical names spaces and forwarding of operations between autonomous data grids

A subset of the capabilities supported by the Storage Resource Broker is listed in Appendix E. Additional capabilities related to data grid administration and project specific access clients such as OpenDAP [55], OAI-PMH [52], DSpace [11], Fedora [13], Perl, Python, Windows, etc. are not listed. The particular capabilities provided by the SRB have been driven by the projects listed in Appendices B and D. We note that capabilities such as latency management are equally important in a preservation environment as in a data grid, when large numbers of records are replicated. A second observation is that many of the transport mechanisms have been driven by the need to interact with network devices such as firewalls, load levelers, and virtual private networks [39]. A third observation is that trust and data virtualization are essential for creating a preservation environment that provides infrastructure independence. A standard example is the creation of access controls that do not change as a record is moved to alternate storage systems. In data grids, the access controls are a constraint between two invariant logical name spaces for files and users. The access controls therefore are location independent.

For each set of capabilities listed in Appendix E, we define the logical name spaces that are used, provide a set of associated operations, and list the state information managed by the data grid to track the status or result of the operation. The state information collectively corresponds to the integrity information needed in a preservation environment. Authenticity information can be stored as descriptive metadata associated with each record. Infrastructure independence is implemented through the logical name spaces and standard operations supported across a wide variety of storage systems and client methods.

6. Workshop Appraisal Results

The Storage Resource Broker provides the essential mechanisms needed to manage distributed data. The capabilities provided by data and trust virtualization are recognized as essential for managing technology evolution, and for ensuring a common governance model across multiple administrative domains. However, multiple workshop participants requested the specification of the minimal set of essential features. This represented a desire to minimize the knowledge required to use data grid technology effectively within preservation environments, and a hope that the set of features needed for preservation would be smaller than the set of features needed to support internationally shared collections.

Based on experience with the eight preservation collaborations listed in Appendix B, the actual impetus is towards more sophisticated environments. The set of features implemented in the SRB have been driven by concerns about scalability and how to manage environments that will hold billions of records. In addition, the sophistication of the access environments is now increasing dramatically through the integration of digital library interfaces on top of the SRB data preservation environment. Both digital library services and workflow environments are being integrated on top of the SRB data grid to support preservation processes, the automated capturing of descriptive metadata, and the automated validation of the integrity of the digital records. Current research activities are focused on the explicit description, management, and application of governance, consistency, and access policies. Collaborations include the NARA-funded project on "Policy Enforcement in Data Grid Environments" and the NSF-funded project on "Constraint-based Knowledge Systems". The expectation is that a new generation of data management technology will emerge that allows the governance policies to be explicitly stated, and dynamically changed as desired without the need to re-write software code. The end goal will be a simpler environment to manage and administer, but at the cost of more sophisticated data management software technology.

The desire to simplify technology for use within preservation environments is perhaps best addressed through the creation of user manuals that identify the most appropriate set of operations for accessioning records, managing descriptive metadata, managing integrity checks, and managing migration of records to new technology. Robert Horton is proposing such a project. This also addresses the desire to minimize the information technology expertise required to manage preservation environments. A preservation manual that identifies appropriate responses to administrative tasks can make the technology easier to understand. The actual maintenance of the software itself will require substantial information technology expertise. As noted in Appendix C, the number of persons who have contributed substantially to the SRB development is twice as large as the staff at SDSC. The contributors typically have a strong background in computer science, and have been able to port new access mechanisms, port new database interfaces, and even collaborate on improved security mechanisms. The collaborations show that it is possible for an institution other than SDSC to both apply and maintain the SRB data grid technology. The level of expertise that is required is typically beyond the level of computer science expertise available within a small archives community.

In contrast to the desire for simpler technology, the desire was also expressed for additional features to improve scalability and interactions with other grid technology. Many of these activities are ongoing through current funded SRB development collaborations. Examples include the porting of the GridFTP access mechanism [16,18] on top of the SRB data grid, the incorporation of the latest version of the Grid Security Infrastructure for authentication [19], and the integration with workflow management systems such as Kepler [28]. Future activities include integration with the Global Grid Forum Storage Resource Manager interface developed at the Lawrence Livermore Berkeley Laboratory [72], and the Cheshire digital library services developed at the University of California, Berkeley and the University at Liverpool [75].

The most heavily expressed desire at the workshop was the assurance of sustainability of the SRB data grid software for bug fixes, ports to new environments, and management of new technology. The long-term availability of technology is essential for its use as core infrastructure in preservation systems. The most effective way to ensure long-term viability of any technology is to have access to multiple independent implementations. For the Storage Resource Broker data grid, the goal is to have at least three sources of the technology:

- An academic version of the SRB is distributed by the San Diego Supercomputer Center. This version aggressively incorporates the features developed in the multiple collaborations listed in Appendix B and by the collaborators listed in Appendix C. The software is distributed as source to academic institutions and US federal agencies. The 70 institutions listed in Appendix D all have copies of the SRB source, and represent a small fraction of the 170 institutions that have downloaded the software in 2004 and 2005.
- A commercial version of the SRB is sold by Nirvana Storage. This version is targeted towards enterprise-level management of distributed data, but still provides both data and trust virtualization.
- The solicitation for the NARA Electronic Records Archive was based on the concepts used for the NARA Research Prototype Persistent Archive. The NARA ERA will provide an independent implementation of these preservation concepts.
- The University of Maryland implemented a data virtualization environment called LPE (Lightweight Preservation Environment) based on Globus Grid technology [14]. The system used the OGSI [54] web services architecture based on WSRF [77]. The system used the Replica Location Service [66] and the RFT Reliable File Transport protocol [65].
- The Preservation Environment Working Group of the Global Grid Forum is promoting deployment of at least three preservation environments that are based on data grid technology. The goal is to demonstrate true infrastructure independence by migrating collections between multiple independent persistent archives implementations.

In practice, the desire for access to source code has been outweighed by the desire for access to consulting support for problems. While all academic sites that use the SRB have copies of the source code, all sites do not have the expertise needed to resolve local problems. The provision of additional expertise for managing the SRB data grid is being addressed through the creation of centers of excellence in the United Kingdom, Australia, the Netherlands, and Taiwan. Each of these groups is establishing a nation-wide data grid for the management of scientific collections, publication of scientific data, and preservation of digital holdings. The hope is that the users of the SRB data grid technology will be able to rely upon centers of national-level expertise. This provides an effective way to build sustainable infrastructure support for international use of the SRB data grid technology.

7. Workshop Specific Recommendations

As part of the workshop, explicit recommendations were made for the current version of the Storage Resource Broker technology (version 3.4 released on October 31, 2005).

- Community participation in SRB development and extended governance to include additional institutions

  A strong desire was expressed for extending the specification of new features beyond the projects that provide funding for software development. Fortunately, the projects listed in Appendix B comprise a small fraction of the groups providing input on distributed data management requirements. Extensive feedback has been received and incorporated from the UK e-Science Data Grid, the NSF Digital Library Initiative Phase I sites, and the InterPARES project (International Research on Permanent Authentic Records in Electronic Systems) [26]. Continued feedback is essential to ensure that the SRB technology addresses the production requirements of the user communities.

  Input from multiple US federal agencies has been acquired through the funded collaborations, and input from standards agencies such as the Global Grid Forum has been sought through multiple working groups. The development of a consortium to filter requests should be done in collaboration with a standards group. Possibilities include the Global Grid Forum (data grid technology), the InterPARES project (preservation technology), and the Chronopolis project (preservation facility). Indeed, the development of an advisory committee is an explicit component of the Chronopolis project for preserving scientific collections. A final possible source of guidance is through the NSF Cyberinfrastructure program, with a focus on the preservation of scientific collections.
- Sustainable version

  The current release of the Storage Resource Broker (version 3.4) is close to providing a final feature set [71]. The rate of feature requests is slowing, and a concerted effort was made in the release of version 3.4 to address all bug reports. Hopefully release version 3.5 will contain the final set of preservation support features. A sustainable version of the software would then focus on porting to new technologies, incorporation of bug fixes as needed, and development of more extensive documentation. The rate of development of new features could be slowed substantially, leading to a stable data grid protocol for managing interactions between data grid servers.

  Funding is still needed to maintain the sustainable version of the software. The current approach of seeking collaborations with large-scale projects in support of application of the technology is still viable. The number of communities expressing interest in managing their own data grids continues to increase. These communities could make feature requests that would require future software releases. Thus the funding model for SRB maintenance will impact the evolution of the software. An ameliorating influence has been the substantial support provided by NARA for development and application of the data grid technology in preservation environments. This has provided a strong focus on preservation capabilities as an essential design component of new features. The continued support by the preservation community of the SRB technology will ensure that the software remains relevant for use in preservation environments.

- New features

  Despite the extensive set of distributed data management features provided by the SRB, current users of the technology have requested additional features related to scaling and interoperability. These features include:

  - Modular composition enabling access to local authentication systems. The goal is to be able to use existing user distinguished name spaces for authentication of users.
  - Interfaces to alternate technologies. This includes support for additional types of storage systems, additional database products for storing metadata, and new workflow management systems. The goal is to be able to incorporate subsystems within the SRB with minimal effort.
  - Integration with Grid technology. This includes developing interfaces to grid services such as the VOM – Virtual Organization Management system, and the OGSA-DAI database interface [16]. The expectation is that the grid services will become the standards used within the preservation community to support interoperability between preservation environments.
  - Standard SRB server interface. This is the publication of the protocol used to support peer-to-peer interaction between SRB servers. The publication of this protocol would improve interoperability between independent implementations of the technology.

    Conversely, one of the current usage models of the SRB federation technology emphasizes the ability to manage independent data grids that do not interact. All data migration or metadata replication is managed externally to the SRB data grid, with the explicit extraction of data from one data grid and the explicit import of the data into a second data grid under administrator control. This is very similar to the design of the Deep Archive, such that interactions with a data grid that is being used for preservation are isolated from the external world.
  - Parallel I/O on bulk loads. The ability to both pack small files before transport, support parallel registration of the small files into the metadata catalog, and use parallel I/O streams to move the packed files appears to be the ultimate data transport request. This would allow optimization of data movement across a wide variety of record sizes, and minimize the number of decisions required by the archivist for tuning performance.
  - Support for improved error handling. The current system returns all errors reported by any of the accessed systems. Tracking which error number corresponds to a particular system component is difficult.

  - Support for administrative and user tools to simplify use of the system. Examples are monitoring tools, automated check-sum verification, recovery of lost or damaged MCAT catalogs, zone and file replication management, and creation of management reports on number of collections, number of files, and storage capacity utilized. The current set of tools have typically been developed for a specific project, and need to be tuned for use in other projects. As the user community expands, a reasonable goal is to seek generic versions of each of the management tools.
- Request for list of alternate technologies. The support provided by the SRB for data and trust virtualization is unique. However there are multiple partial implementations of data virtualization. The implementations provide a subset of the features incorporated in the SRB data grid. Examples include:
  - Sybase/AVAKI data grid. The AVAKI technology was based on the Legion permanent object environment, but has been extended to support shared collections.
  - Oracle database. This provides support for Binary Large Objects and descriptive metadata.
  - IBM High Performance Storage System. This provides support for distributed storage servers and parallel I/O.
  - Veritas. This provides support for backup of files.

8. Directives to the Workshop Panels

Preservation can be viewed as the process of extracting a digital record from its creation environment and then the importation of the digital record into a preservation environment, while preserving authenticity and integrity [35]. Authenticity is the preservation of assertions made by the creator of the record, and includes provenance information. Authenticity information is typically static. Integrity is the set of assertions by the archivist about the state of preservation, and is typically dynamic information that changes after each preservation process. A preservation environment can be viewed as the infrastructure that protects the digital records from changes that occur in the external world. Examples of such changes are technology evolution, emergence of new standards, and even evolution of preservation policies.

A preservation environment is composed of the set of virtualization layers that allow implementations based on both current and future software and hardware systems. The virtualization layers include:

- Object virtualization, the ability to characterize the structure and information content of a digital record independently of the creating application. Examples include persistent objects [43], the Data Format Description Language for scientific data (DFDL) [10], and the Multivalent Browser system [46] for office products.
- Data virtualization, the ability to manage properties associated with the digital records independently of the choice of storage technology. Examples include data grids that manage persistent name spaces for users, files, and metadata while providing a set of standard operations for interacting with storage systems [34].

- Knowledge virtualization, the ability to characterize, organize, and manage both relationships between records and preservation environment management policies, independently of the implementation choice. Examples include Fedora middleware and Cheshire.
- Trust virtualization, the ability to manage authentication and authorization for the preservation environment independently of the administrative domains where the records reside.
- Workflow virtualization, the ability to manage the application of preservation processes and services independently of the choice of execution platform. Examples include web services, Kepler workflow environment, Condor [9], and Grid technologies [16].

8.1 Panel 1: Questions on Features of Digital Preservation Architecture

1) Support for authenticity: The quality of being genuine, not a counterfeit, and free from tampering, and is typically inferred from internal and external evidence, including its physical characteristics, structure, content, and context. From "A Glossary of Archival and Records Terminology" by Richard Pearce-Moses, Chicago, 2005, The Society of American Archivists. For data grids, this is the assurance that the material in the digital archive is correctly linked to descriptions of its origin;

- Assess authenticity assurance. A preservation environment maintains the links between the authenticity metadata that describes the provenance of electronic records and the preserved electronic records. Example capabilities include:
  - Do the namespaces used to identify archivists (for auditing preservation processes), to identify files (for storage of the electronic records), to identify storage resources (for tracking chain of custody), and to manage authenticity metadata remain invariant (persistent) under data management operations?
  - Does the system support Archival Information Packages for recovery from loss of preservation metadata [53]?
  - Does the system provide the ability to automate the extraction of the required preservation metadata for each electronic record?

2) Support for integrity: The quality of being whole and unaltered through loss, tampering, or corruption. From "A Glossary of Archival and Records Terminology" by Richard Pearce-Moses, Chicago, 2005, The Society of American Archivists. For data grids, this is the assurance that the material in the archive is uncorrupted, that the chain of custody can be tracked, and that the information content remains unchanged;

- Assess integrity assurance. A preservation environment provides mechanisms to validate the chain of custody, control the access, and validate data checksums. Example capabilities include:
  - Does the system provide audit trails to track both the storage locations and the archivists who manipulate the records?
  - Does the system support checksums for validating data integrity?

   - Does the system provide access controls on each record, on each metadata attribute, on each storage resource? Are the access controls invariant under data grid operations such as migration?

3) Support for infrastructure independence: The assurance that the digital archives has not imposed any proprietary standards that prevent migration of the contents of the digital archives to another choice of technology.

- Assess infrastructure independence. A preservation environment is viable if the archives (the preserved material and the name spaces used to manage the preserved material) can be migrated into another preservation system or onto other choices of storage and information management technology without loss of authenticity or integrity. Example capabilities include:
   - Does the system support migration of the archives to another preservation environment?
   - Does the system provide the abstractions needed to support technology evolution (storage repository, metadata repository, access protocols, preservation services, preservation state information, encoding formats)?
   - Which preservation environments interoperate (DSpace, Fedora, LOCKSS [29], Greenstone [17], Adore [1], …)?
   - Which access and indexing technologies interoperate (Cheshire, Multivalent Browser, HDF5 [21], OpenDAP [55], …)?

8.1.1 Statement for Panel on Features of Digital Preservation Architecture

Dr. Margaret Hedstrom, University of Michigan

I was asked to assess three aspects of the SRB and data grids digital preservation environment: 1) Authenticity assurance; 2) Integrity Assurance; and 3) Infrastructure independence.

1) Authenticity Assurance

Authenticity is defined as: *The quality of being genuine, not a counterfeit, and free from tampering, and is typically inferred from internal and external evidence, including its physical characteristics, structure, content, and context.* It is important to point out that there is considerable debate among archivists about the meaning of authenticity of digital information. Also, different communities of producers and consumers of persistent archive contents have different definitions and requirement for authenticity.

My primary is whether a stringent requirement for authenticity that is deeply embedded in the architecture is necessary, feasible, and affordable. I have three main questions.

a) Level of granularity. Is it necessary to have detailed preservation metadata associated with each record, object, entity, etc.? This has not been done in the past for most archival collections. The authenticity of individual records is inferred from their presence in a collection whose provenance is known and that has been maintained in a relatively secure environment. The technological controls of audit trails, checksums, restricted write/delete privileges, and version control implemented for files, collections, Archival Information Packages (AIPs) or other aggregations, rather than for every individual records, might be sufficient and would provide much more rigorous means for detecting unauthorized changes or changes the deviate from the archive's policies than current methods of paper records.

b) Preconditions for authenticity assurance. The system is based on assumptions about the presence and quality of creator-supplied metadata. While some data collection/creation systems can be instrumented to automatically capture metadata needed for authenticity, most common personal and office applications lack these capabilities, or (when they exist) they are not implemented. Vendors are enhancing the capabilities of systems to support more automatic collection of such metadata, but the benefits of these capabilities will not be apparent for some time even if they are widely accepted and used. For archival records, at least, even if these systems were deployed tomorrow, the benefits will not be evident for 10, 20, 30 years or even longer when records created tomorrow start to flow into preservation environments. This raises the question of which aspect of "authenticity" the preservation environment ensure. If the producer did not capture metadata on provenance, chain of custody, permissions, authorized users, and previous transformations and migrations, then the archival storage system cannot assure authenticity of the records. It can only demonstrate how, why and who made any changes to the data after it was ingested into the preservation environment.

c) Process controls and human intervention. The architecture provides for restricted permissions to authorized users, and it tracks their actions. That feature is important for authenticity assurance. However, the current workflow requires considerable intervention by "trusted custodians" and "archivists" to validate metadata, document transformations on the data, audit preservation procedures, etc. Many of these interventions may be oversight of technical processes and not activities that involve professional judgment. Automating these processes would reduce the opportunities for operator error or malicious behavior on the part of authorized users and reduce the costs of operating and maintaining the system because it would reduce the number of opportunities for humans to interact with the system.

I do not believe that the system can provide the ability to automatically extract preservation metadata for each electronic record in cases where the producer has not designed the original data collection or recordkeeping system to provide structured, accurate, and adequate metadata. Absent that capability, I do not think the system can scale to the anticipated amount of data to be preserved.

2) Integrity Assurance
Integrity is defines as *"the quality of being whole and unaltered through loss, tampering, or corruption."* The proposed mechanisms (audit trails, checksums and digital signatures, annotations, permissions and controlled access, and version control) in the proper combination are reasonable mechanisms for integrity assurance. It is worth noting that the definitions of authenticity and integrity are similar with the

quality of being genuine as the distinguishing concept for authenticity and the quality of being whole as the distinguishing concept for integrity. Is it necessary to have entirely different mechanisms to manage "genuiness" as opposed to "wholeness" (whatever those distinctions actually mean in practice?)

I question whether integrity assurance needs to be managed at the level of each electronic record (especially because the definition of integrity is based on "being whole"). Typically, chain of custody is tracked for entire collections (or portions of them) not on the level of individual records or documents. Transformations of the data that may be necessary for forward migration typically are conducted on files, data sets, databases, or data types with similar properties.

3) Infrastructure Independence
It is difficult to assess whether the system supports migration to a new preservation environment because there are so many unknowns about which technologies, standards, and services will constitute the new preservation system. It appears that migration is viewed as a one-time event (repeated every few years to take advantage of new technologies) rather than as an ongoing process. At the point of migration, the old preservation environment, data, and metadata are encapsulated and ported to the new preservation environment. Both systems run in parallel until the new environment is validated and then the old system is shut down. For persistent archives with heterogeneous content, a more likely scenario is one where different components of the system are upgraded or replaced independent of the others and where migration is a continuous process. Given the transfer rates for very large files and the requirements for auditing and verifying the accuracy of migrations, migration might be easier to implement as an ongoing or staged process rather than a single event.

I have three other concerns that do not fall into one of the categories above.

4) End user access. The access mechanisms at the fifth level of the architecture are unlikely to satisfy end user requirements because they are designed for archivists. The requirements for archivists and the requirements for the end user are not the same. Many archival collections are freely available to the public without access restrictions and some archives prefer to make their holdings available to anyone without registration. This will require anonymous logins for users with read/copy only permissions. The dissemination information packages seem to be viewed as orders for files (e.g. file discovery) rather than as responses to end user queries. The entire end user interaction is underspecified. It will be important to decide whether end users will query metadata catalogs, data collections, and/or individual records. Public end users prefer search mechanisms that resemble current and widely available web search engines that can search across the entire archive, rather than only metadata catalogs or specific collections.

5) Open- versus closed archives. There is an important distinction between archives (and collections) that are closed rather than open. Closed archives and collections are

those where data collection, collection development, or the recordkeeping process has ceased. Examples of closed collections include the data from an instrument that has been decommissioned; a collection of images that has been digitized and cataloged in its entirety, the records of a government commission that has completed its work and issued its report (e.g. the 911 Commission). Many collections are open, in that they continue to be created and/or revised. Data may come into archival custody as a steady stream as (or shortly after) it is created; or it may come in batches as older records are transferred to archival custody under some schedule. Given that many archives are open (in that they are regularly acquiring new collections or accretions to existing collections), I am concerned about keeping the various replications in synch with each other and about whether the principle of two or three replications is sufficient especially if their contents are not identical.

6) Why authenticity and integrity, not trust? Trust is a broader concept that relies on a variety of social, institutional, and technical mechanisms to increase end users' confidence in the quality, accuracy, authenticity, and integrity of the data in an archives. Recent work on the attributes of a trusted repository and on certification requirements include organizational, financial, technical, and human components. Trust is also built when organizations or services demonstrate competence and develop a track record of high performance, acknowledgement of errors, and implementation of process improvements. If end users trust the preservation environment, then the collections and records in that environment also inherit that trust. I would not contend that errors, malicious activity, and system failures will not occur, and when they are detected they should be acknowledged, damage should be repaired to the extent possible, and policies, processes and technologies should be revised appropriately. It seems, however, that by vesting trust in absolutes of authenticity and integrity (as they are currently defined) given enough time, the environment is almost absolutely certain to fail. Failure to meet an absolute objective will then work against trust in the environment.

8.1.2 Comments on Features of Digital Preservation Architecture

Martha Anderson, Office of Strategic Initiatives, The Library of Congress

In the past 10 years, work on digital preservation systems and digital preservation requirements, suggest that the nature of preservation is organic rather than mechanical. Approaches and processes evolve over time rather than are crafted and established for the long term. The model of diversity may insure the greatest survival rather than the model of normalization. Through the NDIIPP Preservation Partnerships and other work sponsored by the program, there is a strong suggestion that the most significant factor driving approaches to preservation arises from what the OAIS Reference Model describes as the Designated Community. Many of these communities use a single data type such as spatial data, or textual data that then drives the standards, practices, tools and processes for creation, access, and care of the data over time. Each community may have different requirements for such features as authenticity, integrity and infrastructure based upon the scale, adoption of tools and data formats, and access for end-users and caretakers.

1) **Support for authenticity:** The importance of authenticity varies from community to community. Cultural Heritage institutions may have a lower standard for authenticity certification for such content as harvested web sites than a Business Records Management organization that must comply with federal regulations.
    a. Assessing authenticity assurance can be a very costly burden on a preservation system beyond the assurance of integrity. Metadata about the creation and custody chain of digital content is not very well supported by current creation and processing tools. The relationship between producers and preservation entities may be very loose. With the NDIIPP Archives Ingest and Handling Test, the project made the assumption that there was no direct link back to the creator because the content was donated by the public via a web form in the days following September 11, 2001. In the case of this kind of digital folk archive, authenticity before the deposit was not possible or feasible to certify. During the project, the most useful metadata extracted was the technical attributes of the objects. Therefore, only the integrity of the data transferred could be assured and a deeper understanding of the technical characteristics of the data (not its authenticity) brought forward to the receiving system.
2) **Support for integrity:** This is the very basic feature many preservation systems avow to support. At the LC, we have come to refer to this as the bit-preservation level. The tools and methods for assuring data integrity, especially at the file level, are better understood than many other aspects of preservation. That being said, it is not a simple task at large scale to run checksums and manage logs to provide audit trails. Virus scans and checksums for gigabytes of data can consume days of machine time for some systems. It is daunting to think of running these processes at petabyte scale in the current serial mode used by many systems. There is some work being done by one of the DIGARCH NSF projects to probe sampling techniques for hashing that may be encouraging but more investigation into practical approaches for asserting integrity is needed.
3) **Support for infrastructure independence:** Infrastructure independence has been the driver for promotion and adoption of XML and open source tools. However even more important than infrastructure independence is the requirement for an architecture to support the fluid flow of data as it must be migrated to new media, exchanged or transferred to another caretaking entity, or transformed to accommodate format or system obsolescence. The reality is that some proprietary systems used for data creation become the first preservation system. In one of the NDIIPP projects at the University of Maryland, the data that is the focus of preservation is instantiated within the document management system of a law firm. Risks of low adoption may be just as great as risks of proprietary technologies. A widely adopted proprietary format or system may be better understood and offer more preservation advantages in the long run than an open system with low adoption. It is not realistic to believe that the entire digital content producer community will adopt open source systems or even move quickly to adopt data and metadata standards. In some designated communities such as the GIS community, the most widely adopted and supported data formats and tools are proprietary. Vast quantities of business correspondence and legal

documents are created in proprietary formats and on proprietary systems. PDF is an example of a widely adopted format that has gained some support for an archival form in recent years.

4) **Other:**
   a. Scalability is a feature that is likely to override other desired features as the volume of digital content grows. Many current tools and approaches are based upon labor-intensive metadata collection, or normalization rationalizations. As preservation entities collect data and try to manage it, they will be challenged to maintain the current view of approaches. A framework for diversity may be a better approach to preservation architecture. Most work is being done for Designated Communities who focus on the most critical preservation tasks for their data using domain expertise and the most widely adopted tools and systems. In the end it is the ability to transfer this deep understanding forward that preserves the data.

8.1.3 Minnesota Historical Society

Bob Horton, Minnesota Historical Society

I am the state archivist at the Minnesota Historical Society. The MHS is one of the premier cultural heritage institutions in the Midwest and the largest state historical society in the country. We have over 300 FT employees, an annual budget of over $40 million and a record of technological innovation, especially (and I should be straightforward – pretty much exclusively) in our library and archival functions.

The MHS and principally the state archives has worked with the SDSC, particularly Richard Marciano and Reagan Moore, since 1999, and the archivists workbench project. We've worked more closely with the SRB and grid technology in the past three years, as one of the partners in the Persistent Archives Testbed project and as the principal partner in a project to preserve the records of the e-legislature, as we call it. At this point, we're still testing the technology, although the e-legislature project is certainly a more intensive and ambitious test. As I usually explain this, though, we are testing a proven product and concept – the SRB and grid technology – in a new environment. That environment is specifically the state archives, working with government records.

Certainly, my impression is that the SRB works effectively within the environment in which it was designed and first implemented, here at the SDSC, and in other data intensive computing environments. I have seen this most compellingly demonstrated in a slide Richard uses in presentations, which lists volumes of material stored in SRB collections – almost 54,000 GBs for the National Virtual Observatory, 131,000 in the S California Earthquake Center etc. etc. Then 100 GBs in the PAT prototype.

Note the comparative volume. It's a different world. And obviously the technology itself plays a role in the different ways and different rates these different communities have adopted and applied it. But we have tended to articulate those differences in terms of the professional and intellectual aspects of preservation. I want to add something else to that conversation. Again, I'm going to try to articulate the difference in legal, cultural and administrative terms, of a practical implementation, of operation.

At the MHS, at the state archives especially, the staff is simply overstretched. We have limited resources and we have inherited multiple functions. The result is competing and even conflicting priorities. Funding is down. State government is reducing overhead and administrative costs, focusing on core services, with an increasingly ruthless disdain for anything else that adds to the budget. And, last, we don't really even have a state archives anymore – the Society re-organized in December, aggregating all the departments that "collected" into a single collections department, with government records part of a whole that now includes everything from books to wedding dresses to TV news footage to pottery shards.

I don't know where that puts us in the evolutionary scale, whether we're among the first to crawl out of the ocean and try to make it on dry land or whether we're the duck billed platypus of archives. I suspect it's the former – that our experiences will be more, not less common, among archives and cultural institutions. The new National Archives and Library of Canada is one example.

That makes a complex organization like my own even more complex. Our internal applications developed as stovepipe functions (eg CMS, EAD, Aleph). We create and to a lesser extent collect many small data sets, in different formats. We've inherited many variant metadata standards and processes, ones that were usually designed with the idea of professional staff handling, often in multiple interactions, individual items or objects. Traditionally, that system never worked all that well, so I start with some pessimism about any additional expectations that would increase that burden.

We've most often responded to the demands of technology with the idea of forming partnerships – and I hope that brief description of our plight indicates that the partnerships we have in mind are with entities that have more resources than we have. In other words, entities we expect to support us. I've reason to believe that our current definitions of the quid pro quo are not all that compelling.

We absolutely have to present some business case for cooperation with government agencies. There has to be a return on the investment. Authenticity, integrity, enhanced metadata … those terms, frankly, have little resonance in government or business. I recently had a chat with a lawyer who specializes in electronic discovery and information management, who's been involved in some very complex litigation using electronic records. He basically summed up his advice as "look for a solution that's good enough." He suggested that we'd have more support in changing our legal mandates than getting our partners actually to invest in realizing our traditional definitions of records management requirements. He offered to help. And I think he's right, that there is evidence of that "good enough" approach at work in UETA, the Sedona Conference and so on.

As you can gather, I'm leading towards a suggestion that we expand our perspective on the whole issue of managing and preserving digital content because many of the terms we are using and many of the questions we are asking are not those of my partners – the people who are the sources of funds and the creators of the records.

We could dismiss that difference as the result of their naivete – what do they know about preservation? – and in some cases, they are naïve. But we can't simply ignore them because we're asking them to pay for our proposed solutions. We have to speak their language, we have to address their concerns. I can boil those down:

- lower costs – we can't just ask for more
- usability – people want manageable solutions
- less human intervention – that certainly relates to costs, but it also identifies a key criterion used in the evaluation of proposals, the effective and innovative use of technology
- prioritization – there has to be some means to distinguish what's more important, where to focus attention, resource and energies
- add value – and this means a demonstration of some return on investment, some immediate improvement in operations that justifies funding and enhances the current business routine

And I think the SRB and grid technology can address those concerns. So I'll ask, in conclusion, what does resonate with our partners and potential partners? What can we offer, what can we sell?

- Interoperability – our solutions have to function simply as middleware or as modular components to the business applications of the records creators.
- Infrastructure independence – very often defined in terms right now, of moving and sharing data between systems, of XML
- Secure storage – preservation doesn't have the same persuasive force as secure storage. After Katrina and with homeland security concerns, secure offsite storage is very much on the minds of my legislative masters.
- Access –we don't just preserve material, we make it more useful now. Grid technology certainly presents an improvement in that sense over tape backup. There's a question about bandwidth and there are some debates about relative costs, but I think the benefits more than balance those concerns (or will).
- Significance of being innovative– there is certain premium attached to solutions that are perceived as being forward looking. Ask Apple about it.

So in our efforts, especially in the more substantive efforts we're putting in to our e-legislative project, the SRB has passed the scrutiny it's receiving. It answers the needs of some very sophisticated analysts both in the CA and MN legislatures. And it has been favorably received in recent presentations I made to NALIT and a legislative IT task force in VT.

Mentioning the SRB really makes you think of SDSC, though, and the SDSC's role in supporting a technology and application. So far, you really don't get one without the other. So let me close with some suggestions about that partnership.

The organization - the SDSC – is not designed to support software in the same ways that a private company is. The people are great, fabulously helpful, always very pleasant. There is room for confusion. There are some diffuse sets of responsibilities. There are ongoing waves of improvements that tend to make what you do know obsolete very, very quickly. And there are no implementation manuals. We've proposed to Reagan and Richard an idea to work on that and I hope it goes forward.

In all, though, we're very happy about our collaboration. As I mentioned, we don't have the staff, the expertise, the resources, whatever, to develop our own preservation solutions. We barely have the capacity to analyze all the options. We are necessarily going to rely on somebody else. The principal criterion I use to evaluate options, in addition, to those I mentioned in reference to our partners, is the likelihood of being able to specialize – what services or functions can we layer on top of what our partners will do. So I ask what capacity we can develop. And we are, of course, concerned with being part of the developing national cyberinfrastructure, to borrow the term of the moment. With that, I can easily say that we like to work with the SDSC and the SRB.

8.1.4 Comments on Points Raised in Panel 1

Reagan Moore, San Diego Supercomputer Center

The above comments for Panel 1 emphasize the need to choose appropriate management policies for the preservation environment. Many of the features in the SRB data grid have been driven by the multiplicity of management approaches taken across data sharing, data publication, and data preservation environments. We do believe there is a common set of core requirements that are present across all of the data management approaches.

As was described by Margaret Hedstrom, for a given preservation environment, the level of granularity for authenticity and integrity metadata must be chosen. We see multiple variants of this. Through the life-cycle data requirements guide, NARA assigns some attributes to a record group, some attributes to a specific record series, some to a file folder, and some to an individual object. Thus different levels of granularity are used to manage authenticity.

When we work with large collections of scientific data, we find that a reasonable effort has been expended in creating uniform authenticity information. Thus in astronomy, each image is supposed to have an associated FITS header file that details the creation properties of the image. Sky surveys with 5 million images all have FITS headers formed the same way. I agree that records that do not have common authenticity properties will be much harder to manage. On the other hand, they are harder to create in large numbers.

The questions of scalability and cost depend upon the choice of storage system and database that are used to implement the preservation environment. The same data virtualization mechanisms work on PCs and commodity disks as well as on clusters, SANs, and massive archives. Assessments of the cost associated with scaling the size of the archives are definitely needed. I expect management policies (for number of copies, accessibility, frequency of integrity checking) to drive the cost.

The issue of end-user access is also important. I agree that the users should be able to use their preferred search and access mechanisms. This capability again is one of the features that data virtualization systems enable. The access mechanisms are decoupled from the storage protocols, making it possible to add preferred access mechanisms. A noteworthy example is the integration of DSpace and Fedora on top of the SRB data grid. Both Carl Lagoze and MacKenzie Smith describe the preservation support they enable by adding services on top of the SRB data grid in panel 3.

The issue of open versus closed archives is strongly tied to the management policies of the institution. I visited the Australian National Archives in which the records are preserved in a "deep archive" that is inaccessible from the external world. On the other hand, the goal of NARA is to make all records publicly accessible through the web. Obviously, both capabilities can be provided by federating two independent systems.

I also agree that preservation environments are "living entities", in which new technology is continually incorporated. Different components of the preservation environment are upgraded at different times. The management of the technology evolution (when to incorporate new software and hardware) requires careful thought and planning. The goal of infrastructure independence is to ensure that the incorporation of new technology is feasible, with minimum disruption to the parts of the system that are not changing. Again data virtualization helps in this by decoupling metadata management from data management, decoupling management of encoding format from management of data bits, decoupling access mechanisms from storage protocols, decoupling trust management from storage repository management. Indeed one of the goals of the workshop is to decide whether the virtualization layers needed for preservation have been incorporated in the SRB data grid technology (or can be added). I expect a preservation environment to require multiple technologies, including the data virtualization mechanisms provided by the SRB.

8.2 Panel 2: Questions on Preservation Applications Experiences with SRB

1) Scalability: Support for millions of files, support for hundreds of terabytes, support for millions of records.

- Assess scalability. Does the preservation environment scale in size to the number of records that will be archived for future record groups? Does SRB schema extension capability adequately support scalable preservation metadata and provenance metadata (NARA Life Cycle Data Requirements Guide metadata hierarchy) [34]? Example capabilities include:
  - Does record ingestion scale (bulk operations for registration, metadata movement)?
  - Does data movement scale (containers for small files, remote operations for metadata extraction)?
  - Does the MCAT scale in management of system, user-defined and extensible schema?
  - Does the search performance scale?

2) Interactivity: Does the preservation environment provide mechanisms to maintain interactive response?

- Assess interactivity: Distributed storage systems are used within preservation environments to minimize risk of data loss. Can interactivity be maintained in the distributed environment? Example capabilities include:
  - Does the system support multiple metadata catalogs to manage high load levels?

- Does the system support interactions with workflow systems for automated processing?
- Does the system provide interfaces that minimize manual interactions?

3) Extensibility: Does the preservation environment provide mechanisms to integrate new types of access methods, new types of storage systems, new types of data formats?

- Assess extensibility: Can new infrastructure be incorporated into the preservation environment without impacting authenticity and integrity?
  - What are the principal access methods that are needed for preservation environments? How are interactions with workflow systems integrated into the same system along with interactive web browsers?
  - What are the preferred types of storage systems? How will content addressable storage systems or object-based storage systems impact preservation environments?
  - What requirements do specific data types impose on the preservation environment? An example is support for data streaming for displaying video.

8.2.1 Council for the Central Laboratory of the Research Councils

Kerstin Kleese van Dam, CCLRC

CCLRC [6] sees SRB as one of the key technologies to deliver long term data curation, delivering a useful abstraction layer between the physical storage and format of the data and the higher level Information concepts. It is used within a 4 tiered system with:

1) general metadata/data browse, assessment and access through data portals, command line interfaces and programme libraries
2) metadata catalogues and representation information
3) SRB layer for - data ingestion pipeline, logical grouping of data, data delivery
4) physical storage of the data

We believe that SRB fulfills this important role of storage virtualisation very well as well as delivering important information on ownership, data formats and storage format and location.

We see a number of challenges for SRB in the coming years if it wants to become a major player in the long term digital curation work. At present SRB is developed as a research project, whilst it is usually running very stably, monitoring, error handling and documentation are not as well developed as we would hope if run as a production service. There is also the question of being able to guarantee/monitor the successful completion of longer transaction e.g. multi-staged transfers within SRB which would be very useful for such a system. We have experienced some performance problems in areas like the central MCAT where the code is currently written to support a multitude of different database systems and not making use of the available features of more advanced products such as Oracle - a high performance productions version would be desirable. Finally if a wider adaptation of SRB is sought for long term data curation a solution to the licensing problem and support for past versions will be essential.

Overall we would like to iterate that we see SRB as an ideal technology for the challenges of long term digital curation, however to develop its full potential a number of crucial areas will need to be addressed in the future.

8.2.2 National Archives and Records Administration

Mark Conrad, NARA

At the National Archives and Records Administration (NARA) we are developing the Electronic Records Archives (ERA) System. The ERA Vision Statement says, "ERA will authentically preserve and provide access to any kind of electronic record, free from dependence on any specific hardware or software, enabling NARA to carry out its mission into the future." The technology does not presently exist to meet all of the requirements placed on the ERA system. As a result, the ERA Research Division has established a number of partnerships to look at emerging technologies (technologies that may be available in the marketplace in 3-5 years) and to evaluate those technologies' potential for meeting some of the requirements associated with the ERA system.

One of our continuing key partnerships is with the National Science Foundation, Office of Cyberinfrastructure and the San Diego Supercomputer Center (SDSC). We are testing the Storage Resource Broker (SRB) as a data and storage virtualization technology and a key component of the infrastructure in our Virtual Archives Laboratory (VAL or "the lab"). Using the SRB, we have established a "Transcontinental Persistent Archives Prototype". This consists of a data grid that presently has five nodes – the National Archives at College Park, MD and at Washington, DC, SDSC, University of Maryland Institute for Advanced Computer Studies (UMIACS), and Georgia Tech Research Institute (GTRI). The data grid supports most of our collaborative research.

The requirements for the ERA system revolve around three central themes - scalability, evolvability, and extensibility. We have used the SRB to conduct tests in all three areas.

Scalability: We have acquired test collections of electronic records from many agencies of the Federal Government. We are testing and evaluating the management of those collections using the SRB. The collections consist of millions of files. They are logically organized in hierarchical collections (e.g., record group, series, file unit, item, file). We are able to perform both bulk and focused operations (ingest, description, replication, deletion, etc) on the records and their metadata at arbitrary levels of the hierarchy. The SRB has allowed us to carry out operations on these test collections in hours that would have taken months or years or not been possible at all using NARA's present processes for handling electronic records. The results accrued in our research using the SRB enable us to empirically assess data management and architectural factors concerning storage, i/o, bandwidth, and latency as they affect scalability. This information is a tremendous asset for risk mitigation in the development of the operational ERA system. The ERA system will need to handle something on the order of ten trillion digital objects. While we are not currently testing the SRB at that scale, we are running tests at orders of magnitude above most, if not all, existing electronic record repositories.

ERA Research is also supporting the Persistent Archives Testbed project (www.sdsc.edu/PAT). This project is a partnership between SDSC, several state archives, and the Stanford Linear Accelerator (SLAC) Archives and History Office to test the use of the SRB in smaller repositories. Each of these repositories has successfully established an SRB node and they are testing the usability of the SRB in carrying out all archival functions. This project has demonstrated that the SRB can scale down as well as up.

Evolvability: Using the SRB as a brokerage tool informs our understanding of factors contributing to infrastructure independence. During our research we have migrated our test collections across several versions of the SRB without losing a single record or the related metadata. We have stored, managed and moved records across different media, file systems, databases, and operating systems without any data loss.

Extensibility: The SRB is used to support digital libraries and large collections of scientific data around the world. One of the unique requirements that NARA brought to the collaboration with SDSC was the need to support hierarchical archival descriptions of electronic records [70]. At the time we first raised the issue with SDSC, the MCAT was not able to adequately accommodate the metadata used in archival descriptions at NARA. Our research partners at SDSC and UMIACS were able to demonstrate the extensibility of the SRB by developing a web-based, database-driven application that could accommodate hierarchical archival descriptions and directly access the electronic records no matter where they were stored on the SRB-enabled data grid. More recent versions of the SRB have included support for extensible metadata schema. We have not yet tested this functionality, but initial analysis leads us to believe that we should be able to store hierarchical archival descriptions within the MCAT.

We have also tested the SRB as a backend for several other applications. The Producer-Archive Workflow Network (PAWN) is an application developed by our research partners at UMIACS to manage the disposition and transfer of electronic records and their associated metadata from the records creator's system to the archival repository. We are supporting SDSC and our research partners at MIT in the integration of DSpace with the SRB. Our research partners at GTRI are currently modifying the archival processing tools that they developed for the George H. W. Bush Presidential Library so that they can be used in conjunction with the SRB. SRB's use of multiple layers of abstraction provides us with a great deal of flexibility in interfacing with other applications.

The Federal Government produces electronic records in thousands of formats. Our test collections do not contain examples of each of these formats, but the SRB has been able to handle all of the diverse formats present in our collections. Our collections include digital text, audio, video, images, virtual reality, geospatial data, engineering drawings, and complex models in multitudes of formats.

Interactivity: There are many ways to interact with the SRB. NARA represents a different user community than the typical users of the SRB. Our ability to use the SRB effectively is in itself a demonstration of the flexibility and user-friendliness of the SRB. SDSC makes a number of client applications available to use with the SRB. We use three of these on a regular basis. The drag-and-drop capabilities of inQ make it easy to carry out operations on a single file or an entire collection. MySRB provides access to records stored in the SRB to anyone with web access and the proper permissions. The Scommands make up the most powerful set of tools for working with the SRB. While the command line interface may not be as user-friendly as that of inQ and MySRB, the power of these tools makes learning to use them worthwhile. In addition, the SRB includes APIs that allow a great deal of flexibility in integrating the SRB with other applications. In our work with our research partners we have used perl scripts, Kepler workflows, and cgi-based applications, among others, with the SRB.

Summary: Our findings demonstrate that the SRB provides tremendous scalability, evolvability, and extensibility for accessioning, processing, and providing access to archival records. It provides support for the metadata necessary to maintain the provenance of the records. The SRB provides tools that support records integrity and security including, audit trails, access controls, checksums, and the ability to easily create geographically-dispersed copies of records. All of this functionality is available across multiple platforms.

8.2.3 UCSD Libraries

Luc Declerck, UCSD Libraries

The University of California, San Diego Libraries currently uses SRB as the underlying file storage layer for:

- Preservation and access to the Libraries' largest digital collections
- Preservation and access to local technical documentation
- Backup of selected production systems

We started by using an SRB instance located at SDSC and with the help and guidance of SDSC staff successfully transferred, over a period of 10 days, 200,000 .tif image files (approximately 4 Terabytes) from multiple servers in the Libraries.

Since then, we implemented our own SRB instance on local hardware using our own locally assembled grid bricks (approximately 6 Terabytes) and reloaded the 200,000 .tif image files, along with their derivatives files (smaller .jpgs for web and thumbnail display) totaling 800,000 files and roughly 5 Terabytes of data into that SRB instance.

We developed workflow and status tracking procedures for batch loading of multi-file complex digital objects. We also architected our environment such that our file system would remain independent of any infrastructure, including SRB, by:

- Adopting the California Digital Library ARK based persistent file naming convention and naming files with ARKs at both the physical and logical level to enable file recovery at the physical file level, i.e. on the disc (so that we can recover when and if MCAT fails)
- Placing metadata as files in the file system with the same ARK prefix to implicitly bind common files through a convention of a shared file name prefix with variable suffixes. This approach enforces the relationship between a particular content file and a metadata file associated with it. Such a file pair, along with any other files sharing a common ARK prefix, constitute an Archival Information Package (AIP) for complex digital objects with an arbitrary number of components.
- Utilizing the SRB Java Jargon APIs to create a servlet for ARK-identifier-based retrieval of files from SRB.

Authenticity related information is recorded in the metadata file associated with every content file or collection of content files. At this time the associated metadata file is written in accordance with the METS XML schema and the UCSD Libraries have developed specific METS profiles [31] to be used to encode metadata related to particular digital content categories. Thus, support for authenticity is largely relegated to other systems responsible for METS file maintenance.

Support for integrity is likewise largely relegated to other systems responsible for METS file maintenance. However, it would be very helpful if there were more integrated user friendly tools and interfaces like InQ to assist with:

- Ongoing check-sum verification
- Production of management reports (e.g.: number of collections, number of files, storage capacity utilized)
- Zone and file replication management
- Reconstructing a lost or damaged MCAT from raw file systems

During this time, we also collaborated with MIT and SDSC to integrate DSpace with SRB (DSRB project). The project accomplished the following:

- Modification of the DSpace single-item workflow code to enable transparent storage and retrieval of files into SRB as an alternative to “local” file storage.
- Modification of the DSpace batch loading process to permit the bulk registration of pre-existing SRB collections into DSpace.

Our collaboration will continue in 2006, with the PoLicy Enforcement in Data Grid Environments (PLEDGE) project, which has the following goals:

- Identification of necessary policy expression and information life cycle management ontologies to support large-scale digital collection management
- Further specification and development of a more modular, scalable architecture for the DSpace digital library platform (DSpace 2.0)

- Initial development and testing of distributed, federated collections built on data grid storage and managed by DSpace data curators
- Demonstration of SRB support for centralized mechanism to replicate and federate collections across institutional boundaries
- Demonstration of support for preservation of authenticity and integrity during exchange of documents between integrated DSpace/SRB systems

We are also working with UCSD-TV and SDSC to preserve UCSD TV's "Conversations with History Program" videos in SRB (DigArch project). This project involves:

- Modification of the Libraries AIP model for long-term preservation
- Integration of data preservation component in UCSD-TV production workflow with the Kepler workflow system

Our planned future uses of SRB include:

- Implementation of a second local Libraries SRB zone for backup purposes, and an independent 'development' SRB Zone isolated from our 'production' SRB environment
- Implementation of automatic replication of data to other Zones (using Master-Slave, Snowflake, and/or Archival models)
- Exploring opportunities for SRB to SRB transfer into the CDL Digital Preservation Repository and other institutions?

Our preservation applications experiences with SRB:

1) Scalability: Support for millions of files, support for hundreds of terabytes, support for millions of records.
Our primary observation is that scalability of SRB has not been taxed by the UCSD Libraries. Transfer of 5 TB took 10 days. However this was due to serial disk read/write limitations. Moving to a more parallel oriented transfer model could improve this performance, and could shift more burden to the network. Fortunately, such large transfers are generally one-time events which in our environment can generally be tolerated. Ongoing synchronization of collections should be more manageable. SRB offers many options in this arena, but there is a lot of room for clarity in the documentation.

2) Interactivity: Does the preservation environment provide mechanisms to maintain interactive response?
Slow retrieval rates were occasionally experienced on SDSC SRB systems, as a result of files being automatically moved from higher speed cache to slower storage medium, as they age or are left unused, i.e., moved to HPSSS tape system or tape shelves

Probably as a result of early SRB installation difficulties, MCAT (Oracle) performance issues were encountered:

- Retrieving items out of SRB was very slow (easily observable with InQ)

- Proper database indices must be built for optimal performance. However difficulties were encountered when all designated indices were created
- May be related to recent Bug #186

We perceive SRB largely as middleware, however our DSpace/SRB integration work, along with integration of SRB with other locally developed software, has enabled us to achieve satisfactory levels of interactive response.

3) Extensibility: Does the preservation environment provide mechanisms to integrate new types of access methods, new types of storage systems, new types of data formats?
SRB has a rich set of APIs that enable the development of custom access methods. At the UCSD Libraries, most of our focus has been on development using the Java/Jargon APIs, which we have used in our DSpace/SRB integration and other local development projects. We do, however, actively use the S-command interface and are currently eager to have an updated version of those commands for Windows and the latest SRB release. We also expect to be involved in usage of Kepler to aid the workflow involved in data ingestion.

Our local usage of SRB-based storage systems is limited to grid-bricks assembled from large arrays of commodity disk drives. This has proven highly scalable for us, and in this environment data format issues have not emerged.

8.2.4 Academia Sinica

Eric Yen, Academia Sinica
The SRB system in Academia Sinica is used for the long-term preservation of the digital contents produced by the digital archives projects, which are part of the National Digital Archive Projects in Taiwan. The system was deployed by the Academia Sinica Grid Computing Centre (ASGC) in early 2004, which was constituted from 7 sites in different buildings, linked by a dedicated fibre campus network, and provided 60 TB capacity in total (before RAID-5). ASGC is working on a new generation of Grid-based research infrastructure in Academia Sinica and in Taiwan, by using gLite and OSG as the Grid middleware. The DataGrid is a major part of this infrastructure, and the SRB is the first and the largest (in terms of the data volume) DataGrid in our academy right now.

Currently, the SRB DataGrid in Academia Sinica has 1,343,149 files and 28.4 TBs of data, where all files are preserved in two copies on different sites. The status can be found from the monitoring system we developed at http://srb.grid.sinica.edu.tw/asmss_monitoring/index.php. It is estimated that the digital contents generated in Academia Sinica will be 55.75TB in 2005, 49.4TB in 2006 and 88.9TB before 2005. ASGC plans to extend the SRB system by adding 60 TB capacity in January 2006, and another 60TB after the third quarter of 2006. Also in 2006, ASGC is funded by the Taiwan National Science Council to deploy SRB for the whole National Digital Archive Projects, which covers 8 major museums, archives and libraries of Taiwan, dispersed mainly in Taipei and with two sites that are 200KM away from Taipei.

The major problem we encountered in the SRB applications are the performance degradation of MCAT, after having more than 10M files. Our MCAT is implemented with Oracle 10g. All the digital contents are ingested with well formatted metadata for each object and each collection. In the beginning of the digital preservation project, we also wanted to implement the metadata schema in SRB for content-based file search/retrieval, but it seems not so flexible to extend the SRB schema to cope with complicated metadata structure of digital collections. Another issue is about the users themselves. Since InQ has not fully implemented the SRB commands, users have to learn a new set of SRB commands that are totally different to their acquainted use of FTP. Users expect to have a GUI based interface for the migration, retrieval, search and checking of the required files in SRB. We are very happy to work closely with the SRB Team in SDSC for the robust long-term preservation environment based on SRB in the future.

8.3 Panel 3: Questions on SRB as a Digital Preservation Cyberinfrastructure

1) Support for a production environment: What is the relationship between system robustness and the long-term preservation of records?

- Assess production quality. Example capabilities include:
    - Is the system maintainable?
    - Is academic access to source sufficient, or is open-source required?
    - Is the technology continuing to evolve to include new standards?
    - Is the effort needed to migrate to new technologies manageable?
    - Are the labor support requirements for the system manageable?
    - Are the software maintenance requirements manageable?
    - Is the documentation adequate?

2) Support for risk mitigation against data and metadata loss. This includes replication, federation, semantic versions, backups of both data and metadata.

- Assess risk management and disaster recovery: A preservation environment assures against corruption of electronic records and against corruption of the preservation metadata. Example capabilities include:
    - Does the system support replication of data, validation of replicas, and synchronization of replicas across multiple types of storage systems?
    - Does the system support federation of metadata catalogs, and synchronization of metadata catalogs across administrative domains?
    - Does the system provide end-to-end validation mechanisms for assessing data and metadata integrity?
    - Does the system provide mechanisms for disaster recovery?

3) Support for automation of preservation processes. This includes interfaces to workflow environments for executing preservation processes on record groups and record series.

- Assess automation capabilities: Can all aspects of preservation be automated, from appraisal, to accession, to arrangement, to description, to preservation/storage, to access? Example capabilities include:
  - Does the system preserve authenticity, integrity, and infrastructure independence across each preservation process?
  - Does the system provide mechanisms to process audit trails, describe access controls, summarize storage utilization?
  - Does the system provide support for implementing workflow actors to drive each preservation capability?

8.3.1 Fedora middleware and SRB integration

Carl Lagoze, Cornell University

Fedora supports rich information objects to address preservation issues related to object and knowledge virtualization. The native storage of Fedora objects is in XML - a composite document format called FOXML. However, the Fedora digital object model includes a number of richer features including:

1) Remote datastream reference - a datastream may include locally managed data or a reference to external data.
2) Web service integration - disseminations from digital objects can be based on the interaction of contained datastreams and interactions with local or remote web services. The nature of a dissemination (local, remote, static, web-service produced) is opaque from the access perspective.
3) Semantic web integration - The object model provides hooks for RDF-based relationships among objects. These relationships are stored in a triple store that itself is exposed as a web service.

Fedora by itself is not a preservation environment, but it is "preservation worthy" for the following reasons:

1. The object model includes a fine granularity versioning and audit trail capability providing access to the full history of an object through the API. The native FOXML store of digital objects encapsulates all data, metadata, relationships, and service links. An entire Fedora repository can be "rebuilt" based on these XML representations alone.
2. Fedora repositories are compliant with the Reference Model for an Open Archival Information System (OAIS) due to their ability to ingest and disseminate Submission Information Packages (SIPS) and Dissemination Information Packages (DIPS) in standard container formats such as METS and MPEG-DIDL.
3. The implementation of the low-level store in Fedora (LL-Store) is modular and therefore the default file system based storage mechanism can be easily replaced with more preservation targeted storage architectures.

The last feature is the foundation for the integration of Fedora and SRB, which is now being evaluated and will be included in a later release of Fedora. Essentially, this integration maintains all features of Fedora (rich object model, API, management primitives), replacing the basic file storage with the robustness of SRB (location independent, networked, replicated storage). The combination of these features provides a very attractive combination for applications needing persistent access to rich objects. The two systems act in complimentary fashion - Fedora providing the object model, SRB the storage model - to accomplish this.

Note, however, that one aspect of the Fedora object model still remains challenging even in this combined SRB implementation. That is the integration of services into the document model. Given that a dissemination of an object can in Fedora be linked to a web service, how can we preserve that service, and transitively the digital object access point that is dependent on it? Preservation of computational services remains an unanswered question that is not addressed in the framework of SRB, and has only partially been addressed by emulation-type systems.

Finally, the SRB/Fedora integration provides the basis for the persistence needs of the National Science Digital Library [49]. We believe that users of the NSDL need to have access to web-based learning resources independent of their actual persistence in web space. That is, if a teacher or student finds a web-based resource on January 15, they should be able to access it in its same form on February 15, even if the actual web version has changed or disappeared. This motivates our cooperation with SDSC to crawl and archive NSDL resources. It also motivates our goal to implement the NSDL data repository (NDR) as a Fedora versioned repository on top of the SRB storage mechanisms. We hope through this to provide transparency of access in the face of temporal transitions.

8.3.2 DSpace digital library and SRB integration

MacKenzie Smith, MIT Libraries

The DSpace system design was based on the OAIS reference model and implements all of its components. It is intended to address the entire information life cycle of digital objects, primarily those generated by research and cultural heritage organizations. The digital objects in DSpace are compositions of locally-managed datastreams and associated metadata, including descriptive, administrative, and technical, and authenticity metadata. A schema based on the curating organization's structure is used for organizing and relating sets of documents in flexible configurations. DSpace provides a range of pre-built services for ingestion, curation, and access to archived content by end-users.

1. Submission workflow services support ingestion of new digital objects and associated metadata by their creators or their intermediaries. Batch ingestion and registration are also supported. SIPs (Submission Information Packages) are structured using a registered METS profile.
2. Administration and curation services are provided for Archival Information Packages and for administering the archive over time using administrative (technical, rights) and provenance metadata (ongoing collection management and preservation events). Some administrative and preservation policies can be established in the system by digital archivists, and much research and development is being invested in this aspect of the system.
3. Access services are provided via default and customizable Web user interfaces, via common protocols such as Open Archives Initiative (OAI) [52], and via a set of Web Services (e.g. standard search and retrieval services). SIPs can be

provided by these Web Services using the registered METS profile and a number of other packaging standards (e.g. IMS Content Packages [25] required for educational technology systems).

Preservation concerns are met in DSpace by several means. Straightforward storage of the structured objects and associated data managed by the DSpace administrative user interface; auditing services to monitor content for media faults, data corruption, etc.; a set of policies encoded in the system by the digital archivists (e.g. digital file format support policies, access policies, etc.); a History system that tracks preservation and other collection management events and related management reports (i.e. provenance metadata); documented format-specific preservation strategies that are developed by curators.

DSpace supports a storage layer based on the standard file system and has recently added support for an SRB-accessible data grid as an alternative to manage data distributed across multiple storage systems, replication of files, and federation with other DSpace instances. This integration allows for the registration into DSpace of collections already stored in SRB, or placement of DSpace submitted items into SRB-managed storage. DSpace 1.3 with SRB support was released last summer and there are a number of DSpace sites using a pilot SRB-based data grid (hosted at SDSC). There is demand from the DSpace community (150 production digital archives at the present time) for data grid-based storage services and storage utilities, and SRB integration has been an important advance towards that goal. Using SRB-based storage to achieve geographically-distributed content replication is a high priority for the entire DSpace community.

Cyberinfrastructure business models: The data grid as instantiated by SRB is now well tested with a variety of preservation platforms (e.g. DSpace) and institutions are beginning to build significant local collections of digital objects that would benefit from large-scale distributed, federated storage. At the present time these institutions can further test this infrastructure using storage provided locally at SDSC, but before moving into true operations we need to understand the business models behind the data grid, and the cyberinfrasture more generally.

Issues that must be addressed by the digital library and archives community before serious investment in data grid technology can be made include

– Organizational models for cyberinfrastructure provision
– Cost models for providing cyberinfrastructure services such as data grid storage utilities to various user communities (i.e. market segmentation and appropriate service definitions with appropriate recovery costs)
– Governance and social dynamics of interaction between cyberinfrastructure providers and user communities (e.g. for the SRB open source software)
– Legal frameworks for contracts or other types of relationships between cyberinfrastructure providers and user communities

In normally developing markets such services are provided by commercial (or non-commercial) vendors with distinct product or service offerings that have well-defined associated cost models and business plans. The cyberinfrastructure has no such business framework and won't for some time.

It is still unclear whether formal non-profit governance models such as exist for the Internet (i.e. the IETF [24] and ICANN [23]) would be most appropriate, or whether a more informal approach such as the W3C [76] for the Web is best.

It is also unclear how actual services will emerge. For example in the case of SRB and data grid utilities, should SDSC provide a revenue-generating service to any organization who needs storage and is willing to adopt SRB? Only to accredited research institutions or government agencies? Or offer different pricing models for different organizational types? With what service-level agreements? Or should different communities (e.g. research libraries, bioinformatics data archives, national archives) build their own community data grids and develop shared operations and cost recovery solutions among themselves?

As we move out of the initial phase of research and experimentation into serious examination of SRB as a production data grid technology we need to define a process for the business context to emerge, or we risk building dependencies on ephemera.

8.3.3 Real-time Data Management Systems

Frank Vernon, University of California San Diego, Scripps Institution of Oceanography

Real-time, gridded sensor networks such as ROADNet [20,67,60] collect large volumes of multidisciplinary sensor data that must be buffered (held in accessible storage) for immediate analysis and redistribution, as well as archived for future re-examination for long-term trends and comparison of recent to historic events. In order to be useful to the wide range of users, data contributors, science teams and monitoring operations, these diverse datasets need to be accessible in some kind of centralized manner, though with component data sets often distributed amongst different research groups. Data access methods must be straightforward for end-users and for authors of analysis, processing, and display operations. Furthermore, any such system must be able to handle the diverse system and domain metadata, which may vary widely from subdiscipline to subdiscipline, often blurring the boundary between data and metadata. Finally, although the resultant system may be used for research-based prototypes and short-term (low investment) monitoring experiments, the need for constant availability and for support of large-scale, mission-critical scientific and monitoring operations requires a robust cyberinfrastructure characteristic of a hardened production system.

For the preservation of scientific data, at least as far as the data structures to support multiple types of data is concerned it seems we have made a lot of progress. The key issue here is perhaps scalability of raw storage capability and access methods that allow full exploration of the data structures being archived. Clearly the SRB has a lot to offer in those aspects. The harder issues are the preservation of scientific analysis tools and visualization systems cited in item 1. Historically there has been a continued evolution of the code base for these tools, where the code has been maintained and upgraded by one or another form of community investment. Constant changes in the hardware, operating system, and supporting-application environments drive the common truism that "software rusts." Thus, for preservation of multiple types of scientific data structures, it's quite possible that systems based on data format description languages will be sufficient. However for the preservation of the scientific analysis tools and visualization systems, we are skeptical that the description languages alone will suffice. The latter seem to require a continued investment in the software connectivity between those data structures and the state of the art tools that use them, or at the very least between those data structures and modernized versions of the tools that were originally written for them.

Regarding item 3, "Can all aspects of preservation be automated," the lesson from real-time monitoring systems and the automation of data processing appears to be a slightly more conservative statement: "Over time, as our community understanding of the processes in the 'manual workflow' matures, more and more of them can be progressively automated, with large advances over the long term." Witness the progress made in earthquake seismology, where the community has steadily moved into automated digital data collection, then automated phase picking with hand location of earthquakes, then automated earthquake locations, and now automated generation of many sophisticated data products such as magnitudes, moment tensors, shakemaps, etc. It is hard to predict the exact timescale for substantial automation of large-scale data-preservation tasks, or the final level of sophistication, however one imagines that with enough investment, great strides are possible. For an interesting object lesson in the leaps possible for automation, compare the capabilities for automated gene-sequencing at the end of the human genome project (2003) to the small-scale hand analyses that were being performed at the project's inception fifteen years ago [quick overview page at http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml].

Finally, a note on open-source licensing, pertaining to the question "Is academic access to source sufficient, or is open-source required?" This requires careful thought, probably focused on the end goal of maximizing large-scale data preservation. Again citing the Human Genome web-page above, tailoring some kind of transfer of technology to the private sector can catalyze large-scale industries that can further the project's cause by encouraging application development. The Open Source Initiative [http://www.opensource.org] lists over 50 different approved "open-source" license models, and if the lessons from the recent 2005 PHP conference [http://zend.kbconferences.com/] are at all indicative, the development of business models for applying open-source code to enterprise-level computing seems destined for explosive growth. Given the similar scales shared by commercial enterprise computing and the data-preservation applications apparently under discussion here, it seems the data preservation community would want to harness as much of the potential available in the business / academia interface by choosing an appropriate approach to licensing.

8.3.4 Additional examples of data management technologies

1) PAWN, Multivalent Browser, and SRB integration. The NARA research prototype persistent archive uses the PAWN (Producer Archive Workflow Network) [57] to manage the accessioning of digital records into a preservation environment. PAWN implements the Ingest Process defined in the Open Archival Information System reference model. PAWN uses METS schema to encapsulate content, structural, descriptive, and preservation metadata. PAWN manages the staging and assembly of data, the transport to the archives, and the verification of metadata, bitstreams and preservation information after reception at the archives. Three SRB data grids are used to manage copies of the digital records. The data grids are federated to ensure consistency of both digital records and authenticity information. The Multivalent Browser is used to provide access to records written in the pdf format, and thus serves as the object virtualization layer.

2) Cheshire, Multivalent Browser, Kepler, and SRB integration. Cheshire is a digital library system developed at the University of California, Berkeley, and extended by the University of Liverpool. Cheshire incorporates an internal workflow management system for application of preservation processes, which is being integrated with the Kepler workflow environment (developed at SDSC and the University of California, Davis). Cheshire provides indexing tools and an information retrieval system. The Multivalent Browser technology provides media adaptors for parsing documents, and a separate set of behaviors for manipulating the parsed data. Media adaptors exist for pdf, LaTex, XML, HTML, and office products (through integration with OpenOffice). The SRB data grid technology provides support for data and trust virtualization.

3) Digital Format Registry. The University of Maryland has demonstrated a scalable and secure design based upon the use of web technologies for the registration of digital formats [15].

8.3.5 Comments on Points Raised on Panel 3

Reagan Moore, San Diego Supercomputer Center

The points raised by MacKenzie Smith have a wider implication. While UCSD owns the intellectual property rights for the SRB, UCSD licensed the commercial rights to Nirvana Storage. If an institution charges service fees for use of the SRB, then the institution has to pay a royalty on the service fees to Nirvana Storage. This arrangement is slowly being worked out and eventually will be published on the web site. Note that a royalty on a service fee of $0 is $0. Thus academic use without charging a fee incurs no royalty.

Any long-lived service needs a funding model, either through fees raised for use of the service, or through a commitment by the institution that provides the service, or through pre-paid support for the service (author publication fees). My personal objective is to have every collection of standard digital reference sets supported by multiple institutions. Thus if one institution stops support, the other institutions can continue to provide the service. In essence, this means replication of the data across the three institutions. One would then use data grid technology to federate access across the institutions. The SRB supports this model of sustainability.

Cost models are necessary for any institution. At the moment, the cost for archival storage at SDSC (including equipment amortization, maintenance, operations) is $600 per TB per year. The cost for disk storage is about $1500 per TB per year. The cost models are always point in time statements. We expect the cost to decrease substantially (historically each new technology provides twice the capacity at the same media cost). We expect holographic storage to dramatically reduce the cost of storage. Thus SDSC provision of 5 TBs of storage to DSpace is an effective cost of $3000 per year for tape storage.

A major requirement for a preservation environment is the evaluation of the trustworthiness of the software that is used. This trust is usually assigned by picking a vendor that has proven trustworthy over the years. If open source software is used, then the institution that is creating the preservation environment has to do its own assessment of the trustworthiness of the software (does the software contain trapdoors that skirt security requirements, does it work on the particular set of hardware, does it interact with other systems correctly, does it have the robustness that is needed). This assessment requires a higher degree of expertise (read higher administrative support cost). The savings one gets from open source software is balanced by the higher administrative cost needed to assign trust to the software. It is not obvious where the breakeven point is.

One way to avoid loss of governance is to rely upon replication of the digital records across three sites. If a site stops participating, one then looks for another collaborating institution and replicates the data. The goal is to avoid having a dependency upon a single institution. This works as long as the amount of time needed to do the replication is less than the lifetime of media. Thus the legal framework turns into an assessment of the archival capacity (how much data can be managed at a single site given that all of the data may need to be replicated to another site). As long as the replication is possible, the partner site can disappear.

The questions on cyberinfrastructure are effectively questions about whether NSF will provide long-term infrastructure for data management. This is well worth asking.

My version of the questions is the next to-last question. Each community needs to build their own data grid (control their own destiny), develop shared operations and cost recovery mechanism among themselves. Cyberinfrastructure is just one of the possible partners in building the larger environment. Additional questions then are:

- How many institutions are needed to build a viable preservation environment?
- What are the minimal requirements on institutions for new ones to participate and old ones to depart (minimum notice of leaving that is greater than the time to replicate the data)?
- What are the federation requirements on the digital records for replication across the partner institutions (Does every institution have a complete copy of data and metadata)?

- Does preservation introduce more stringent requirements on institution partnerships than joint digital libraries?

9. Acknowledgements

This workshop was funded as a collaboration between two NSF awards: IIS045995: *Robust Technologies for Automated Ingestion and Long-Term Preservation of Digital Information* and IIS0456055: *Digital Preservation Lifecycle Management: Building A Demonstration Prototype for the Preservation of Large Scale Multimedia Collections*. More information about preservation environments based on the SDSC Storage Resource Broker can be found at http://www.sdsc.edu/srb/ and http://www.sdsc.edu/NARA/.

10. References

1. Adore – A Modular and Standards-Based Digital Object Repository at the Los Alamos National Laboratory, TCDL Bulletin, Vol. 2, Issue 1, 2005.
2. APAC – Australian Partnership for Advanced Computation, http://www.apac.edu.au/
3. BaBar high-energy physics data grid, http://www.slac.stanford.edu/BFROOT/
4. Baru, C., R, Moore, A. Rajasekar, M. Wan,"The SDSC Storage Resource Broker," Proc. CASCON'98 Conference, Nov.30-Dec.3, 1998, Toronto, Canada, p. 5.
5. BIRN – Biomedical Informatics Research Network, http://www.nbirn.net.
6. CCLRC – Council for the Central Laboratory of the Research Councils, http://www.cclrc.ac.uk/
7. CDL/DPR – California Digital Library, Digital Preservation Repository, http://www.cdlib.org/programs/digital_preservation.html
8. Chen, A. E., E. Yen, Perspective of Grid Computing and e-Science in Taiwan, International Syposium on Grid Computing 2005, Taiwan, April, 2005.
9. Condor high throughput computing system, http://www.cs.wisc.edu/condor/
10. DFDL – Data Format Description Language, http://forge.gridforum.org/projects/dfdl-wg
11. DSpace digital library, http://www.dspace.org/
12. ENZO cosmology application, http://cosmos.ucsd.edu/enzo/
13. Fedora digital object repository system, http://www.fedora.info/
14. Foster, I., Kesselman, C., "The Grid: Blueprint for a New Computing Infrastructure," Chapter 5, "Data Intensive Computing," Morgan Kaufmann, San Francisco, 1999
15. Geremew, M., S. Song, J. JaJa, "Using Scalable and Secure Web Technologies to Design a Global Digital Format Registry Prototype: Architecture, Implementation, and Testing", Submitted to Archiving 2006.
16. Global Grid Forum, http://www.gridforum.org/
17. Greenstone Digital Library Software, http://www.greenstone.org/cgi-bin/library
18. GridFTP, http://www.globus.org/toolkit/docs/4.0/data/gridftp/
19. GSI – Grid Security Infrastructure, http://www.globus.org/toolkit/docs/4.0/security/

20. Hansen, T., K. Lindquist, F. Vernon, J. Orcutt(2004). VORBrouter: A dynamic data routing system for Real-Time Seismic networks. Eos Trans. AGU, 85(47), Fall Meet. Suppl., Abstract, SF41A-0749.
21. HDF5 – Hierarchical Data Format version 5, http://hdf.ncsa.uiuc.edu/HDF5/
22. Hedstrom, M., et. al. It's About Time: Final Report of the Workshop on Research Challenges in Digital Archiving and Long-term Preservation, April 12-13, 2002. Washington, D.C.: National Science Foundation and The Library of Congress, August 2003.
23. ICANN, Internet Corporation for Assigned Names and Numbers, http://www.icann.org/
24. IETF, Internet Engineering Task Force, http://www.ietf.org
25. IMS Global Learning Consortium, http://www.imsglobal.org
26. InterPARES - International Research on Permanent Authentic Records in Electronic Systems, http://www.interpares.org/
27. KEK high energy accelerator research organization BELLE data grid, http://belle.kek.jp/
28. Kepler scientific workflows, http://kepler-project.org/
29. LOCKSS – Lots of Copies Keep Stuff Safe, http://lockss.stanford.edu/
30. Ludäscher, B., R. Marciano, R. Moore, "Preservation of Digital Data with Self-Validating, Self-Instantiating Knowledge-Based Archives," ACM SIGMOD Record, 30(3), 54-63, 2001.
31. METS – Metadata Encoding and Transmission Standard, http://www.loc.gov/standards/mets/
32. Moore, R., M. Wan, A. Rajasekar, "Storage Resource Broker: Generic Software Infrastructure for Managing Globally Distributed Data", Proceedings of IEEE Conference on Globally Distributed Data, Sardinia, Italy, June 28, 2005.
33. Moore, R., F. Berman, B. Schottlaender, A. Rajasekar, D. Middleton, J. JaJa, "Chronopolis: Federated Digital Repositories Across Time and Space", Proceedings of IEEE Conference on Globally Distributed Data, Sardinia, Italy, June 29, 2005.
34. Moore, R., R. Marciano, "Technologies for Preservation", book chapter in "Managing Electronic Records", editors Julie McLeod and Catherine Hareto, Facet Publishing, 2005.
35. Moore, R., R. Marciano, "Prototype Preservation Environments", Library Trends Volume 54, Number 1 (Summer 2005), "Digital Preservation: Finding Balance", Johns Hopkins University Press.
36. Moore, R., J. JaJa, R. Chadduck, "Mitigating Risk of Data Loss in Preservation Environments", NASA / IEEE MSST2005, Thirteenth NASA Goddard / Twenty-second IEEE Conference on Mass Storage Systems and Technologies, April 2005, pp. 39-48.
37. Moore, R., A. Rajasekar, M. Wan, "Data Grids, Digital Libraries and Persistent Archives: An Integrated Approach to Publishing, Sharing and Archiving Data", Special Issue of the Proceedings of the IEEE on Grid Computing, Vol. 93, No.3, pp. 578-588, March 2005.
38. Moore, R., W. Underwood, "Preservation Environments for Digital Entities," InterPARES II report, June 2004.

39. Moore, R., “Operations for Access, Management, and Transport at Remote Sites,” Global Grid Forum, GFD.46, December 2003.
40. Moore, R., A. Merzky, “Persistent Archive Concepts,” Global Grid Forum, December 2003.
41. Moore, R., “Common Consistency Requirements for Data Grids, Digital Libraries, and Persistent Archives”, Grid Protocol Architecture Research Group, Global Grid Forum, Tokyo, Japan, March 5, 2003.
42. Moore, R., C. Baru, “Virtualization Services for Data Grids”, Book chapter in "Grid Computing: Making the Global Infrastructure a Reality", pp. 409-436, John Wiley & Sons Ltd, 2003.
43. Moore, R., “The San Diego Project: Persistent Objects,” Archivi & Computer, Automazione E Beni Culturali, l’Archivio Storico Comunale di San Miniato, Pisa, Italy, February, 2003.
44. Moore, R., “Preservation of Data, Information, and Knowledge,” Proceedings of the World Library Summit, Singapore, April 2002.
45. Moore, R., C. Baru, A. Rajasekar, B. Ludascher, R. Marciano, M. Wan, W. Schroeder, and A. Gupta, “Collection-Based Persistent Digital Archives – Parts 1& 2”, D-Lib Magazine, April/March 2000, http://www.dlib.org/
46. Multivalent Browser - http://multivalent.sourceforge.net/
47. NARA Research Prototype Persistent Archive, http://www.sdsc.edu/NARA/
48. NDIIPP - National Digital Information Infrastructure Preservation Program, http://www.digitalpreservation.gov/
49. NSDL – National Science Digital Library, http://nsdl.org/
50. NSF Teragrid, http://www.teragrid.org/
51. NVO – National Virtual Observatory, http://www.us-vo.org/
52. OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting, Van de Somple, H., M. Nelson, C. Lagoze, S. Warner, “Resource Harvesting Within the OAI-PMH framework”, D-Lib magazine, vol. 10, number 12, http://www.dlib.org/dlib/december04/vandesompel/12vandesompel.html
53. OAIS – Reference Model for an Open Archival Information System, http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html
54. OGSI – Open Grid Services Infrastructure, http://www.ggf.org/ogsi-wg/
55. OpenDAP – Open-source project for a Network Data Access Protocol, http://www.opendap.org/
56. PAT – Persistent Archives Testbed, http://www.sdsc.edu/PAT/
57. PAWN – Producer Archive Workflow Network, http://narawiki.umiacs.umd.edu/twiki/bin/view/Main/PAWN
58. Rajasekar, A., M. Wan, R. Moore, W. Schroeder, “Data Grid Federation”, PDPTA 2004 - Special Session on New Trends in Distributed Data Access, June 2004.
59. Rajasekar, A., R. Moore, F. Berman, B. Schottlaender, “Digital Preservation Lifecycle Management for Multi-media Collections, Lecture Notes in Computer Science, vol. 3815/2005, pp. 380-384, November 2005.
60. Rajasekar A., F. Vernon, K. Lindquist, J. Orcutt,, S. Lu, R. Moore, “Accessing Sensor Data Using Meta data: A Virtual Object Ring Buffer Framework,” DMSN 2005, Trondheim, Norway Aug 2005.

61. Rajasekar, A., Michael Wan, Reagan Moore, Arun Jagatheesan, George Kremenek, “Real Experiences with Data Grids - Case studies in using the SRB”, International Symposium on High-Performance Computer Architecture, Kyushu, Japan, December, 2002.
62. Rajasekar, A., R. Moore, B. Ludäscher, I. Zaslavsky, “ The Grid Adventures: SDSC’s Storage Resource Broker and Web Services in Digital Library Applications:, 4th Russian Conference on Digital Libraries, Dubna, Russia, October, 2002.
63. Rajasekar, A.,M. Wan, R. Moore, “mySRB and SRB, Components of a Data Grid”, 11th High Performance Distributed Computing conference, Edinburgh, Scotland, July 2002.
64. Rajasekar, A.,R. Moore, "Data and Metadata Collections for Scientific Applications", High Performance Computing and Networking (HPCN 2001), Amsterdam, Holland, June 2001, pp. 72-80.
65. RFT – Reliable File Transport, http://www-unix.globus.org/toolkit/docs/3.2/rft/
66. RLS – Replica Location Service, http://www.globus.org/rls/
67. ROADnet - Real-time Observatories, Applications, and Data management Network, http://roadnet.ucsd.edu/
68. SCEC - Southern California Earthquake Center, http://epicenter.usc.edu/cmeportal/index.html
69. Shibboleth authentication, http://shibboleth.internet2.edu/
70. Smorul, M., J. JaJa, F. McCall, S. F. Brown, R. Moore, R. Marciano, S-Y. Chen, R. Lopez, R. Chadduck, “Recovery of a Digital Image Collection Through the SDSC/UMD/NARA Prototype Persistent Archive,” SDSC Technical Report 2003-06, September 2003.
71. SRB home web page, http://www.sdsc.edu/srb/
72. SRM – Storage Resource Manager, http://sdm.lbl.gov/srm-wg/
73. UCSD Libraries image digital library, http://gort.ucsd.edu/dlp/program.htm
74. United Kingdom e-Science Data Grid, http://www.e-science.clrc.ac.uk/web/projects/storage_resource_broker
75. Watry, P., R. Larson, “Cheshire 3 Framework White Paper: Implementing Support for Digital Repositgories in a Data Grid Environment”, Proceedings of IEEE International Symposium on Globally Distributed Data, Sardinia, June, 2005, http://cheshire.liv.ac.uk/
76. W3C – WorldWide Web Consortium, http://www.w3.org/MarkUp/
77. WSRF – Web Services Resource Framework, http://www.globus.org/wsrf/

Appendix A: Workshop Attendees

The name of each attendee is given, along with their home institution, the area of data management expertise, whether they are already a user of SRB data grid technology, and a project represented at the workshop.

Attendee	Institution	Area	SRB User	Project
Stephen Abrams	Harvard	Digital Library		NDIIPP
Martha Anderson	Library of Congress	Preservation		NDIIPP
Lucy Barber	NHPRC	Preservation		
Linda Barnhart	UCSD Libraries	Digital Library	Yes	
Fran Berman	SDSC	Preservation	Yes	Chronopollis
Peter Berrisford	Rutherford Laboratory	Data Grid	Yes	e-Science data grid
Leesa Brieger	SDSC	Data Grid	Yes	NVO
Robert Chadduck	NARA	Preservation	Yes	Persistent Archive
Sheau-Yen Chen	SDSC	Data Grid	Yes	Persistent Archive
Mark Conrad	NARA	Preservation	Yes	Persistent Archive
Charles Cowart	SDSC	Data Grid	Yes	NSDL
Antoine De Torcy	SDSC	Data Grid	Yes	Persistent Archive
Luc Declerck	UCSD Libraries	Digital Library	Yes	DigArch
Tim DiLauro	JHU	Digital Library		NDIIPP
Max Evans	NHPRC	Preservation		PAT
Declan Fleming	UCSD Libraries	Digital Library	Yes	
James French	NSF			
Chris Frymann	UCSD Libraries	Digital Library	Yes	DigArch
Lucas Gilbert	SDSC	Data Grid	Yes	BIRN
Margaret Hedstrom	U. Michigan	Preservation		
Robert Horton	Minnesota Historical Soc.	Preservation	Yes	PAT
Arwen Hutt	UCSD Libraries	Digital Library	Yes	
Yoshimi Iida	KEK	Data Grid	Yes	BELLE
Kohki Isikawa	KEK	Data Grid	Yes	BELLE
Arun Jagatheesan	SDSC	Data Grid	Yes	LSST
Joseph JaJa	U Maryland	Preservation	Yes	Persistent Archive
Mark James	UCSD	Data Grid	Yes	BIRN
Keith Johnson	Stanford	Preservation		NDIIPP
Kerstin Kleese van Dam	Daresbury Laboratory	Data Grid	Yes	e-Science data grid
Ardys Kozbial	UCSD Libraries	Digital Library	Yes	DSpace
Harry Kreisler	UCB	Preservation	Yes	DigArch
George Kremenek	SDSC	Data Grid	Yes	Teragrid
Wilko Kroeger	SLAC	Preservation	Yes	PAT
John Kunze	UCOP	Preservation	Yes	CDL/NDIIPP
Jim Kupsch	U. Wisconsin			
Carl Lagoze	Cornell University	Digital Library	Yes	Fedora
William LeFurgy	Library of Congress	Preservation		NDIIPP
Sifang Lu	SDSC	Data Grid	Yes	ROADnet
Philip Maechling	USC	Digital Library	Yes	SCEC
Martha Maiden	NASA			
Fillia Makedon	NSF			

Richard Marciano	SDSC	Preservation	Yes	PAT
Mike McGann	U Maryland	Preservation	Yes	Persistent Archive
Stephen McMahon	ANU	Data Grid	Yes	APAC
Don Middleton	NCAR	Data Grid	Yes	Chronopolis
Glen Moloney	U. Melbourne	Data Grid	Yes	BELLE
Gabriela Montoya	UCSD Libraries	Digital Library	Yes	
Reagan Moore	SDSC	Data Grid	Yes	Persistent Archive
Richard Moore	SDSC	Data Grid	Yes	Teragrid
Jean-Yves Nief	IN2P3	Data Grid	Yes	BaBar
Mike Norman	UCSD	Data Grid	Yes	ENZO
Roman Olschanowsky	SDSC	Data Grid	Yes	Teragrid
Michael Pagels	DARPA			
Bernard Pailthorp	U. Queensland	Data Grid	Yes	APAC
Arcot Rajasekar	SDSC	Data Grid	Yes	Persistent Archive
Trish Rose	UCSD Libraries	Digital Library	Yes	
Chris Rusbridge	Edinburgh University	Preservation		DCC
Takeshi Sasaki	KEK	Data Grid	Yes	BELLE
Brian Schottlaender	UCSD Libraries	Digital Library	Yes	Chronopolis
Wayne Schroeder	SDSC	Data Grid	Yes	BaBar
Kenneth Sharp	Stanford	Data Grid	Yes	SLAC
MacKenzie Smith	MIT	Digital Library	Yes	DSpace
Mike Smorul	U Maryland	Preservation	Yes	Persistent Archive
Andreas Stanescu	OCLC	Preservation		
David Valentine	SDSC	Data Grid	Yes	GEON
Frank Vernon	UCSD/SIO	Data Grid	Yes	ROADnet
Mike Wan	SDSC	Data Grid	Yes	Persistent Archive
Tim Warnock	SDSC	Data Grid	Yes	NEESgrid
Brad Westbrook	UCSD Libraries	Digital Library	Yes	DigArch
Eric Yen	Academia Sinica	Data Grid	Yes	
Bing Zhu	SDSC	Data Grid	Yes	NSDL

Appendix B: Projects currently funding SRB development or application

Funding Agency	Project	Application
NARA	Research Prototype Persistent Archives	Persistent Archive
NARA	Policy Enforcement in Data Grid Environments	Persistent Archive
NHPRC	Persistent Archives Testbed	Persistent Archive
NHPRC	E-legislature Project	Persistent Archive
NHPRC	California Geospatial Records Preservation	Persistent Archive
NSF	NSDL – National Science Digital Library	Persistent Archive
NSF/LC	Digital Preservation Lifecycle Management	Persistent Archive
LC/CDL	Digital Preservation Repository	Persistent Archive
U British Columbia	GIS preservation – Vancouver VanMap	Persistent Archive
DOE	Fusion Portals	Data Grid
DOE	Particle Physics Data Grid	Data Grid
NSF	National Virtual Observatory	Data Grid
NSF	Real-time Observatories, Applications, and Data management Network	Data Grid
NSF	TeraGrid	Data Grid
NSF	ITR – Constraint-based Knowledge Systems	Data Grid
NIH	BIRN – Bio-medical Informatics Research Network	Data Grid
UCOP/LLNL	Scientific Data Management	Data Grid
NSF	Southern California Earthquake Center	Digital Library
NSF	Partnership for Biodiversity Informatics	Digital Library

Appendix C: SRB developers

More than 22 international collaborators have contributed to the development of the Storage Resource Broker technology. The dominant contribution has been the porting of additional access mechanisms to enable specific applications to access SRB shared collections. The second most common contribution has been the port of the SRB technology to support use of additional database systems or storage systems. The contributed software is available at http://www.sdsc.edu/srb/

Site/Contributor	Technology	Contributed Software
Ohio State University / Mario Lauria	Client	MPI-IO port
KISTI, Korea / Oh-kyoung Kwon	Client	SRB enabled globus-url-copy
U. Bristol, UK / Simon Metson	Client	GMCat Grid Giggle interface
U. Maryland / Mike Smorul	Client	Perl load library
Poland University / Michal Wronski	Client	Perl load library
Johns Hopkins University / Anthony Kolasny	Client	Upload/Download utilities
BIRN / Tim Warnock	Client	Linux Userland File System
BIRN / Tim Warnock	Client	Bulk operations
BIRN / Tim Warnock	Client	SRB bash source file
U. Iowa / Karen Pease	Client	Unix I/O library
Halcyon Systems / Jose Zero	Client	OpenDAP/DODS
UCSD Libraries / David Little	Client	DSpace port
ANL / John Bresnahan	Client	GridFTP
NCHC, Taiwan / Barz Hsu	Client	FUSE mountable file system
Ohio State University / Mario Lauria	Driver	Windows driver
U. Maryland / Mike Smorul, Mike McGann	Driver	Informix database driver
NCSA / Randy Sharpe	Driver	ObjectStore driver
Aerospace Corporation / Craig Lee	Driver	PostgreSQL database driver
Ohio State University / Joel Saltz	Driver	DataCutter remote procedures
NCSA / Peter Cao	Driver	HDF5 remote procedures
UK e-Science data grid / Ananta Manandhar	Driver	GSI certificate delegation
UK e-Science data grid / Michael Doherty	Documentation	Installation manual
BIRN / Roman Olschanowsky	Administration	Audit trail reports
UK e-Science data grid / Adil Hasan	Administration	Python test scripts
Academia Sinica / Huimin Lin	Administration	System Report Generator

The SRB development team at the San Diego Supercomputer Center supports the productization of the Storage Resource Broker. New features are added in response to requirements from both the funding projects and the projects applying the SRB technology.

SDSC staff member	**SRB support**
Reagan Moore	PI
Michael Wan	SRB architect
Arcot Rajasekar	SRB manager, information architect
Wayne Schroeder	SRB productization, security architect
Charlie Cowart	INQ client, NSDL persistent archive
Lucas Gilbert	Jargon java client, DSpace/Fedora integration
Bing Zhu	Perl, Python, Windows load libraries
Antoine de Torcy	MySRB web browser, NARA collections
Sheau-Yen Chen	SRB administration
George Kremenek	SRB collections, Teragrid
Arun Jagatheesan	WSDL services, Matrix workflow
Leesa Brieger	HyperAtlas, NVO services
Sifang Lu	OpenDAP client, ROADnet application
Richard Marciano	SALT persistent archives
Chien-yi Hou	Preservation workflows
Students	Bug fixes

Appendix D: Sites that have downloaded SRB source

Representative sites within the United States out of 86 sites that downloaded SRB:

Project	Institution
Database and Information Systems Laboratory	University of California Davis
Chemistry/Biochemistry	University of California Los Angeles
Consortium of Universities for the Advancement of Hydrol Science, Inc., Digital Library Server	University of California Merced
Computer Science & Engineering	University of California San Diego
ITR - constraint based data management, Computer Scien Department	University of California San Diego
Marine Physical Laboratory	University of California San Diego
National Center for Microscopy and Imaging	University of California San Diego
Cosmology, Physics Department	University of California San Diego
National Center for Microscopy and Imaging, TeleScience	University of California San Diego
University of Florida Research Grid (HPS)	University of Florida
Bioinformatics	University of Kansas
Department of Computer Science	University of Maryland
Network for Earthquake Engineering Simulation	University of Minnesota
Library archive	University of Pittsburgh
Rapid Unified Generation of Urban Databases (RUGUD)	US Army Research Activity
P2Tools Design & Development Team Leader	US Environmental Protection Agency
EPA Data Grid initiative	US Environmental Protection Agency
Government Agency	US Navy
Oceanography collections	Woods Hole Oceanographic Institute

Project	Institution
National Virtual Observatory	Caltech
Cooperative Institute for Research in Environmental Sciences /Center for Integrated Space Weather Mode	Colorado University
Institute for Astronomy	Hawaii University
Common Instrument Middleware Architecture, Natio Middleware Initiative	Indiana University
Indiana University Cyclotron Facility	Indiana University
Dspace digital library	MIT
Atmospheric Sciences Data	NASA
NOAO data grid	National Optical Astronomy Observatory
Web-at-Risk National Digital Information Infrastruct and Preservation Program (CDL)	New York University Libraries
MPI-IO interface	Ohio State Univiversity
Computer Science	Oregon State University
BioPilot	Pacific Northwest National Laboratory
TeraGrid project	Purdue University
Fusion Portal	San Diego State University
SDSC Production SRB system	San Diego Supercomputer Center
Texas Advanced Computing Center	Texas University
Network for Earthquake Engineering Simulation	Texas University
NCAR Visualization	UCAR
Network for Earthquake Engineering Simulation	University at Buffalo

Representative International Sites, out of 84 international sites that downloaded SRB:

Project	Institution
Data mangement project	British Antarctic Survey, UK
eMinerals	Cambridge e-Science Center, UK
Sickkids Hospital in Toronto	Canada
Welsh e-Science Centre	Cardiff University, UK
Visualization in scientific computing	Chinese Academy of Science, China
Australian Partnership for Advanced Computing Data Grid	Commonwealth Scientific and Industrial Research Organization, Australia
Consorzio Interuniversitario per il Calcolo Automatico d Nord Orientale, HPC-EUROPA project	Italy
Center for Advanced Studies, Research, and Development	Italy
LIACS(Leiden Inst. Of Comp. Sci)	Leiden University,The Netherlands
Australian Partnership for Advanced Computing Data Grid	Melbourne, Australia
Monash E-Research Grid	Monash University, Australia
Computational Materials Science	Nanyang Technological University, China
Virtual Tissue Bank	Osaka University, Japan
Cybermedia Center	Osaka University, Japan
Belfast e-Science Centre	Queen's University, UK
Information Technology Department	Sejong University, South Korea
Nanyang Centre for Supercomputing	Singapore
National University (Biology data grid)	Singapore
Swiss Federal Institute (Ecole Polytechnique Federale de Lausanne)	Switzerland

Project	Institution
CERN- GridFTP	Switzerland
Protein structure prediction	Taiwan University, Taiwan
Trinity College High Performance Computing (HPC-Europa)	Trinity College, Ireland
National Environment Research Council	United Kingdom
Universidad Nacionale Autonoma de Mexico Grid	Universidad Nacionale Autonoma de Mexico
Parallab(HPC-EUROPA project)	University of Bergen, Norway
Physics Labs	University of Bristol, UK
Laboratory for Bioimages and Bioengineering	University of Genoa, Italy
Bio Lab	University of Genoa,Italy
School Computing	University of Leeds, UK
Dept. of Computer Science	University of Liverpool, UK
Worldwide Universities Network	University of Manchester, UK
Large Hadron Collider Computing Grid	University of Oxford, UK
Computational Modelling	University of Queensland, Australia
Instituto do Coracao	University of Sao Paulo,Brazil
White Rose Grid	University of Sheffield. UK
Australian Partnership for Advanced Computing Data Grid	University of Technology, Australia
Computational Chemistry environment	University of Z rich, Switzerland
Australian Partnership for Advanced Computing Data Grid	Victoria, Australia

Appendix E: Storage Resource Broker Capabilities

Storage Resource Broker logical name spaces, global data manipulation operations, and global state information for the functional areas of trust virtualization and data virtualization.

	Logical naming	Standard operations	State information
Trust Virtualization	Logical user names	Add or delete user	User:Group:Zone
		GSI authentication	Certificate authority location
		Challenge-response authentication	Encrypted user password
		Issue ticket-based authentication	Time to live and number of allowed accesses
	User roles	List user roles	Curate, audit, annotate, read, write, group administration, superuser, public
		Set access control by role for user	Access controls on users
	Group names	Set access control by role for group	Access controls on groups
		Set access control on metadata for user	Access controls on metadata
		Set access control on resource for user	Access controls on resources
		Turn on audit trails	Audit trails
		Enable client-based encryption	Encryption key
		Resolve error number	System log of all accesses
Data Virtualization	Logical entity names	Define SRB physical file name structure	SRB physical file pathname structure
		Load a file into SRB collection (Sput)	Physical location where SRB stores file
		Unload a file from a SRB collection (Sget)	
	Shadow links	Register existence of external file	Location of external file
		Register existence of external directory	Location of external directory
	Logical container names	Create container	Physical file in which data is aggregated
		Create checksum	Checksum
		Verify checksum	
		Synchronize replicas	Dirty bit for writes
		Synchronize remote files with SRB files	
		Synchronize SRB files with remote files	
		Synchronize SRB files between two SRB collections	
		Posix I/O - partial read and write	Replica location
		Delete file	
		Recursive directory registration	
		Register a file as a replica of existing file	Owner, size
		Create version	Version number
		Create backup	Backup time
		Lock a file	Lock status
		Register SQL command	Data type
		Issue a registered SQL command	
		Create and issue a Datascope query	
		Register URL	

Storage Resource Broker logical name spaces, global data manipulation operations, and global state information for the functional areas of latency management, collection management and federation management.

	Logical naming	Standard operations	State information
Latency Management	Logical resource names	Load leveling	Quotas on storage and usage of storage
		Fault tolerant replication	Replication state
	Compound resources	File staging	Names for file system cache
		Automated access control setting	Sticky bits to inherit access controls of parent collection
		Client and server initiated parallel I/O on access	Creation time, update time
		Client and server initiated bulk file registration	
		Client and server initiated remote procedures	Location in SRB of remote procedures
		Client and server initiated bulk metadata load	
		Bulk delete - trash can	Deletion flag
		Automated checksum verification on load	
		Third party transfer	
		Store files in a logical container	
Collection Managment	Descriptive metadata	Extensible metadata	Descriptive metadata for SRB file
	Collection hierarchy	Create/delete subcollection	Parent collection identity
		Create collection metadata	Descriptive metadata for SRB collection
		Extensible schema	Table structure of metadata
		Create soft link between two logical files	Soft link
		Import of XML files	
		Export of XML and HTML files	
		Remote template-based metadata extraction	Location in SRB of templates
		Synchronize slave catalog with master catalog	Location of slave catalog
		Queries on descriptive and state information	
Federation Management	Distinguished zone names	Access zone authority to register zone name	Zone name and port number
	Zone authority name	User authentication by home zone	
		Cross-registration of resources between zones	
		Synchronization of user names between zones	
		Synchronization of file names between zones	
		Synchronization of metadata between zones	