Integration of Object-based Storage into Preservation Environments

Reagan W. Moore
San Diego Supercomputer Center
9500 Gilman Drive, La Jolla CA 92093-0505
Phone:+1-858-534-5073    FAX: +1-858-534-5152
E-mail: moore@sdsc.edu

Presented at the THIC Meeting at the Sony Auditorium,
3300 Zanker Rd, San Jose CA 95134-1940
April 19-20,2005
Preservation

• Archival processes through which a digital entity is extracted from its creation environment, and then supported in a preservation environment, while maintaining authenticity and integrity information.

• Extraction process requires insertion of support infrastructure underneath the digital material

• Goal is infrastructure independence, the ability to use any commercial storage system, database, or access mechanism

• Can Object-based Storage Devices support preservation functions?
Preservation Communities

- InterPARES - diplomats
  - Preservation of records
- NARA - life cycle management model
  - Preservation of records from federal agencies
- State archives
  - Preservation of submitted “collections”
- Australia - continuum model
  - Preservation of active data with records
InterPARES - Diplomatics

- Authenticity - maintain links to metadata for:
  - Date record is made
  - Date record is transmitted
  - Date record is received
  - Date record is set aside [i.e. filed]
  - Name of author (person or organization issuing the record)
  - Name of addressee (person or organization for whom the record is intended)
  - Name of writer (entity responsible for the articulation of the record’s content)
  - Name of originator (electronic address from which record is sent)
  - Name of recipient(s) (person or organization to whom the record is sent)
  - Name of creator (entity in whose archival fonds the record exists)
  - Name of action or matter (the activity for which the record is created)
  - Name of documentary form (e.g. E-mail, report, memo)
  - Identification of digital components
  - Identification of attachments (e.g. digital signature)
  - Archival bond (e.g. classification code)
InterPARES - Diplomatics

- Integrity - maintain links to metadata for
  - Name(s) of the handling office / officer
  - Name of office of primary responsibility for keeping the record
  - Annotations or comments
  - Actions carried out on the record
  - Technical modifications due to transformative migration
  - Validation
Support Infrastructure

- Manages the electronic records
- Manages the provenance metadata (authenticity)
- Manages the integrity metadata
- Manages the name spaces used to control the electronic records
Data Grid Support for Preservation

• Authenticity - the assurance that provenance metadata remains linked to the electronic records
  – Link authenticity metadata (descriptive metadata) to files
  – Location independent naming convention for files

• Integrity - the assurance that the electronic records are not corrupted
  – Link integrity metadata (audit trails, access controls, checksums)
  – Support for distributed environments (replication, federation)

• Infrastructure Independence
  – Standard operations across databases
  – Standard operations across storage repositories
Data Grids

• Manage shared collections that are distributed in space
  – Location of item, access controls, checksums

• Implement infrastructure independence
  – Standard operations for interacting with storage repositories

• Implement presentation independence
  – Standard APIs to support porting of user interfaces
Managing Distributed Data Name Spaces

Data Access Methods (Web Browser, Java, Perl, “C”)

Storage Repository
- Storage location
- User name
- File name
- File context (creation date,…)
- Access constraints

Naming conventions provided by storage systems
Data Grids Provide a Level of Indirection for Each Naming Convention

- Storage location
- User name
- File name
- File context (creation date, ...)
- Access constraints

Data Collection

- Logical resource name space
- Logical user name space
- Logical file name space
- Logical context (metadata)
- Control/consistency constraints

Data Access Methods (C library, Unix, Web Browser)

Data is organized as a shared collection
Authenticity

- Package authenticity metadata with the electronic record in an Archival Information Package (AIP)
  - Reference Model for an Open Archival Information System (OAIS).
  - Metadata Encoding and Transmission Standard (METS)
- Validate AIPs for conformance with preservation metadata standard METS profile
  - Standard evolves over time
Integrity

- Package integrity metadata with electronic records in an AIP
- Update AIP on every operation on the electronic record
  - Audit trail
  - Date of checksum validation
  - Transformative migrations of encoding format
Data Grid Operations

• Remote operations
  – Unix file system (open, close, seek, read, write, stat,..)
  – Bulk operations for latency management
  – Remote procedures for data filtering
  – Data transformations
  – Third party transfer

• Collective operations
  – Load leveling
  – Fault tolerance
  – Replication

The Premier Advanced Recording Technology Forum
Latency Management - Bulk Operations

- Bulk register
  - Create a logical name for a file
- Bulk load
  - Create a copy of the file on a data grid storage repository
- Bulk unload
  - Provide containers to hold small files and pointers to each file location
- Bulk delete
  - Mark as deleted in metadata catalog
  - After specified interval, delete file
- Bulk metadata load
  - Support parsing of metadata from a remote file at remote storage
- Requests for bulk operations for access control setting, ...
Operations Performed by Object-based Storage Device

• Manipulation
  – Bulk operations for metadata extraction, registration of digital entities, load and unload of data

• Integrity
  – Validation of checksums
  – Validation of AIPs
  – Updates of Archival Information Packages

• Presentation
  – Transformative migration
    • Conversion of encoding format for display
Implications

- Global properties are managed by the data grid
  - Name spaces
  - Collective operations

- OSD can support local operations
  - Metadata extraction from files
  - AIP manipulation

- Security will be distributed between the data grid and the OSD
  - Data grids manage data distributed across administrative domains
Security Management

• Within the shared collection, the digital entities are owned and managed by the data grid
  – Files, URLs, SQL commands, database binary large objects can be registered into the shared collection
  – Files are stored under an account ID representing the data grid

• Access controls are managed by the data grid
  – Files / metadata / storage systems

• Access controls are defined for multiple roles
  – Schema extension, create new metadata
  – Modify metadata
  – Add annotations
  – Turn on audit trails
  – Write data
  – Read data
Federation of Data Grids

• A data grid provides a single sign-on environment
  – The data grid manages the name space for the user

• To authenticate persons between data grids
  – Define user name to be
    Home data grid / Group / User
  – Access Home data grid for authentication
Federation Between Enterprises

Data Access Methods (Web Browser, DSpace, OAI-PMH)

Data Collection A

Data Grid
- Logical resource name space
- Logical user name space
- Logical file name space
- Logical context (metadata)
- Control/consistency constraints

Data Collection B

Data Grid
- Logical resource name space
- Logical user name space
- Logical file name space
- Logical context (metadata)
- Control/consistency constraints

Access controls and consistency constraints on cross registration of digital entities

The Premier Advanced Recording Technology Forum
Preservation Environments (Based on Storage Resource Broker)

- NARA research prototype persistent archive
- NHPRC Persistent Archive Testbed
- NSF National Science Digital Library persistent archive
- University of California - Digital Preservation Repository
- UCSD image archive
- MIT DSpace/SRB preservation environment
## Storage Resource Broker Collections at SDSC

(4/18/2005)

<table>
<thead>
<tr>
<th>Collection</th>
<th>GBs of data stored</th>
<th>Number of files</th>
<th>Number of Users</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Grid</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSF/ITR - National Virtual Observatory</td>
<td>53,862</td>
<td>9,536,751</td>
<td>100</td>
</tr>
<tr>
<td>NSF - National Partnership for Advanced Computational Infrastructure</td>
<td>33,196</td>
<td>6,878,235</td>
<td>380</td>
</tr>
<tr>
<td>Static collections Æ Hayden planetarium</td>
<td>8,013</td>
<td>161,352</td>
<td>227</td>
</tr>
<tr>
<td>Pzone Æ public collections</td>
<td>7,511</td>
<td>3,918,644</td>
<td>67</td>
</tr>
<tr>
<td>NSF/NPACI - Biology and Environmental collections</td>
<td>22,179</td>
<td>54,695</td>
<td>67</td>
</tr>
<tr>
<td>NSF/NPACI Æ Joint Center for Structural Genomics</td>
<td>6,785</td>
<td>913,430</td>
<td>50</td>
</tr>
<tr>
<td>NSF - TeraGrid, ENZO Cosmology simulations</td>
<td>165,470</td>
<td>1,360,795</td>
<td>3,267</td>
</tr>
<tr>
<td>NIH - Biomedical Informatics Research Network</td>
<td>10,674</td>
<td>7,168,846</td>
<td>268</td>
</tr>
<tr>
<td><strong>Digital Library</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSF/NPACI - Long Term Ecological Reserve</td>
<td>256</td>
<td>9,033</td>
<td>36</td>
</tr>
<tr>
<td>NSF/NPACI - Grid Portal</td>
<td>2,620</td>
<td>53,048</td>
<td>460</td>
</tr>
<tr>
<td>NIH - Alliance for Cell Signaling microarray data</td>
<td>559</td>
<td>71,318</td>
<td>21</td>
</tr>
<tr>
<td>NSF - National Science Digital Library SIO Explorer collection</td>
<td>2,655</td>
<td>1,052,550</td>
<td>27</td>
</tr>
<tr>
<td>NSF/ITR - Southern California Earthquake Center</td>
<td>107,470</td>
<td>2,304,282</td>
<td>64</td>
</tr>
<tr>
<td><strong>Persistent Archive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NHPRC Persistent Archive Testbed (Kentucky, Ohio, Michigan, Minnesota)</td>
<td>96</td>
<td>378,806</td>
<td>28</td>
</tr>
<tr>
<td>UCSD Libraries archive</td>
<td>4,147</td>
<td>408,050</td>
<td>29</td>
</tr>
<tr>
<td>NARA- Research Prototype Persistent Archive</td>
<td>1,449</td>
<td>883,982</td>
<td>58</td>
</tr>
<tr>
<td>NSF - National Science Digital Library persistent archive</td>
<td>3,572</td>
<td>26,931,909</td>
<td>136</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>430 TB</td>
<td>62 million</td>
<td>5,285</td>
</tr>
</tbody>
</table>
Scalability

• Major challenge is the large number of files
  – Databases scale to a billion records
  – File systems scale to 20-50 million files

• Use containers to aggregate data before storage
  – Minimizes the number of files seen by the storage system

• Distribute files across storage systems
  – Load leveling

• Distribute files across data grids
Conclusion

• Object-based storage devices can support operations essential to preservation environments
  – Association of authenticity and integrity metadata with each file
  – Operations on the metadata (update, validation)
  – Integrity checking of files
  – Operations on the files (metadata extraction, transformative migration)
SDSC SRB Team (left to right)

- Arun Jagatheesan
- George Kremenek
- Sheau-Yen Chen
- Arcot Rajasekar (SRB development lead)
- Reagan Moore (SRB PI)
- Michael Wan (SRB architect)
- Roman Olschanowsky (BIRN)
- Bing Zhu
- Charlie Cowart
- Lucas Gilbert
- Tim Warnock
- Wayne Schroeder (SRB product)
- Adam Birnbaum (SRB production)
- Antoine De Torcy
- Vicky Rowley (BIRN)
- Marcio Faerman (SCEC)
- Students & emeritus
  - Erik Vandekieft
  - Reena Mathew
  - Xi (Cynthia) Sheng
  - Allen Ding
  - Grace Lin
  - Qiao Xin
  - Daniel Moore
  - Ethan Chen
  - Jon Weinburg

Premier Advanced Recording Technology Forum
For More Information

Reagan W. Moore
San Diego Supercomputer Center

moore@sdsc.edu

http://www.sdsc.edu/srb/