Data Grids and Data Management

Bing Zhu
bzhu@sdsc.edu
http://www.sdsc.edu/srb
Topics

• Architecture of the system
  • How we connect to SDSC, RO, MHS and replicate and transfer files)

• Storage options (offline, online)
  • Ramifications in terms of disaster recovery

• Demonstration of the different interfaces
  • Scommands, inQ, MySRB

• Resources available on your website
  • SRB wiki, srb-chat, presentations, manuals

• Response to questions
Using a Data Grid – in Abstract

- User asks for data from the data grid
- The data is found and returned
  - Where & how details are hidden
• User asks for data
• Data request goes to SRB Server
• Server looks up data in catalog
• Catalog tells which SRB server has data
• 1st server asks 2nd for data
• The data is found and returned
Using a Data Grid - Details

- Data Grid has arbitrary number of servers
- Complexity is hidden from users
Shared Collections

- **Data grids** support the creation of shared collections that may be distributed across multiple institutions, sites, and storage systems.

- **Digital libraries** publish data, and provide services for
  - Curation
  - Discovery
  - Presentation

- **Persistent archives** preserve data, managing the migration to new technology
  - Manage authenticity and integrity
Persistent Archives

- Authenticity metadata for each record
  - Manage provenance metadata for source of record
  - Tracking chain of custody
- Integrity metadata for each record
  - Checksums
  - Replicas
  - Synchronization flags
  - Access controls
  - Versions
  - Audit trail
<table>
<thead>
<tr>
<th>Project</th>
<th>GBs of data stored</th>
<th>1000وثs of files</th>
<th>GBs of data stored</th>
<th>1000وثs of files</th>
<th>Users with ACLs</th>
<th>GBs of data stored</th>
<th>1000وثs of files</th>
<th>Users with ACLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Grid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSF / NVO</td>
<td>17,800</td>
<td>5,139</td>
<td>51,380</td>
<td>8,690</td>
<td>80</td>
<td>100,990</td>
<td>13,217</td>
<td>100</td>
</tr>
<tr>
<td>NSF / NPACI</td>
<td>1,972</td>
<td>1,083</td>
<td>17,578</td>
<td>4,694</td>
<td>380</td>
<td>34,830</td>
<td>7,239</td>
<td>380</td>
</tr>
<tr>
<td>Hayden</td>
<td>6,800</td>
<td>41</td>
<td>7,201</td>
<td>113</td>
<td>178</td>
<td>8,013</td>
<td>161</td>
<td>227</td>
</tr>
<tr>
<td>Pzone</td>
<td>438</td>
<td>31</td>
<td>812</td>
<td>47</td>
<td>49</td>
<td>23,099</td>
<td>13,287</td>
<td>68</td>
</tr>
<tr>
<td>NSF / LDAS-SALK</td>
<td>239</td>
<td>1</td>
<td>4,562</td>
<td>16</td>
<td>66</td>
<td>115,178</td>
<td>146</td>
<td>67</td>
</tr>
<tr>
<td>NSF / SLAC-JCSG</td>
<td>514</td>
<td>77</td>
<td>4,317</td>
<td>563</td>
<td>47</td>
<td>17,095</td>
<td>1,775</td>
<td>55</td>
</tr>
<tr>
<td>NSF / TeraGrid</td>
<td>80,354</td>
<td>685</td>
<td>2,962</td>
<td>202,226</td>
<td>4,443</td>
<td>3,267</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIH / BIRN</td>
<td>5,416</td>
<td>3,366</td>
<td>148</td>
<td>16,288</td>
<td>15,306</td>
<td>361</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digital Library</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSF / LTER</td>
<td>158</td>
<td>3</td>
<td>233</td>
<td>6</td>
<td>35</td>
<td>236</td>
<td>34</td>
<td>36</td>
</tr>
<tr>
<td>NSF / Portal</td>
<td>33</td>
<td>5</td>
<td>1,745</td>
<td>48</td>
<td>384</td>
<td>2,620</td>
<td>53</td>
<td>460</td>
</tr>
<tr>
<td>NIH / AfCS</td>
<td>27</td>
<td>4</td>
<td>462</td>
<td>49</td>
<td>21</td>
<td>733</td>
<td>94</td>
<td>21</td>
</tr>
<tr>
<td>NSF / SIO Explorer</td>
<td>19</td>
<td>1</td>
<td>1,734</td>
<td>601</td>
<td>27</td>
<td>2,605</td>
<td>1,121</td>
<td>27</td>
</tr>
<tr>
<td>NSF / SCEC</td>
<td>15,246</td>
<td>1,737</td>
<td>52</td>
<td>167,140</td>
<td>3,471</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistent Archive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NARA</td>
<td>7</td>
<td>2</td>
<td>63</td>
<td>81</td>
<td>58</td>
<td>2,916</td>
<td>2,004</td>
<td>58</td>
</tr>
<tr>
<td>NSF / NSDL</td>
<td>2,785</td>
<td>20,054</td>
<td>119</td>
<td>5,653</td>
<td>50,600</td>
<td>136</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCSD Libraries</td>
<td>127</td>
<td>202</td>
<td>29</td>
<td>190</td>
<td>208</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NHPRC / PAT</td>
<td>1,336</td>
<td>519</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>28 TB</td>
<td>6 mil</td>
<td>194 TB</td>
<td>40 mil</td>
<td>4,635</td>
<td>701 TB</td>
<td>113 mil</td>
<td>5,393</td>
</tr>
</tbody>
</table>
Shared Collections

• Purpose of SRB data grid **software** is to enable the creation of a collection that is shared between academic institutions
  • Register digital entity into the shared collection
  • Assign owner, access controls
  • Assign descriptive, provenance metadata
  • Manage state information
    • Audit trails, versions, replicas, backups, locks
    • Size, checksum, validation date, synchronization date, …
  • Manage interactions with storage systems
    • Unix file systems, Windows file systems, tape archives, …
  • Manage interactions with preferred access mechanisms
    • Web browser, Java, WSDL, C library, …
• Digital libraries now build upon data grids to manage distributed collections
  • DSpace digital library - MIT and Hewlett Packard
  • Fedora digital library - Cornell University and University of Virginia

• Persistent archives build upon data grids to manage technology evolution
  • NARA research prototype persistent archive
  • California Digital Library - Digital Preservation Repository
  • NSF National Science Digital Library persistent archive
• URLs for educational material for all grade levels registered into repository at Cornell
• SDSC crawls the URLs, registers the web pages into a SRB data grid, builds a persistent archive
  • 750,000 URLs
  • 13 million web pages
  • About 3 TBs of data
Control of water pollution from agriculture - FAO irrigation and drainage paper 55

Table of Contents
Demonstrate preservation environment
- Authenticity
- Integrity
- Management of technology evolution
- Mitigation of risk of data loss
  - Replication of data
  - Federation of catalogs
- Management of preservation metadata
  - LCDRG metadata hierarchy
- Scalability
  - Types of data collections
  - Size of data collections

Federation of Three Independent Data Grids

Original data at NARA, data replicated to U Md & SDSC
Replicated copy at U Md for improved access, load balancing and disaster recovery
Deep archive at SDSC, no user access
BaBar High-energy Physics

- Stanford Linear Accelerator
- Lyon, France
- Rome, Italy
- San Diego
- RAL, UK

- A functioning international Data Grid for high-energy physics

Moved over 170 TBs of data
Astronomy Data Grid

- Chile
- Tucson, Arizona
- NCSA, Illinois

A functioning international Data Grid for Astronomy

Moved over 400,000 images
Storage Resource Broker 3.3.1

Federation Management

Consistency & Metadata Management / Authorization, Authentication, Audit

Logical Name Space  Latency Management  Data Transport  Metadata Transport

Database Abstraction  Storage Repository Abstraction

Databases - DB2, Oracle, Sybase, Postgres, mySQL, Informix  Archives - Tape, Sam-QFS, DMF, ORB  File Systems Unix, NT, Mac OS X  Databases - DB2, Oracle, Sybase, Postgres, mySQL, Informix

C Library, Java  Unix Shell  Linux I/O C++  NT Browser, Kepler Actors  DLL / Python, Perl, Windows  DSpace, OpenDAP, GridFTP, Fedora  http, Portlet, WSDL, OAI-PMH  Linux I/O C++  DLL / Python, Perl, Windows  DSpace, OpenDAP, GridFTP, Fedora  Unix Shell  Unix Shell  Unix Shell  Unix Shell  Unix Shell

C++  DLL / Python, Perl, Windows  DSpace, OpenDAP, GridFTP, Fedora  Unix Shell  Unix Shell  Unix Shell  Unix Shell  Unix Shell

DB2, Oracle, Sybase, Postgres, mySQL, Informix  Archives - Tape, Sam-QFS, DMF, ORB  File Systems Unix, NT, Mac OS X  Databases - DB2, Oracle, Sybase, Postgres, mySQL, Informix  Unix Shell  Unix Shell  Unix Shell  Unix Shell  Unix Shell
Types of Risk

- **Media failure**
  - Replicate data onto multiple media
- **Vendor specific systemic errors**
  - Replicate data onto multiple vendor products
- **Operational error**
  - Replicate data onto a second administrative domain
- **Natural disaster**
  - Replicate data to a geographically remote site
- **Malicious user**
  - Replicate data to a deep archive
How Many Replicas

• Three sites minimize risk
  • Primary site
    • Supports interactive user access to data
  • Secondary site
    • Supports interactive user access when first site is down
    • Provides 2nd media copy, located at a remote site, uses different vendor product, independent administrative procedures
  • Deep archive
    • Provides 3rd media copy, staging environment for data ingestion, no user access
Deep Archive

Deep Archive

Staging Zone

Remote Zone
Server initiated I/O

No access by Remote zones

PVN

Pull

Register

Z3: D3: U3

Z2: D2: U2

Firewall
Demonstrations - InQ

inQ a browser/query tool of SRB for Windows platform.
mySRB is a Web-based Browser and Query Tool for SRB.
Demonstrations - Scommands

- Connecting to SRB Server - Sinit
- Listing files - Sls
- Creating sub-collection - Smkdir
- Put and Get of files - Sput & Sget
- Browse on metadata - Sufmeta, Sls
- Data replication - Sreplicate
- Listing of resources - SgetR
- Federation - Sls /
SRB Web Site: http://www.sdsc.edu/srb

- SRB software downloads
- SRB Manual pages
- SRB Tutorials
- Bug Report
- Hot pages
- FAQ
- etc
Reagan Moore - PI
Richard Marciano - SALT persistent archives
Michael Wan - SRB Architect
Arcot Rajasekar - SRB Manager
Wayne Schroeder - SRB Productization
Charlie Cowart - inQ
Lucas Gilbert - Jargon
Bing Zhu - Perl, Python, Windows
Antoine de Torcy - mySRB web browser
Sheau-Yen Chen - SRB Administration
George Kremenek - SRB Collections
Arun Jagatheesan - Matrix workflow
Sifang Lu - ROADnet Application

Contributors from UK e-Science, Academia Sinica, Ohio State University, Aerospace Corporation, ...