Preservation Environments

Reagan W. Moore

San Diego Supercomputer Center

9500 Gilman Drive, MC-0505

La Jolla, CA 92093-0505

moore@sdsc.edu

NASA/IEEE MSST 2004

12th NASA Goddard/21st IEEE Conference on Mass Storage Systems & Technologies

The Inn and Conference Center

University of Maryland University College

Adelphi MD USA

April 13-16, 2004

Topics - Persistent Archives

- Persistent archive definition
- Properties of preservation environments
- How to build a persistent archive using data grids
- Implications for massive archives
  - Replication environments
  - Deep archives

Persistent Archive Definition

- Data grid community
  - Persistent Archive is the infrastructure that manages storage technology evolution while preserving a collection
- Archivist community
  - Persistent Archive is the collection that is being preserved in some choice of infrastructure
- Together they define a preservation environment

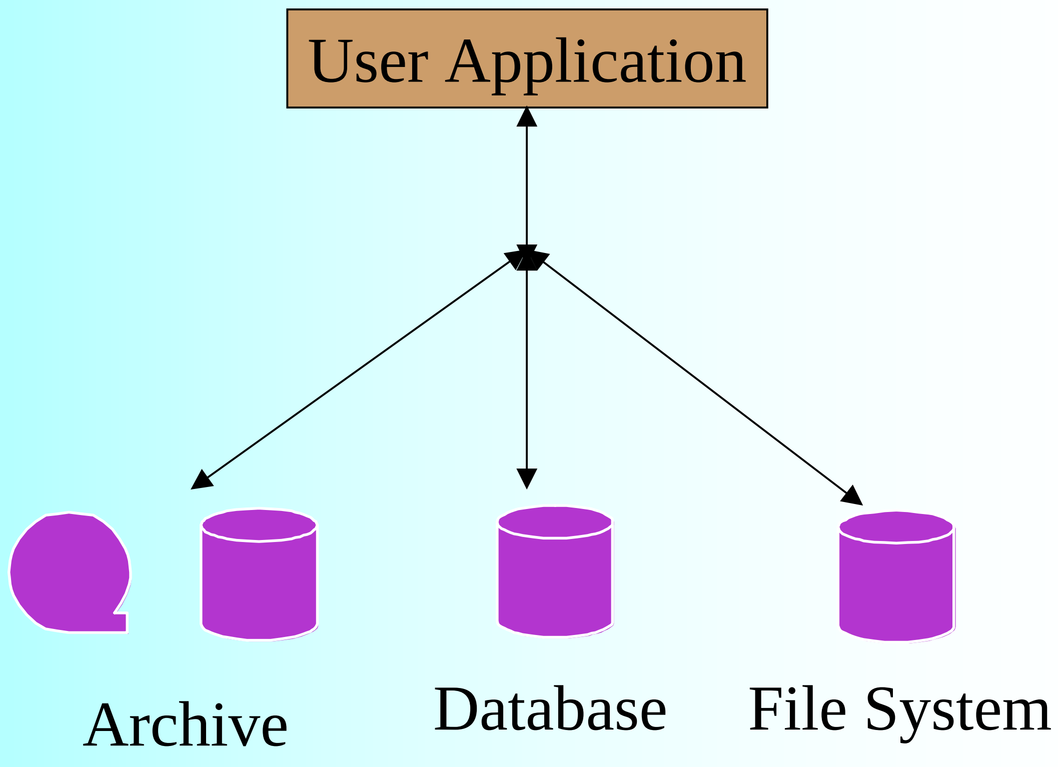
Persistent Archive Requirements

- Persistent identifier
  - Name remains unchanged when the file is moved
- Management of context
  - Preservation, administrative, integrity metadata
- Management of content
  - Containers, transformative migrations, replication
- Consistency between context and content
  - Assert that all operations on material are tracked and controlled
  - Manage technology evolution

Data Grids

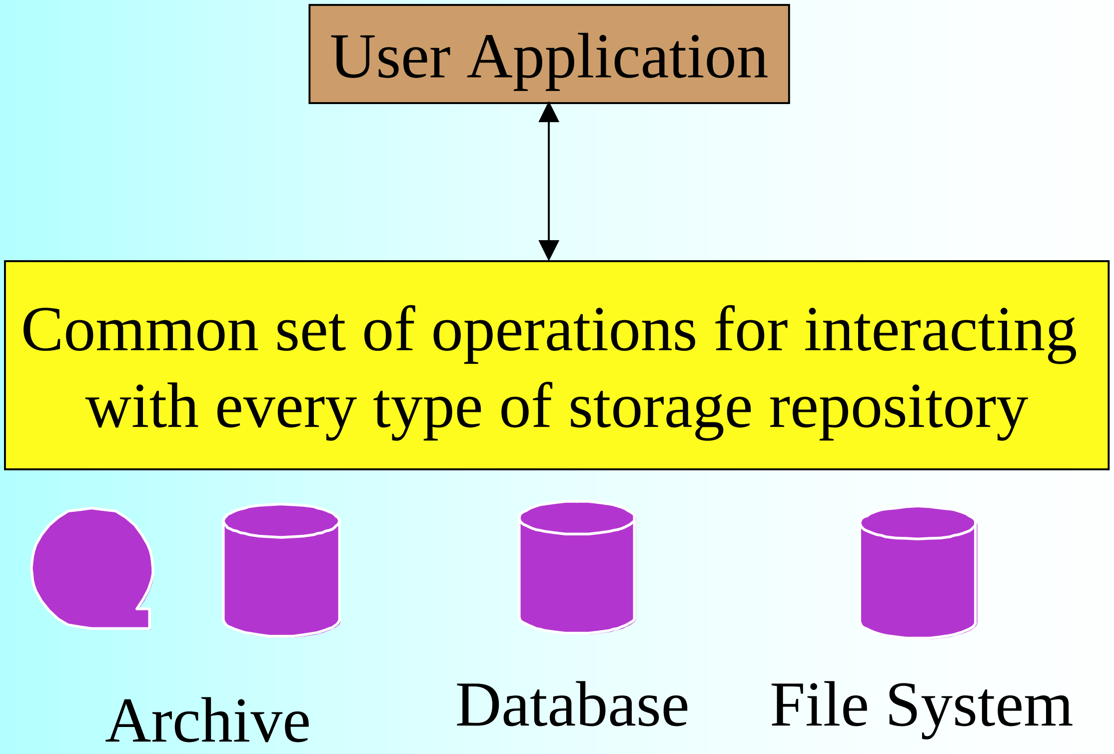
- Virtualization mechanisms to manage storage
  - Storage repository virtualization
  - Information repository virtualization
- Virtualization mechanisms to manage data
  - Data virtualization
  - Access virtualization
- Simplify management of data distributed across multiple sites and across multiple types of storage repositories

Storage Repository Virtualization



Storage Repository Virtualization

Remote operations
- Unix file system
- Latency management
- Procedures
- Transformations
- Third party transfer
- Filtering
- Queries
- Containers



Containers

- Archivists store hardcopy in “cardboard boxes”
- A container is the digital equivalent, the aggregation of digital files into a single file, with an associated “packing list”

Data Stored at SDSC

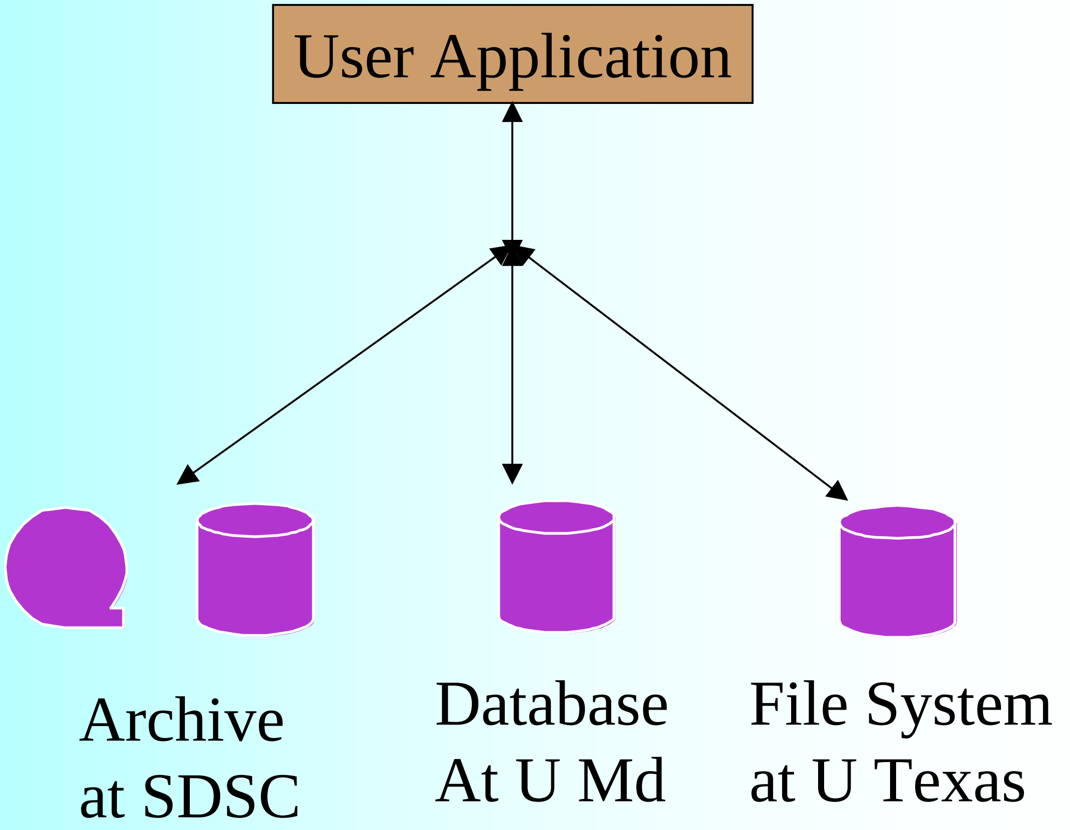
- HPSS archive
  - Stores 1 Petabyte of data
  - Stores 17 million files
- Storage Resource Broker data grid
  - Stores 114 Terabytes of data
  - Stores 31 million files
  - Containers are used to aggregate files before loading into HPSS

SRB Collections at SDSC

Project Instance	As of 12/22/2000 Data_size (in GB)	As of 12/22/2000 Count (files)	As of 5/17/2002 Data_size (in GB)	As of 5/17/2002 Count (files)	As of 3/3/2004 Data_size (in GB)	As of 3/3/2004 Count (files)	As of 3/3/2004 Users
Data Grid							
Digsky	7,599.00	3,630,300	17,800.00	5,139,249	45,939.00	8,685,572	80
NPACI	329.63	46,844	1,972.00	1,083,230	13,700.00	4,050,863	379
Hayden			6,800.00	41,391	7,835.00	60,001	168
SLAC			514.00	77,168	3,432.00	446,613	43
LDAS/SALK			239.00	1,766	2,002.00	14,427	66
TeraGrid					22,563.00	452,868	2,585
BIRN					892.00	2,472,299	160
Digital Library							
DigEmbryo	124.30	2,479	433.00	31,629	720.00	45,365	23
HyperLter	28.94	69	158.00	3,596	215.00	5,110	29
Portal			33.00	5,485	1,610.00	46,278	374
AfCS			27.00	4,007	236.00	42,987	21
NSDL/SIO Exp			19.20	383	1,217.00	193,888	26
TRA			5.80	92	92.00	2,387	26
SCEC					12,311.00	1,730,432	47
UCSDLib					127.00	202,445	29
Persistent Archive							
NARA/Collection			7.00	2,455	72.00	82,192	58
NSDL/CI					1,529.00	12,658,072	116
TOTAL	8 TB	3.7 million	28 TB	6.4 million	114 TB	31 million	4230

** Does not cover data brokered by SRB spaces administered outside SDSC.
Does not cover databases; covers only files stored in file systems and archival storage systems
Does not cover shadow-linked directories

Data Virtualization

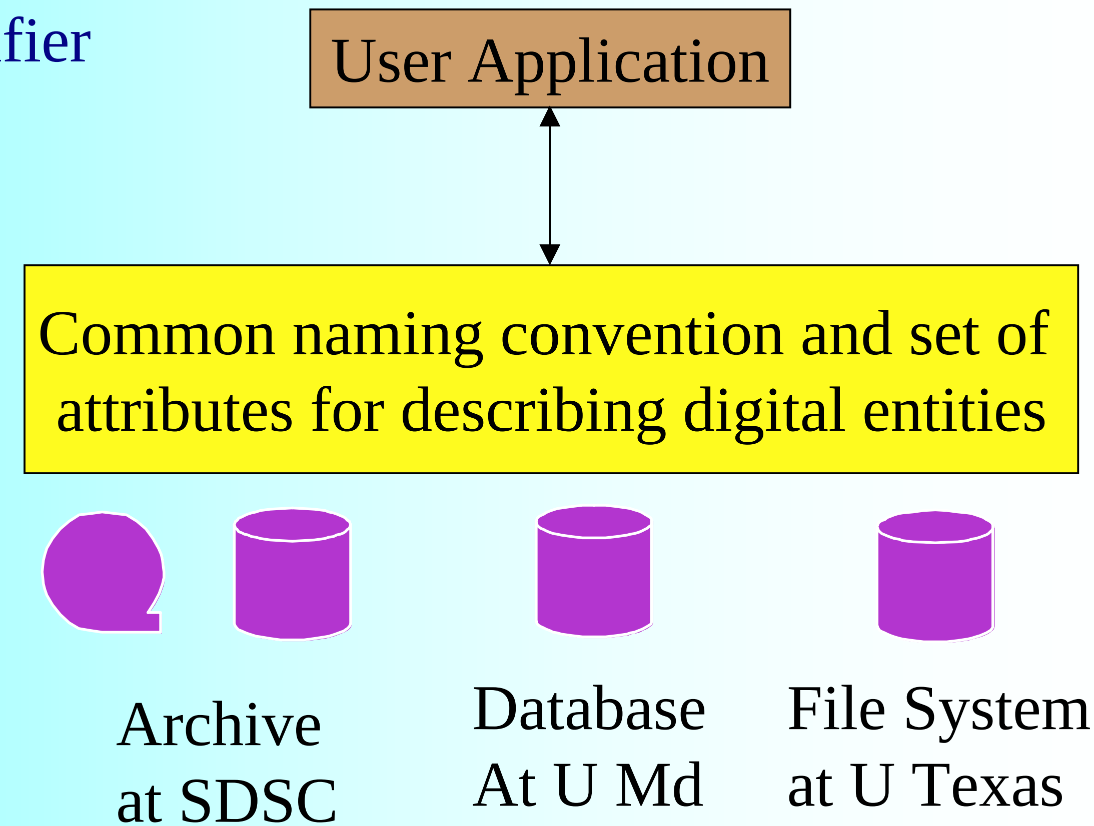


Data Virtualization

Logical name space
- Location independent identifier
- Persistent identifier

Collection owned data
- Access controls
- Audit trails
- Checksums
- Descriptive metadata

Inter-realm authentication
- Single sign-on system



Data Virtualization

- Associate information context with each digital entity
  - Organize information context as metadata in a collection
- Administrative metadata
  - Location, file name in storage, size, creation time, update time, owner, container, replica
- Descriptive metadata
  - Provenance, record series attributes, discovery attributes
- Integrity metadata
  - GUID, checksum, access controls, audit trails
- Structural metadata
  - Encoding format, component order
- Behavioral metadata
  - Operations that can be applied for presentation and manipulation

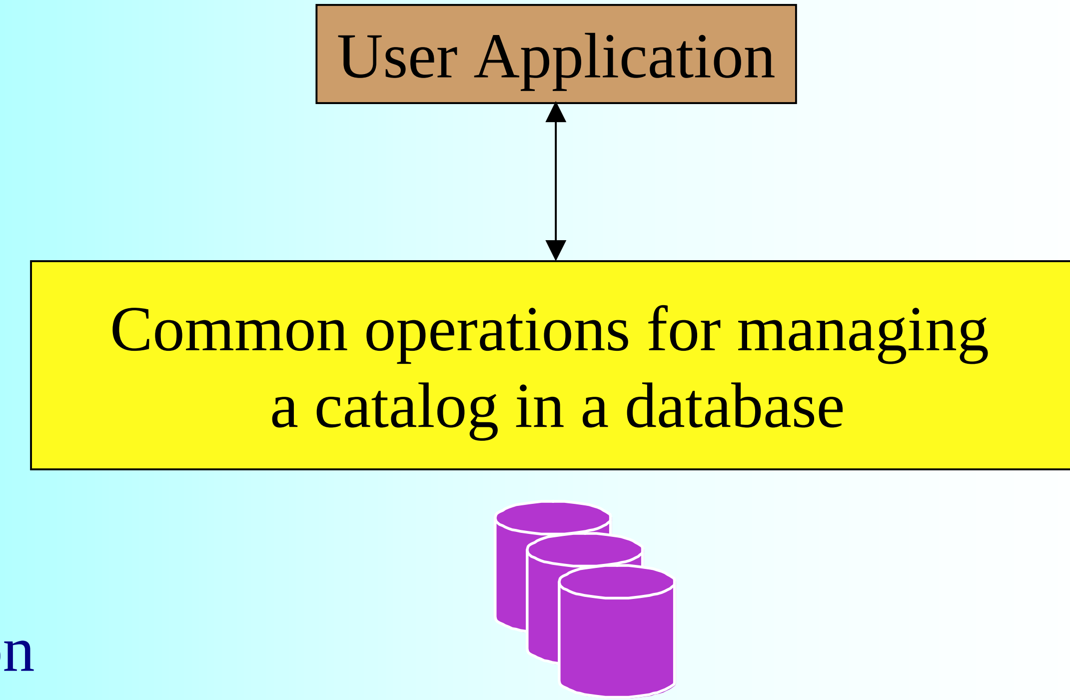
Logical Name Space for Files (Persistent Identifiers)

- Infrastructure independent naming convention for files
  - Map Global Unique Identifier to the logical file name
  - Map physical location to the logical file name
  - Map descriptive metadata to the logical file name

Information Repository Virtualization

Operations used to manage administrative, descriptive, user-defined metadata

- Import from XML file
- Export to XML file
- Bulk load
- Bulk unload
- Schema extension
- Access controls
- Dynamic SQL generation



Choice of database for Metadata Catalog

Map from API to remote operations

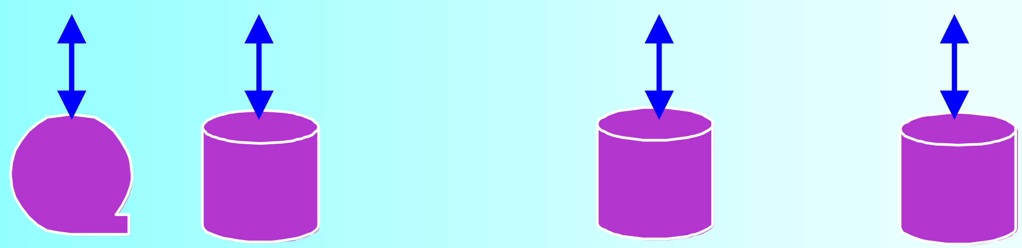
- Unix file system
- Latency management
- Procedures
- Transformations
- Third party transfer
- Filtering
- Queries

Access Virtualization

Application

C, C++, Libraries	**Linux I/O**	**Unix Shell**	**Java, NT Browsers**	**DLL / Python**	**GridFTP**	**OAI WSDL**

Common operations performed on all storage repositories



Technology Evolution

- All components of the "Persistent Archive" will evolve
  - Hardware systems
  - Software systems
  - Protocols
  - Access methods
  - Encoding syntax for digital entities
- Create drivers for each new storage repository protocol
  - Migrate data to each new storage system
- Manage evolution of the encoding syntax through either transformative migration or emulation

Are Repeated Media Migrations Feasible?

- At SDSC, cartridge capacity has increased from 200 Mbytes to 200 Gbytes for same cartridge cost
- Only migrate to new technology when the cost per Gigabyte is a factor of two lower
- Then the media cost is fixed when sum over all migrations

  $(1 + 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + \ldots) = 2$
- SDSC migrates to new media to reduce cost
  - All tape are stored in robots to minize labor costs

Transformative Migration versus Emulation

- Transform the encoding format to a new standard
  - Read the digital entity and apply transformations
  - Can combine encoding format transformation with media migration
- Emulation
  - Create a transportable parser for the original encoding format
  - Example - Multivalent Browser (written in Java) for parsing pdf

Persistent Archives

- When migrate from an old technology to a new technology, both versions are available.
- Virtualization mechanisms used for federation across space can be used to manage migration over time
- Persistent archives can be built on data grid infrastructure

Automation of Archival Processes

Archival Process	Functionality
Appraisal	Assessment of digital entities
Accession	Import of digital entities
Description	Assignment of preservation metadata
Arrangement	Logical organization of digital entities
Preservation	Long-term storage
Access	Discovery and retrieval

Data Grid Core Capabilities and Functionality

Storage repository abstraction
Storage interface to at least one repository
Standard data access mechanism
Standard data movement protocol support
Containers for data
Logical name space
Registration of files in logical name space
Retrieval by logical name
Logical name space structural independence from physical file
Persistent handle

Information Repository Abstraction

Collection owned data
Collection hierarchy for organizing logical name space
Standard metadata attributes (controlled vocabulary)
Attribute creation and deletion
Scalable metadata insertion
Access control lists for logical name space
Attributes for mapping from logical file name to physical file
Encoding format specification attributes
Data referenced by catalog query
Containers for metadata

Distributed Resilient Architecture

Specification of system availability
Standard error messages
Status checking
Authentication mechanism
Specification of reliability against permanent data loss
Specification of mechanism to validate integrity of data
Specification of mechanism to assure integrity of data

Virtual Data Grid

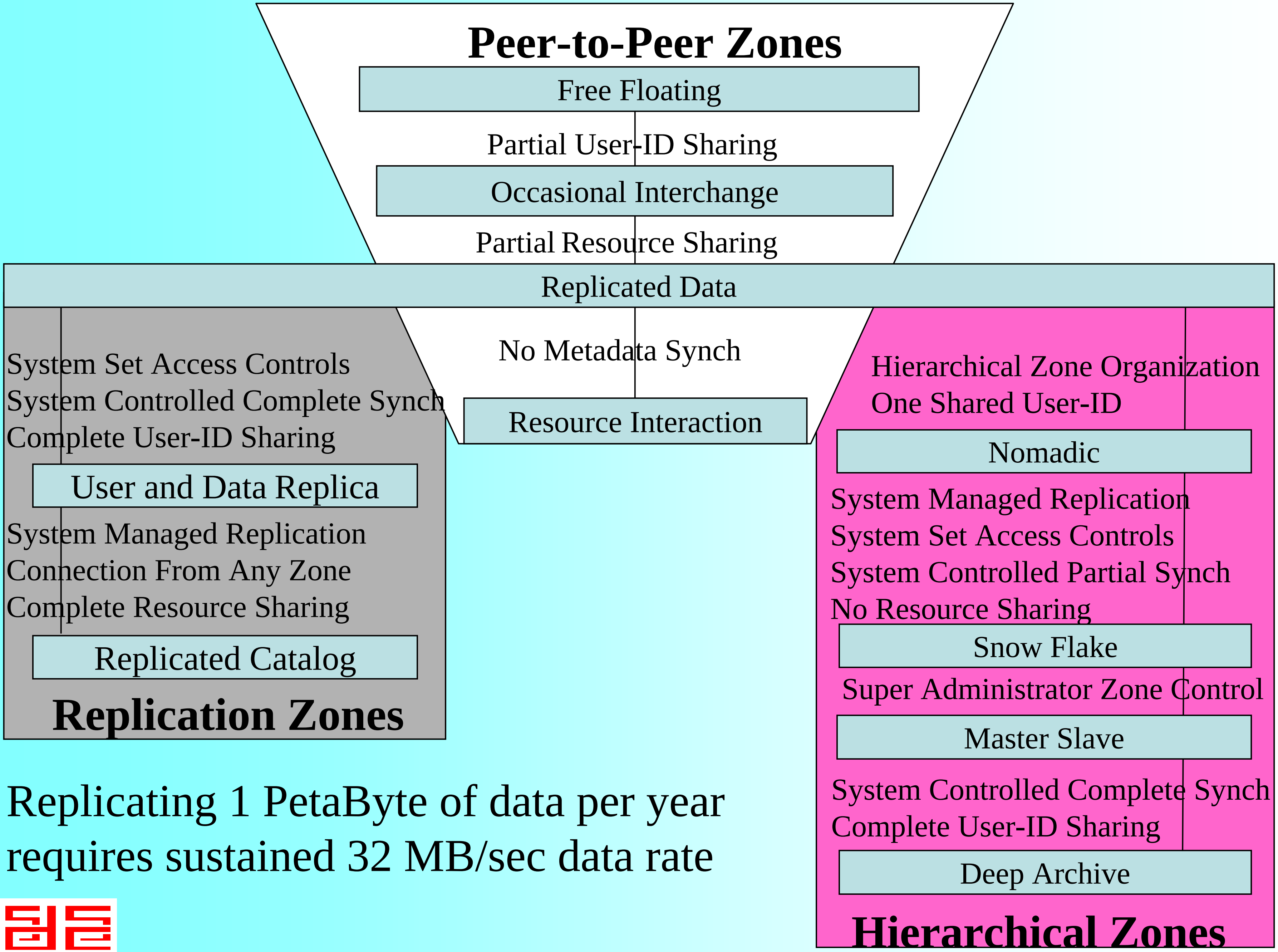
Knowledge repositories for managing collection properties
Characterization of the application of transformative migrations on encoding format
Characterization of the application of archival processes

Data Grid Federation

- Data grids provide the ability to name, organize, and manage data on distributed storage resources
- Federation provides a way to control sharing of resources, users, data and metadata between independent data grids.
- We call each data grid a “zone”, hence zoneSRB

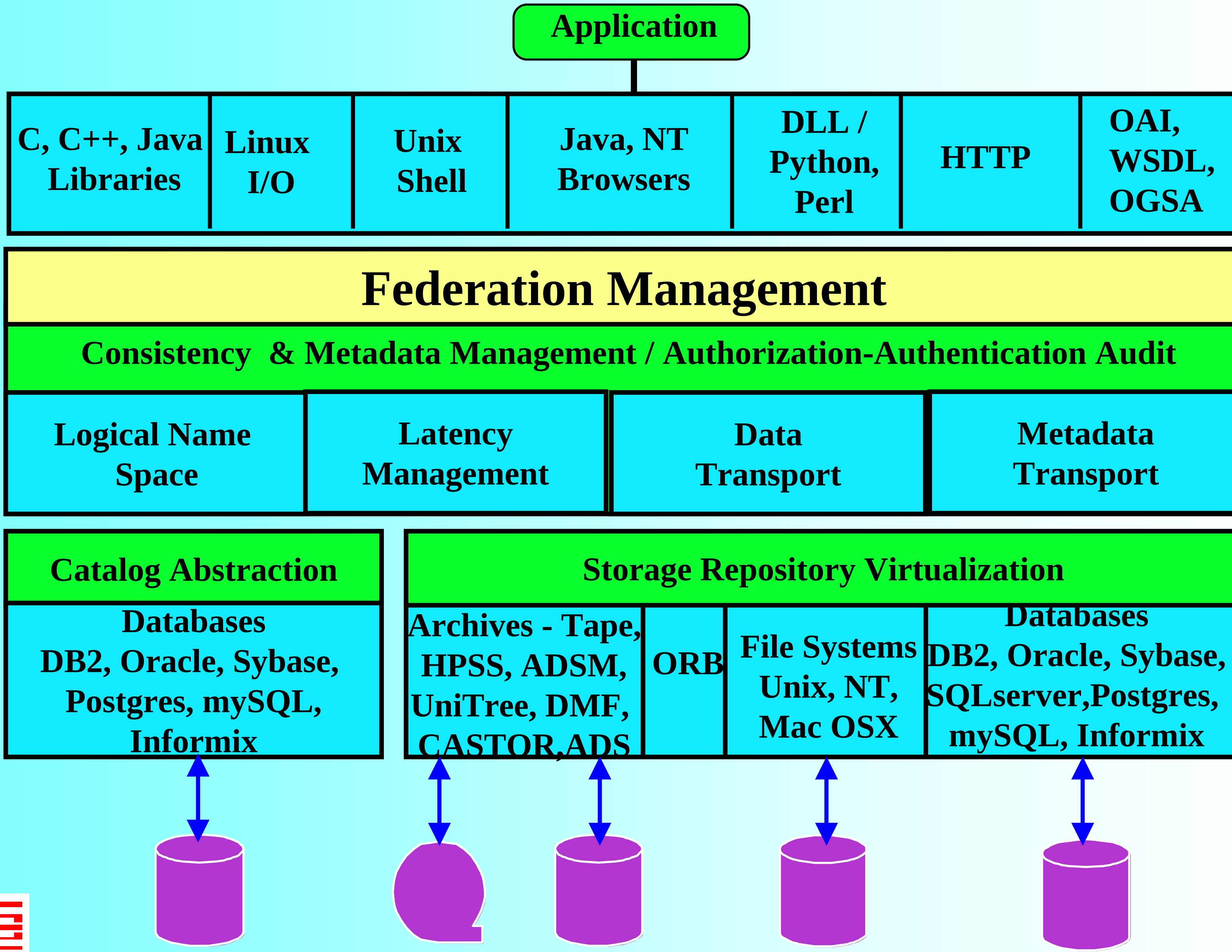
Deep Archive

- Impose sharing constraints:
  - Only system administrator access
  - Selected replication of files
  - Write once, with versions created on changes to data
- Impose consistency constraints
  - Coordinate update of preservation metadata with file replication
- Use federation to guarantee preservation against
  - Local hardware and software failures
  - Local operation errors
  - Local disasters



Replicating 1 PetaByte of data per year requires sustained 32 MB/sec data rate

Data Grid Federation - zoneSRB



For More Information

Reagan W. Moore
San Diego Supercomputer Center

moore@sdsc.edu

http://www.npaci.edu/DICE

http://www.npaci.edu/DICE/SRB

http://www.npaci.edu/dice/srb/mySRB/mySRB.html