

*Information Based Computing for  
High Performance Data Systems*

Reagan W Moore

San Diego Supercomputer Center, 10100 Hopkins Dr

La Jolla CA 92093

Phone: +1-858-534-5073

E-mail: [moore@sdsc.edu](mailto:moore@sdsc.edu)

Presented at the THIC Meeting at the DoubleTree Hotel

Del Mar CA 92130-2539

on January 19, 2000



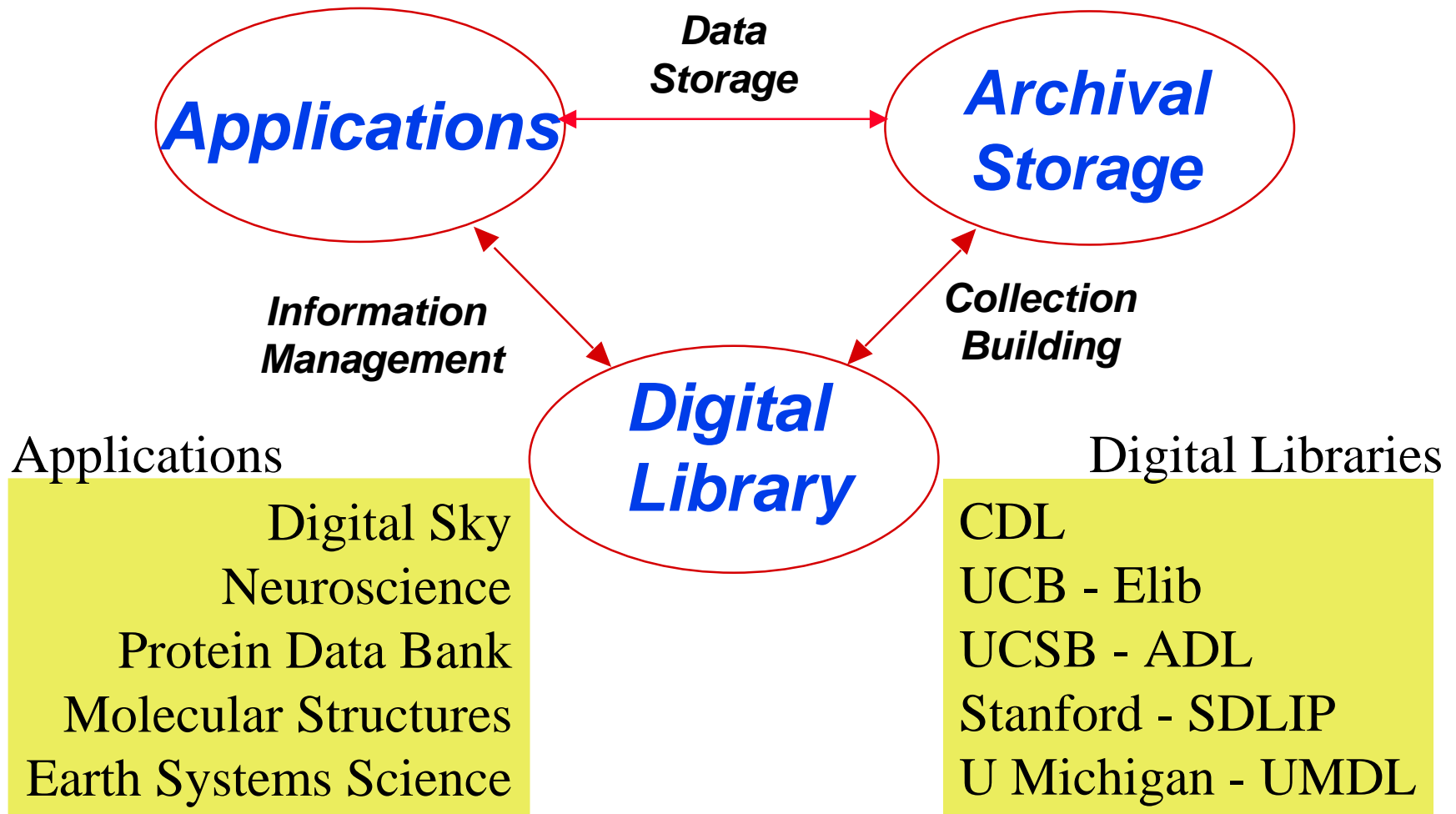
***Information Based Computing  
for  
High Performance Data Systems***

**Reagan W. Moore  
San Diego Supercomputer Center**

**moore@sdsc.edu  
<http://www.npaci.edu/DICE>**



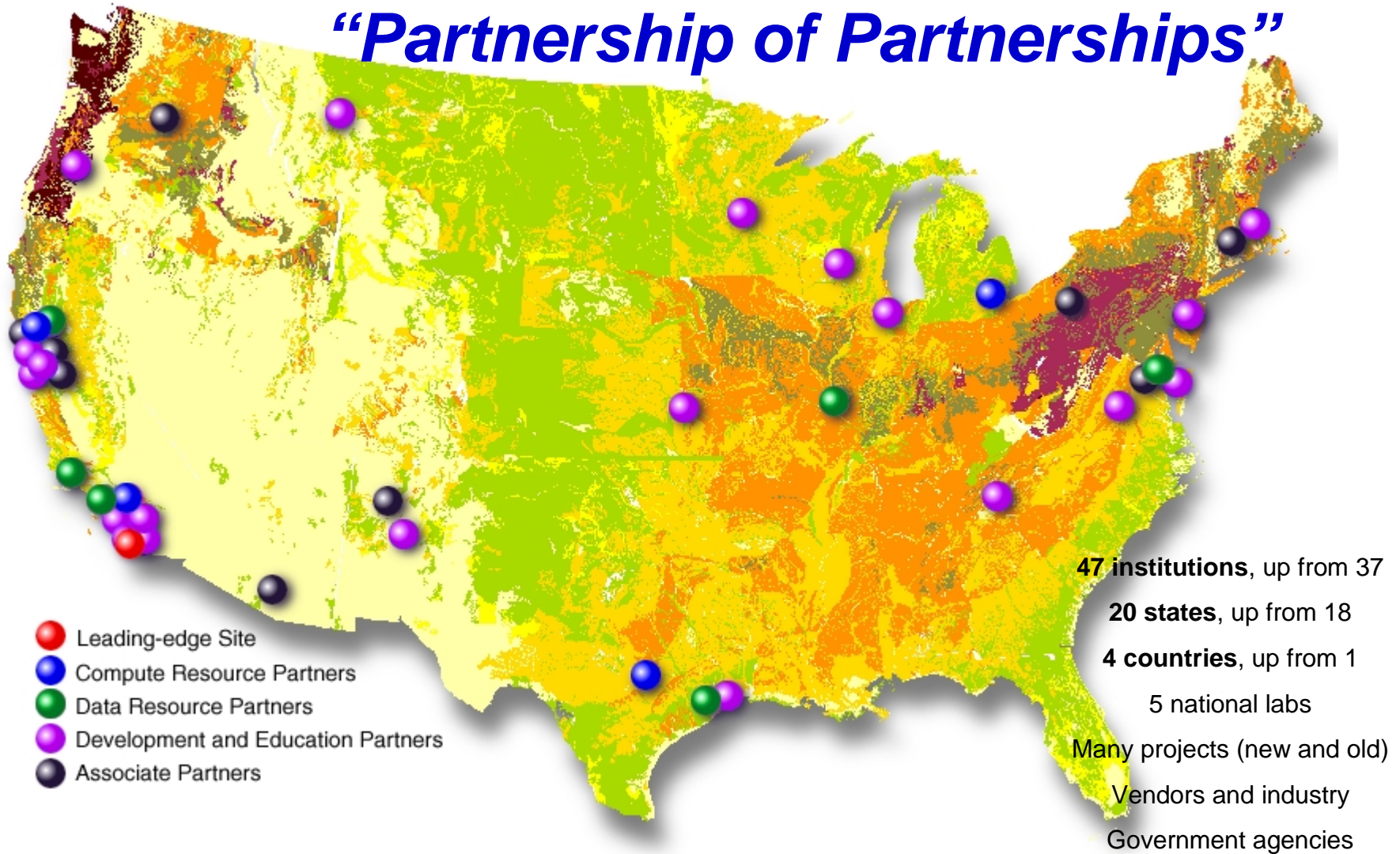
# Information Based Computing



# *Access to Data Collections*

- **Scientific Data Collections**
  - Publication of scientific data sets
  - Information discovery mechanisms
- **Information Models for Data**
  - eXtended Markup Language (XML) Document Type Definition (DTD)
    - Information model for digital objects, data collections, and presentation interfaces
  - Application to scientific data collections
    - Digital sky, Protein Data Bank, Neuroscience brain images
    - California Digital Library - Art Museum Image Consortium

# ***NPACI is a Highly Leveraged National “Partnership of Partnerships”***



# ***NPACI Application Thrust Areas***

- **Earth Systems Science - Bernard Minster (UCSD/SIO)**
  - Remote sensing data
- **Neuroscience - Mark Ellisman (UCSD)**
  - Federating brain image data
- **Molecular Science - Russ Altman (Stanford)**
  - Protein Data Bank
- **Engineering - Tinsley Oden (U Texas)**
  - Reservoir modeling

# Data Collections

<b>Collection Site</b>	<b>Project</b>	<b>Archive (TB)</b>	<b>Database (GB)</b>
U Wisconsin	Digital Insight / NPACI-EOT	10-20	5
UCOP	California Digital Library archive	5	10
Caltech	2MASS images / NPACI-DICE	2-20	2
UCSB	Alexander Digital Library / NPACI-DICE	2	2
UCOP	Art Museum Image Consortium / CDL	1.5	30
SDSC	CEED / Ecological Society of America	1	2
UCB	Elib flora collection / NPACI-DICE	1	60
U Maryland	HPSS federation / NPACI-DICE	1	1
SDSC	Human Brain Project / NPACI-Alpha NS	1	10
UCLA	Human Brain Project / NPACI-NS	1	2
Washington U	Human Brain Project / NPACI-NS	1	10
SDSC	Molecular Structures / NPACI-Alpha MS	1	10
SDSC	Protein DataBank	0.5	2
UCSD	PRDLA Chinese text collection	0.5	10
SDSC	Visualization image collection	0.5	5
SDSC	NARA – USPTO patents	0.3	70
UCSC	REINAS / NPACI-ESS	0.1	1
U Michigan	UM Digital Library / NPACI-DICE	0.1	30
SDSC	SRB Production system / NPACI-DICE		75

# Bio-Informatics

Application		Infrastructure
Structural Comparison (n x n)	Information Model	User interface / Analysis tools / Portals
Mediation of Information using XML / Extensible Meta-data Catalog		Federation / Mediation of Collections
Protein Data Bank Services	Markup Language	Digital Library Services
PDB / Genome / Molecular Trajectory Collections		Collection Management
Storage Resource Broker	Meta-data	Data Handling System
HPSS / file system		Archival Storage



# *Service Architecture*

- **Massive data scales**
  - Petabytes of data
  - Billions of data sets
  - Millions of data collections
- **Challenges**
  - Information discovery
  - Data movement from collection to application
  - Streaming of data through multiple cache levels

# *Challenges*

- **Facilitate access to high-end resources**
  - Support data intensive computing
- **Facilitate access to distributed data resources**
  - Support information discovery
- **Minimize complexity of user interfaces**
  - Provide unifying data access system
- **Requires information management infrastructure**

# *Collection Based Access*

- **Abstract data set naming and administration away from physical storage resource**
  - Data sets defined by attributes
  - Logical collection used to group data sets across storage systems
    - Enables support for replication of data
  - Collection owned data
- **Authentication controlled by data handling system**
- **Persistence controlled by data handling system**

# *Context Based Objects*

- **For data to be useful, the context must be defined**
  - Data format - binary/integer representation
  - Physical meaning - units
  - Structure - geometry
  - Relevance - feature annotation
  - Semantics - data dictionary for attributes
- **Context is preserved as meta-data attributes**
- **Are there common XML attributes for scientific data?**

# *Common Information Model*

- **eXtensible Markup Language (XML)**
  - Use tags to define semantic context for components of the data set
- **Document Type Definition (DTD)**
  - Provides semi-structured representation for organizing tags that can be applied to groups of digital objects
- **Development of standards for tags**
  - Digital sky, Protein Data Bank, Neuroscience brain images
  - California Digital Library - Art Museum Image Consortium

# *Hierarchy of Information Contexts*

- **Digital object context**
  - Meta-data to define the structure of the object
  - When publishing a digital object, must also publish the context of the object
- **Use collections to organize objects**
  - Meta-data to define the structure of the collection
  - When publishing a collection, must also publish the information needed to organize the collection.
- **Use presentation context to control access**
  - Meta-data to define structure of presentation

# *Information Management Hierarchy*

- **Persistent Archives**
  - Storage of information model, data model, along with data
- **Data Grid**
  - Access to data in a different administration domain
- **Digital Library - Presentation / Information Discovery**
  - Interlib - ADEPT, UC Berkeley Digital Library
- **Data Collection**
  - Extensible Meta-data catalog - EMCAT
- **Data handling**
  - SDSC Storage Resource Broker - SRB
- **Archival Storage**
  - High performance storage system - HPSS

# Synergistic Projects

Agency	Project	Archival storage	Data Handling	Collection	Digital Library	Data Grid	Persistent Archive
NARA	Persistent Archive	Scaling	Containers for data, XML DTD for meta-data, X.509 interface	DTD fitting, Meta-data mining on AMICO, general collection creation software	Schema evolution tools	Security architecture	Evaluation of Excelon/Oracle, conversion of patents to XML, prototype
NASA	IPG		SRB drivers for accessing NASA data sources	EMCAT extensible schema, XML DTD interface	Digital Library interface to archive	MCAT & LDAP & SDLIP integration	
NHPRC	Spatial Data	Lightweight archive		DTD fitting for GIS data	Spatial algebra for GIS, GIS querying	Spatial mediator for GIS, MIX wrapper for Web sites	Persistent archive infrastructure for local collections
NLM	Digital Embryo		Containers for data	Digital Embryo image meta-data		Replication of data between caches	
NSF	NPACI	Scaling, Performance, Petabyte archive	Distributed data access, containers for data, ADL backup into archive	EMCAT extensible schema, bulk meta-data ingestion, schema for neuroscience, 2MASS collection	Implement DOM/VXD with SDLIP	Federation of neuroscience collections, MIX wrapper for genome, integration with Globus & Legion	





# *Generality of Information Infrastructure*

- **Same information model needed to manage**
  - Federation in space
    - Metacomputing environment
    - Interoperable services for digital libraries
  - Migration over time
    - Collection creation and update
    - Persistent archive
- **Same storage systems needed to support**
  - Supercomputer center data
  - Discipline specific data collections
  - Digital library collections

# *Applications*

- **Support for distributed data collections**
- **Federation of data collections into digital library**
- **Integration of digital libraries with archives**
- **Finding aids for federation of digital libraries through mediation of information**
- **Data grids for application access**
- **Persistent archives**

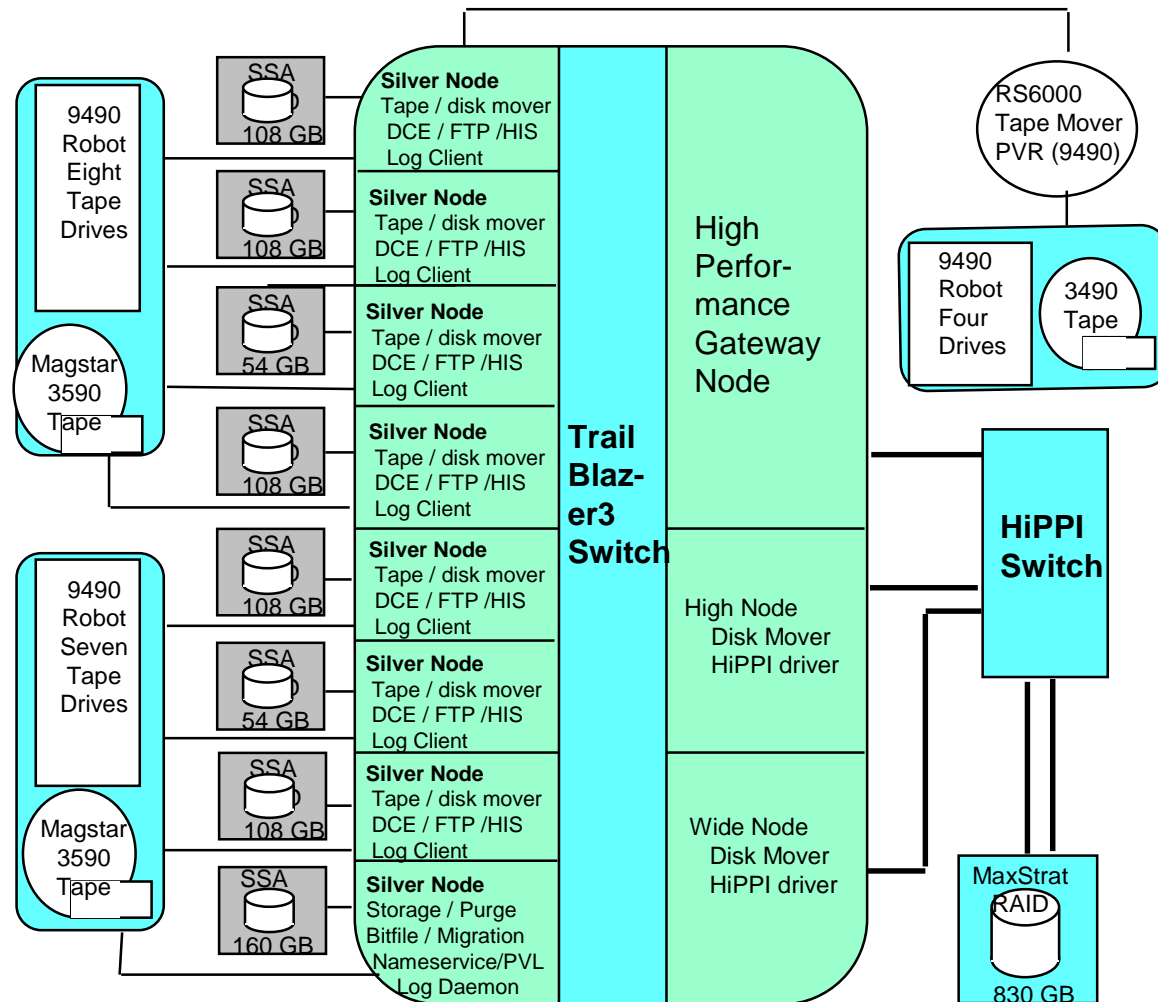
# *Common Infrastructure*

- **Archival storage system**
  - Long term storage of data
  - Migration to low cost media
  - Massive data storage
- **Projects**
  - NPACI archival storage for teraflops system
  - NASA IPG
  - NPACI data collections
  - CDL, ADL, ELIB, UMDL digital library
  - NARA Persistent archives

# *High Performance Storage System*

- **Archival storage to support teraflops computer**
  - Capacity 500 TB in 3 tape robots
  - 2-5 TBs of disk
  - 70 MB/sec data access
- **Usage of system**
  - 110 TB of data stored
  - 11 million files
  - 16,000 transactions per day

# HPSS Archival Storage System



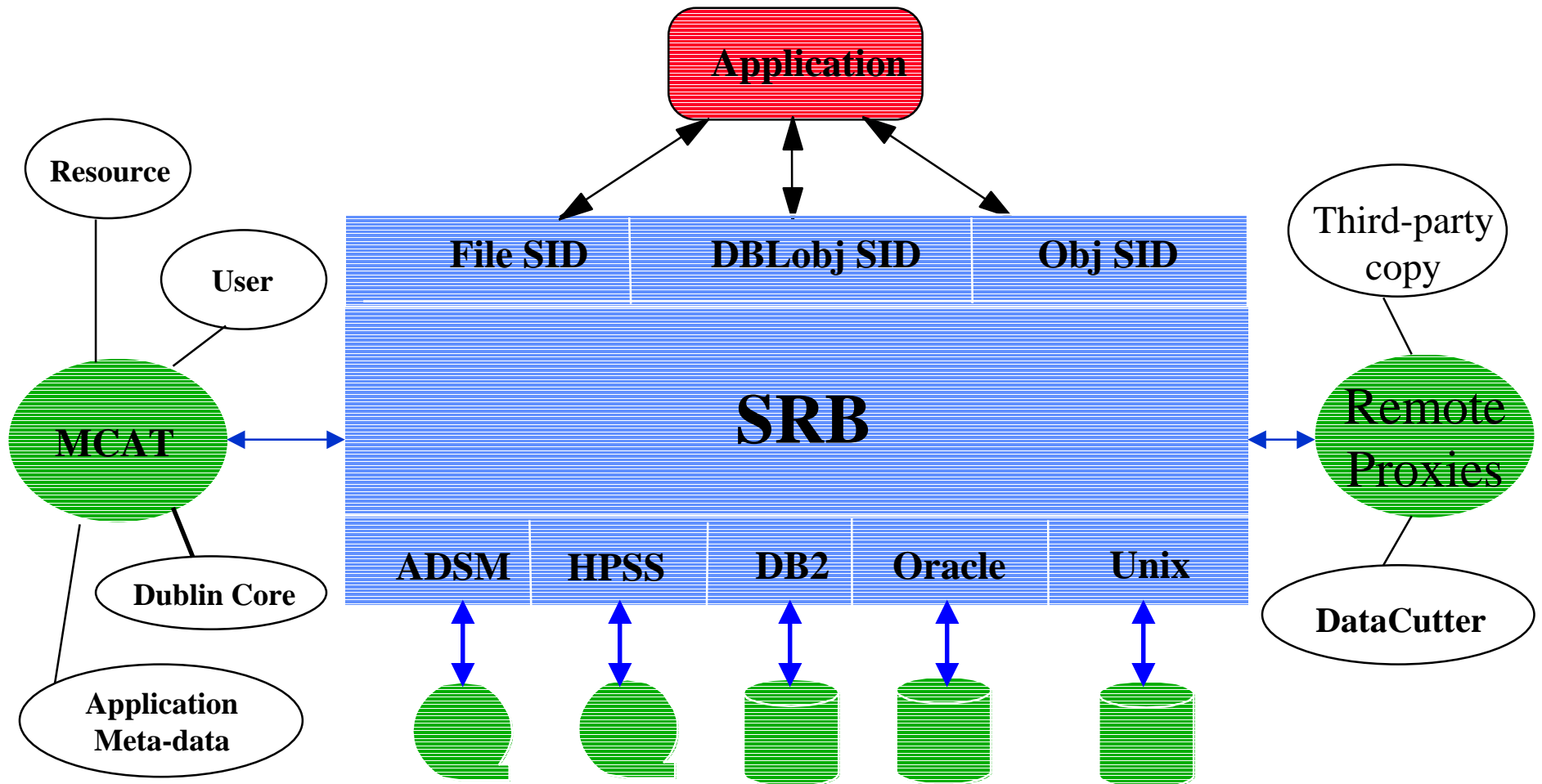
# *Common Infrastructure*

- **Data Handling System**
  - Remote data access
  - Heterogeneous storage resource access
  - Provision of a common API
- **Projects**
  - NPACI remote data access
  - ASCI linkage of archives to analysis platforms
  - NASA IPG
  - CDL federation of data collections
  - NARA persistent archives

# *Grid Transparencies*

- **Location transparency**
  - Distribution of data collection across multiple physical resources
- **Name transparency**
  - Attributed based access to data
- **Protocol transparency**
  - Common API for access to remote data resources
- **Time transparency**
  - Minimization of data access latency

# SDSC Storage Resource Broker & Meta-data Catalog





# SRB Production Sites

- **SRB Servers: 19 sites, 25 hosts, 46 resources, 90 users, 350,000 data sets**

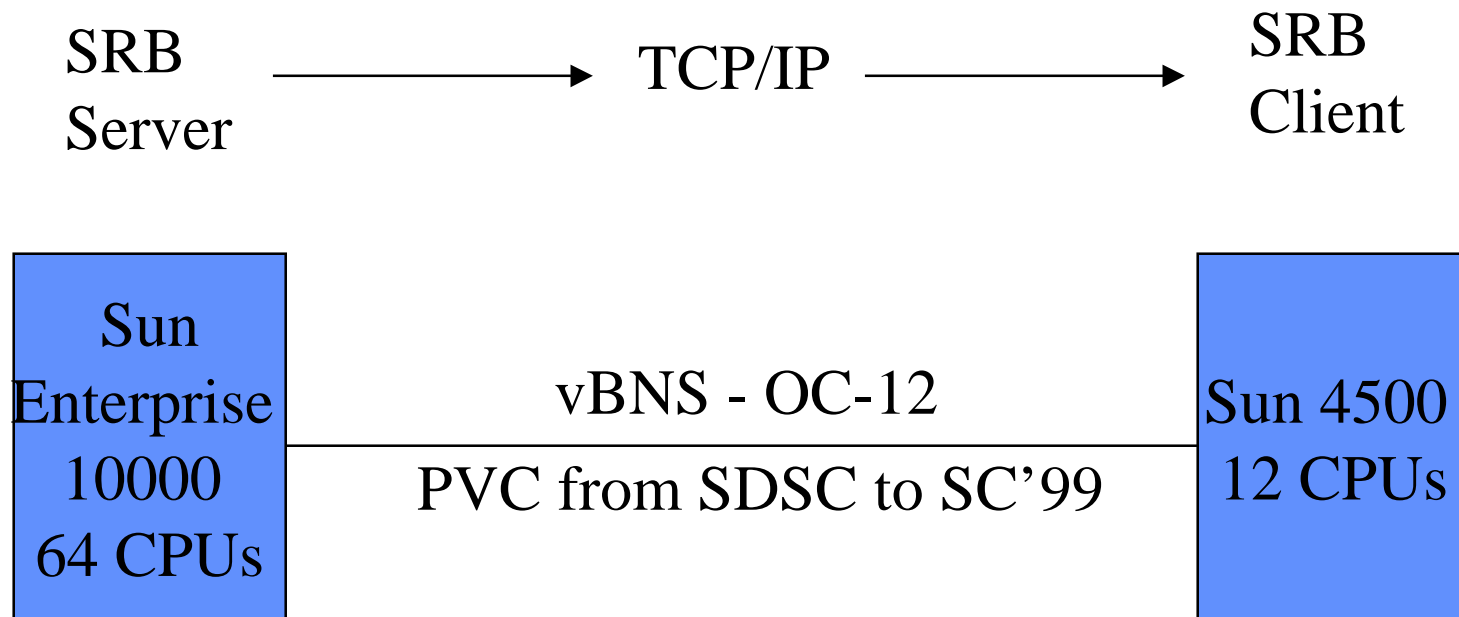
SDSC - 4 hosts                      V1.1.4 (HPSS, DB2, Oracle, Illustra, UnixFS, C-90Unicos)  
U. Maryland    V1.1.4 (HPSS, UnixFS)  
U. Michigan    V1.1.3 (ADSM, UnixFS)  
UIUC(NCSA)                      V1.1 (Oracle, UnixFS)  
Rutgers U.        V1.1.2 (UnixFS)  
CalTech         V1.1.4 (HPSS, UnixFS)  
UC Berkeley     V1.1.4 (UnixFS)  
Montana State U                      V1.1.4 (UnixFS)  
UCLA            V1.1.4 (UnixFS)  
UCSB            V1.1.3 (UnixFS)  
U Texas, Austin                      V1.1.3 (DMF, UnixFS)  
UC Davis        V1.1.3 (UnixFS), V1.1 (UnixFS)  
Washington U, StL                      V1.1.4 (UnixFS), V1.1 (UnixFS)  
U Houston        V1.1.3 (UnixFS)  
UCSC            V1.1.4 (Oracle, UnixFS)  
UCSD - 2 hosts                      V1.1.4 (UnixFS)  
LBL              V1.1.3 (UnixFS)  
LLNL             V1.1.3 (DB2, UnixFS)  
U Chicago        V1.1.4 (UnixFS)

- **MCAT:**

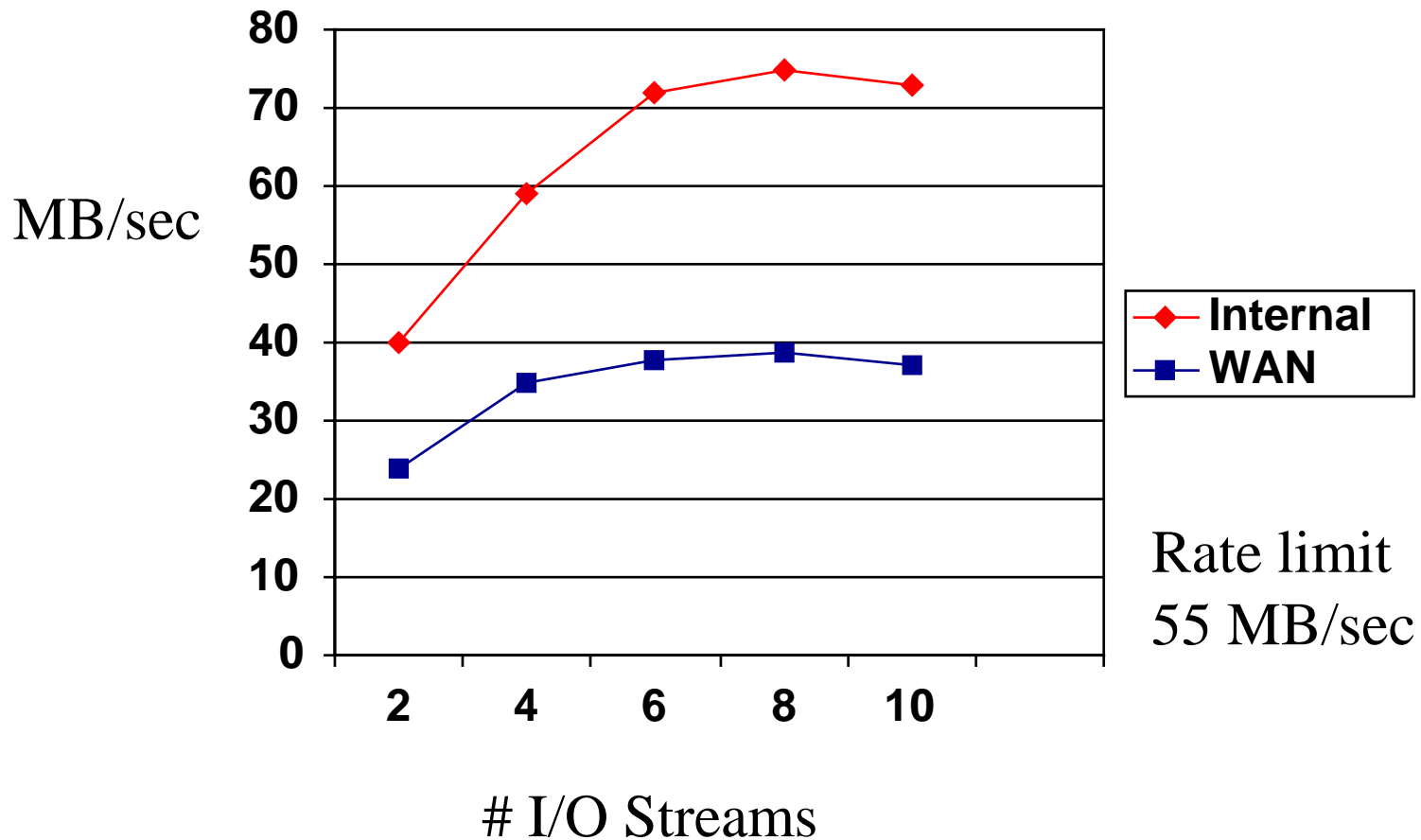
SDSC (UDB(DB2), Oracle)  
LLNL (UDB(DB2))  
LBL (MySQL)  
UCSC (Oracle)



# Test Configuration



## Storage Resource Broker Wide Area Network Performance



# *Common Infrastructure*

- **Data Collection**

- Data ingestion
- Data organization
- Schema support
  - Meta-data attributes
  - Attribute organization

- **Projects**

- AMICO image collection
- 2MASS digital sky image collection
- PRDLA chinese text
- PDB, ESA

# *Art Museum Image Consortium*

- **Demonstrated**

- Support for heterogeneous digital objects
- Automated conversion of meta-data to XML DTD
- Validation of meta-data
- XSL style sheet for presenting information

# AMICO Meta-data Conversion to XML

The image displays two side-by-side browser windows illustrating the conversion of AMICO metadata to XML. The left window shows the original metadata text, and the right window shows the resulting XML code.

**Left Window (Original Metadata):**

Address: <http://www.npaci.edu/DICE/AMICO/Demo/amico-source-records-col>

Links: scratch.html Principles of Database Systems -- Notes Persistent

XIDAKAG.1964.16.tif)-XDEFull View)-XPUAlbright-Knox  
-XRTReproduction)-XAMImage)-XFO)-XFFtiff)-XCMsep  
-XFD1650 x 1361)-XFF8983372)-XFCnone)-XRE)-XRYIs  
-XRIAKAG.1964.16)-XRSCopyright Albright-Knox Art Ga  
-XVD19980430)-XVV1)-XLY1998)-

AIDASIA.1979.251)-OTYCeramics)-OTG)-OTNTEa Leaf.  
12 in. (30.5 cm); D. 9 1/2 in. (24.1 cm)-OMG)-OMDStonew  
with overglaze enamels and silver (Kyoto ware)-CRG)-CR  
Ninsei)-CRNNonomura Ninsei)-CRCJapanese)-CDTc. 15'  
-OCG)-OCTEdo period, mid-17th century)-OCPJapan, Kyo  
Prefecture)-OOG)-OONAsia Society)-OOPNew York, Nev  
-OOA1979.251)-OOCAsia Society: The Mr. and Mrs. John  
Rockefeller 3rd Collection)-ORG)-ORLhttp://www.asiasoci  
-RIG)-RIPY)-RIDFull view: front)-RIRHasFormat)  
-RILASIA.1979.251.a.tif)-RIG)-RIPN)-RIDFull view: bac  
-RIRHasFormat)-RILASIA.1979.251.b.tif)-RDG)-RDDAs  
Handbook of the Mr. and Mrs. John D. Rockefeller 3rd Coll  
York: Asia Society, [1981], p. 109.-RDRReferences)-RDG  
-RDDMeech-Pekarik, Julia. "Notable Japanese Ceramics."  
(November 1983), p. 433.-RDRReferences)-RDG)-RDDSh  
Shujiro, ed. Zaigai Nihon no shuho (Japanese Art Treasures  
Tokyo: Mainichi Shinbun, 1981, vol. 9, pl. 76.-RDRReferen  
-AVD19980430)-AVV1)-ALY1998)-ADPParsed OCT into  
-)-

XIDASIA.1979.251.a.tif)-XDEFull view: front)-XPUAsia So  
-XDA19980318)-XRTReproduction)-XAMImage)-XFO)-2  
-XFD1024 x 1536)-XFF4.5 MB)-XFCNone)-XRE)-XRYIs  
-XRIASIA.1979.251)-XRSCopyright, Asia Society)-XVD15

Done

Internet

**Right Window (XML Output):**

Address: <http://www.npaci.edu/DICE/AMICO/Demo/amico-objects.xml>

Links: scratch.html Principles of Database Systems -- Notes Persistent URL Home Page

```
<?xml version="1.0" ?>
<!DOCTYPE am_objects (View Source for full doctype...)>
- <am_objects>
+ <am_object>
+ <am_object>
+ <am_object>
- <am_object>
  <AID>ASIA.1979.251 </AID>
  <OTY>Ceramics</OTY>
+ <OTG>
  <MET>H. 12 in. (30.5 cm); D. 9 1/2 in. (24.1 cm)
  </MET>
+ <OMG>
+ <CRG>
- <OCG>
  <OCT>Edo period, mid-17th century</OCT>
</OCG>
<OCP>Japan, Kyoto Prefecture</OCP>
- <OOG>
  <OON>Asia Society</OON>
  <OOP>New York, New York, USA</OOP>
  <OOA>1979.251</OOA>
  <OOC>Asia Society: The Mr. and Mrs. John D.
    Rockefeller 3rd Collection</OOC>
</OOG>
- <ORG>
  <ORL>http://www.asiasociety.org</ORL>
</ORG>
- <RTG>
```

Internet

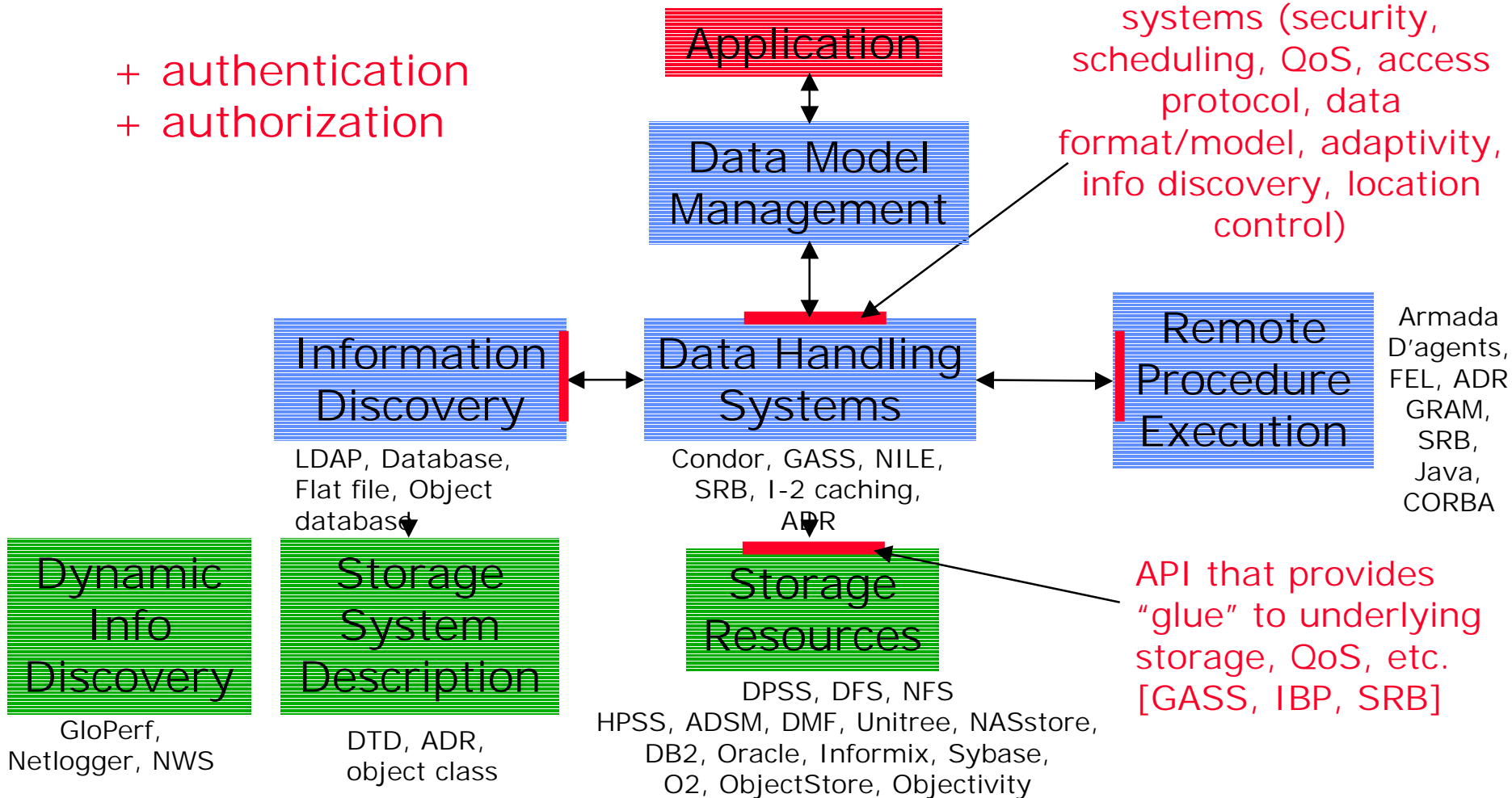
# *Common Infrastructure*

- **Data Grid**
  - Distributed data access
  - Information discovery
  - Access to distributed services
  - Access to computation servers
- **Projects**
  - Grid Forum
  - NPACI data grid
  - NASA IPG
  - ASCI DVC

# Grid Forum Data Access Architecture

+ authentication  
+ authorization

API that provides "glue" to underlying data handling systems (security, scheduling, QoS, access protocol, data format/model, adaptivity, info discovery, location control)



API that provides "glue" to underlying storage, QoS, etc. [GASS, IBP, SRB]





## *Further Information*

<http://www.npaci.edu/DICE>