Storage Resource Broker
Global Data Grids

Reagan W. Moore

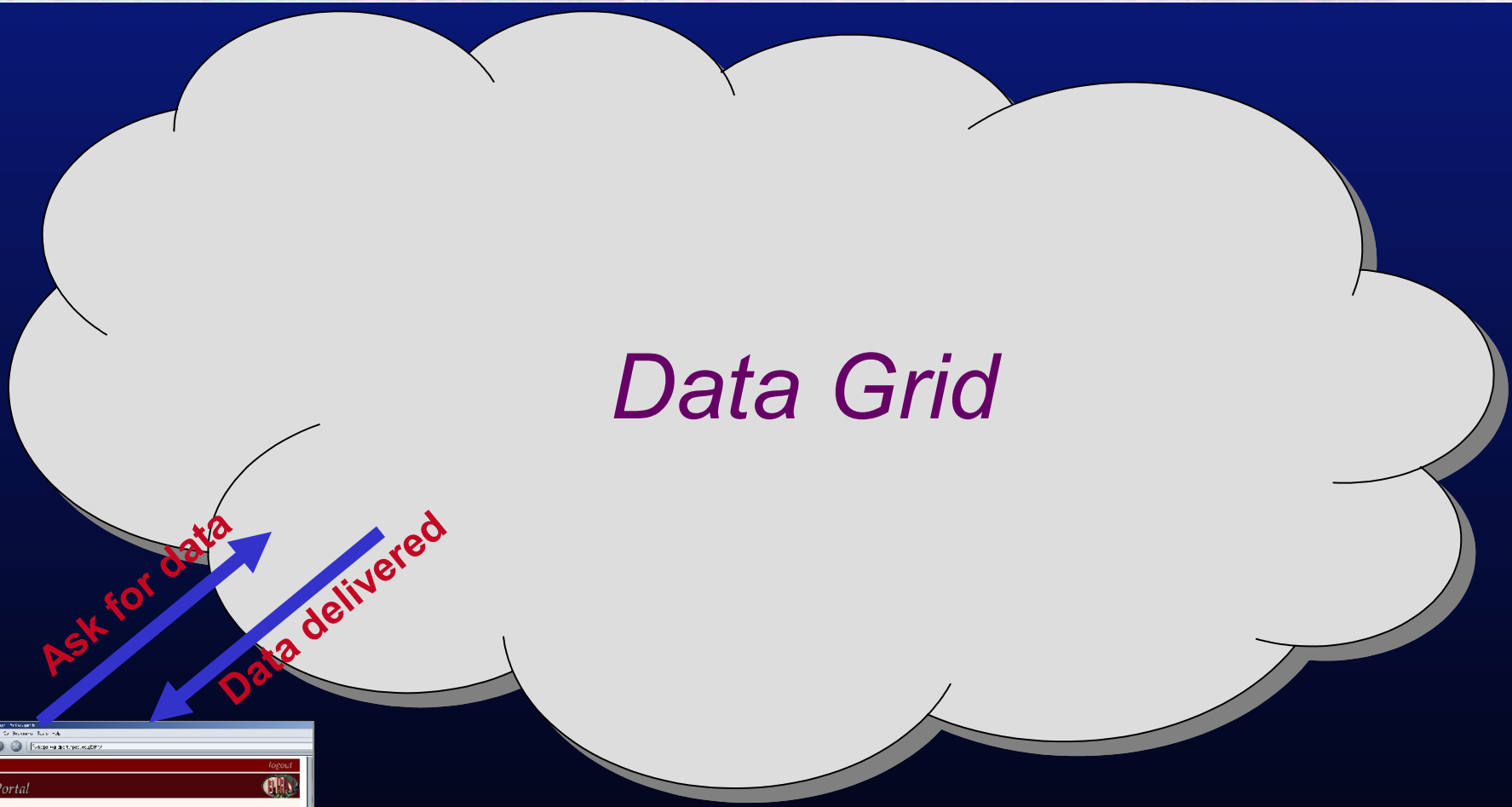
Arcot Rajasekar

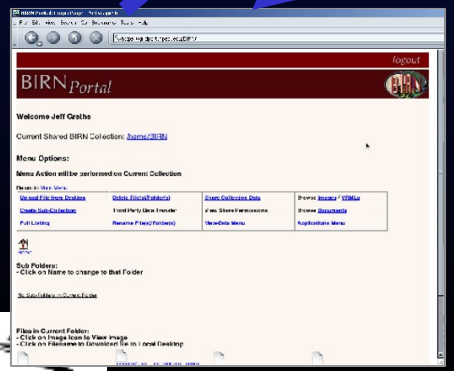
Michael Wan

{moore,sekar,mwan}@sdsc.edu

http://www.sdsc.edu/srb

Using a Data Grid – *in Abstract*



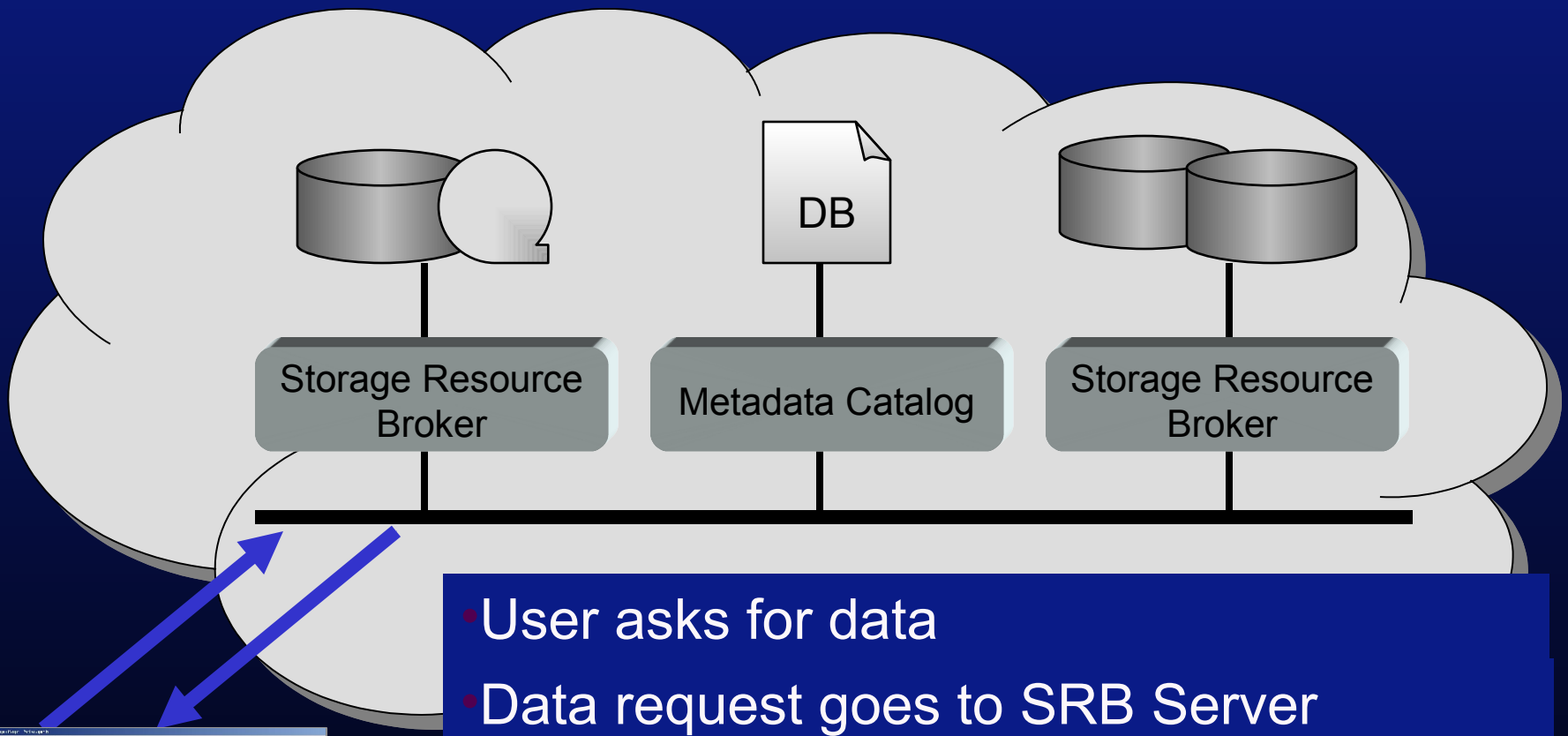


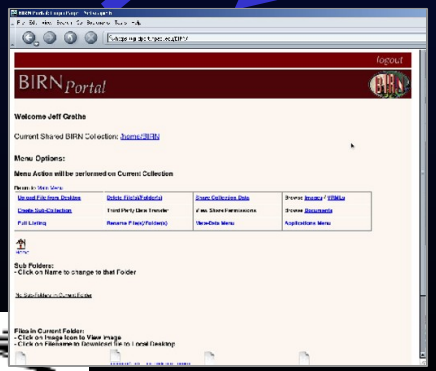
- User asks for data from the data grid
- The data is found and returned
  - Where & how details are hidden

UCSD

SDSC

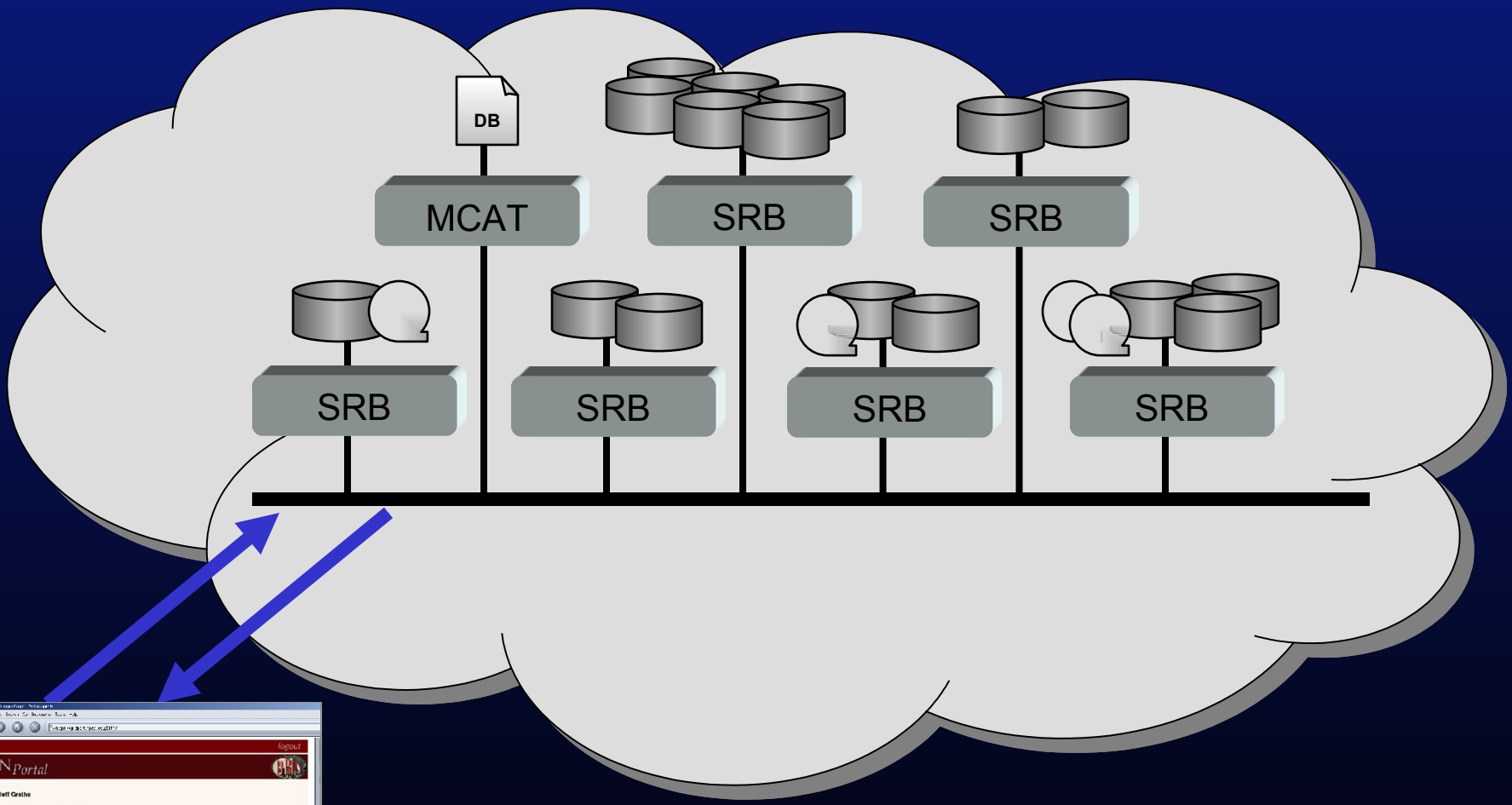
Using a Data Grid - *Details*

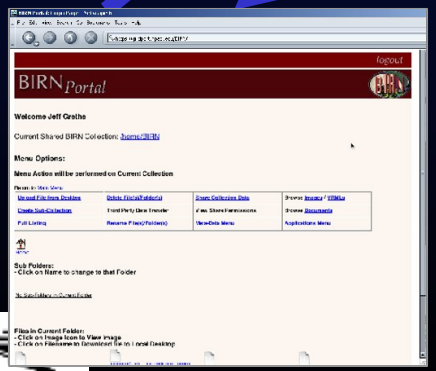




- User asks for data
- Data request goes to SRB Server
- Server looks up data in catalog
- Catalog tells which SRB server has data
- 1st server asks 2nd for data
- The data is found and returned

Using a Data Grid - *Details*





- Data Grid has arbitrary number of servers
- Heterogeneity is hidden from users

NSF / NVO	17,800	5,139	51,380	8,690	80	100,990	13,217	100
NSF / NPACI	1,972	1,083	17,578	4,694	380	34,830	7,239	380
Hayden	6,800	41	7,201	113	178	8,013	161	227
Pzone	438	31	812	47	49	23,099	13,287	68
NSF / LDAS-SALK	239	1	4,562	16	66	115,178	146	67
NSF / SLAC-JCSG	514	77	4,317	563	47	17,095	1,775	55
NSF / TeraGrid			80,354	685	2,962	202,226	4,443	3,267
NIH / BIRN			5,416	3,366	148	16,288	15,306	361
NSF / LTER	158	3	233	6	35	236	34	36
NSF / Portal	33	5	1,745	48	384	2,620	53	460
NIH / AfCS	27	4	462	49	21	733	94	21
NSF / SIO Explorer	19	1	1,734	601	27	2,605	1,121	27
NSF / SCEC			15,246	1,737	52	167,140	3,471	73
NARA	7	2	63	81	58	2,916	2,004	58
NSF / NSDL			2,785	20,054	119	5,653	50,600	136
UCSD Libraries			127	202	29	190	208	29
NHPRC / PAT						1,338	519	28

Shared Collections

- **Purpose of SRB data grid 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, …

Generic Infrastructure

- **Digital libraries now build upon data grids to manage distributed collections**
  - DSpace digital library - MIT and Hewlitt Packard
  - Fedora digitial 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
- **Real-time sensor systems**
  - Use data grids to federate multiple independent systems

Biomedical Informatics Research Network BIRN

Data Grid

Requirements

- **Extensibility**
  - Add new nodes to data grid dynamically
  - Minimize administrative support
- **Uniform name spaces**
  - Users
  - Files
  - Metadata
- **Interactive response**
  - Master-slave metadata catalogs

Requirements

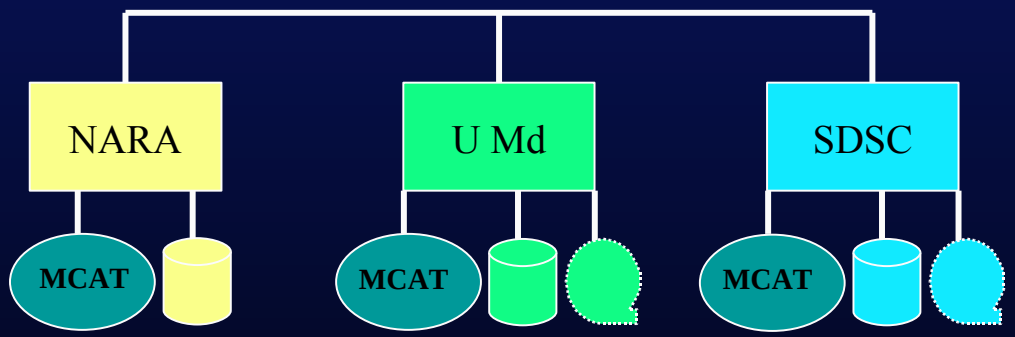
- **HIPAA patient confidentiality**
  - Authentication of users
  - Access controls on data
  - Access controls on metadata
  - Access controls on storage systems
  - Pinning of files to storage systems
  - Audit trails
  - End-to-end encryption

NARA Persistent Archive

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
- 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

Active archive at SDSC, user access

UCSD

SDSC

Federation

- **Authentication**
  - User defined by triplet:
  - (zone: user-name.domain-name)
- **Federation of zones**
  - Requires establishment of trust between data grids
    - Remote zone sends authentication requests to user's home zone
    - Authorization (access controls) applied by remote zone

Federation

- **Support high availability**
  - Independent zone administration
- **Improve interactivity**
  - Local metadata catalog
- **Improve reliability**
  - Preservation environment
- **Control user access**
  - Deep archive

Logical Name Spaces

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

Data Collection

Storage Repository

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

Data Grid

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

Data is organized as a shared collection

UCSD

SDSC

Federation Between Data Grids

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

Data Collection A

Data Collection B

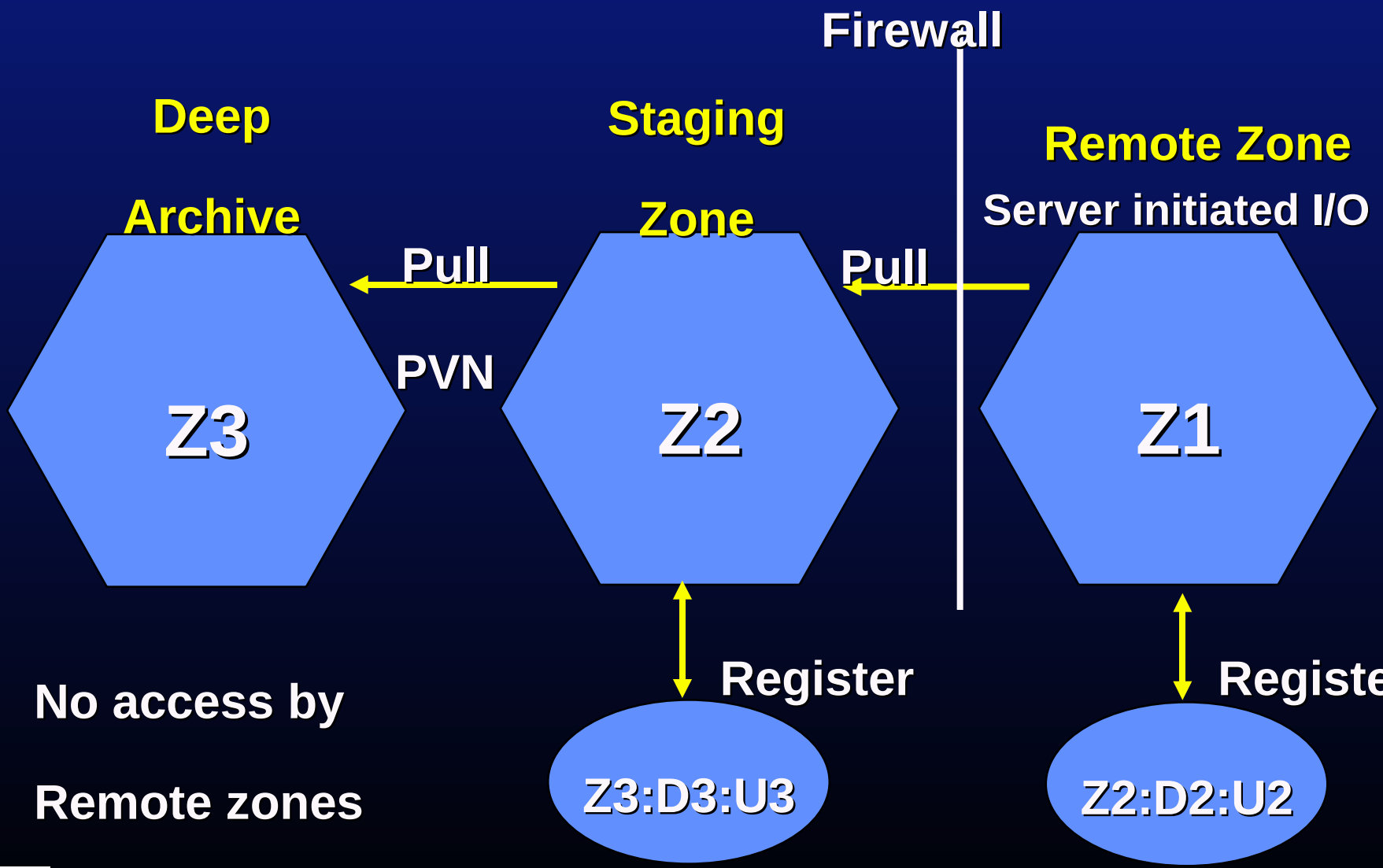
Data Grid		Data Grid
• Logical resource name space	⟷	• Logical resource name space
• Logical user name space	⟷	• Logical user name space
• Logical file name space	⟷	• Logical file name space
• Logical context (metadata)	⟷	• Logical context (metadata)
• Control/consistency constraints	⟷	• Control/consistency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

Access controls and consistency constraints
on cross registration of digital entities

UCSD

SDSC

Deep Archive



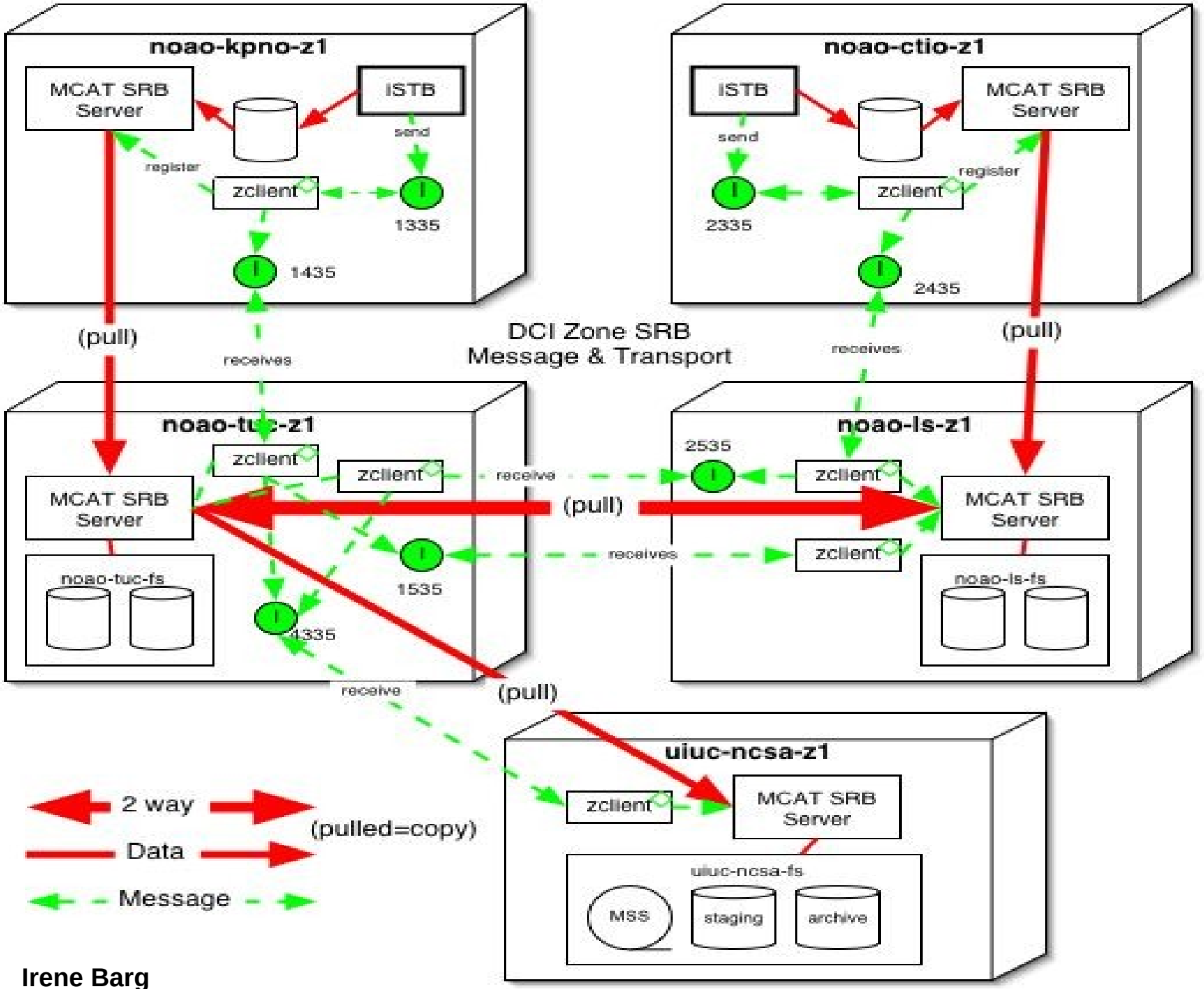
National Optical Astronomy Observatory

- Federation of 5 data grids between Chile and the US
- Management of preservation environment
  - **Assured replicas**
- **High availability**
  - **Independent management systems**

- **NOAO/NCSA DTS Team:**
  Irene Barg, Ray Plante, Phil Warner

UCSD

SDSC


noao-kpno-z1
MCAT SRB Server
ISTB
send
register
zclient
1335
1435
noao-ctio-z1
ISTB
MCAT SRB Server
send
register
zclient
2335
2435
(pull)
receives
DCI Zone SRB
Message & Transport
(pull)
receives
noao-tuc-z1
zclient
zclient
receive
MCAT SRB Server
noao-tuc-fs
1535
1335
receive
(pull)
noao-ls-z1
2535
zclient
zclient
MCAT SRB Server
receives
noao-ls-fs
(pull)
uiuc-ncsa-z1
zclient
MCAT SRB Server
uiuc-ncsa-fs
MSS
staging
archive
2 way
Data
Message
(pulled=copy)

Worldwide University Network Data Grid

- SDSC
- Manchester
- Southampton
- White Rose
- NCSA
- U. Bergen

Manchester-SDSC mirror

- A functioning, general purpose international Data Grid for academic collaborations

UCSD

SDSC

WUNGrid Shared Collections

- **BioSimGrid**
  - Molecular structure collaborations
- **White Rose Grid**
  - Distributed Aircraft Maintenance Environment
- **Medieval Studies**
- **Music Grid**
- **e-Print collections**
  - DSpace digital library
- **Astronomy**

UCSD

SDSC

GGF SRB Data Grid Federation

Data Grid	Country	SRB version	Demouser ggfsdsc	SRB Zone name	Storage Resource Logical Name	I/O MB/sec
APAC	Australia	3.4.0-P	yes	AU	StoreDemoResc_AU	3.9
NOAO	Chile/US	3.4.1	yes	noao-ls-t3-z1	noao-ls-t3-fs	
ChinaGrid	China	CGSP-II	(software)			
IN2P3	France	3.4.0-P	yes	ccin2p3	LyonFS4	[25.]
DEISA	Italy	3.4.0-P	yes	DEISA	demo-cineca	
KEK	Japan	3.4.0-P	yes	KEK-CRC	rsr01-ufs	7.4
SARA	Netherlands	3.4.0-P	yes	SARA	SaraStore	
IB	New Zealand	3.4.1	yes	aucklandZone	aucklandResc	(0.3)
ASGC	Taiwan	3.4.0-P	yes	TWGrid	SDSC-GGF_LRS1	(0.1)
NCHC	Taiwan	3.4.0-P	yes	ecogrid	ggf-test	
RAL	UK	3.4.0-P	yes	tdmg2zone		
IB	UK	3.4.1	yes	avonZone	avonResc	
WunGrid	UK	3.3.1	(hardware)	SDSC-wun	sfs-tape	
Purdue	US	3.4.0-P	yes	Purdue	uxResc1	(2.5)
Teragrid	US	3.4.0-P	yes	SDSC-GGF	sfs-disk	
U Md	US	3.4.0-P	yes	umiacs	narasrb02-unix1	

Trust Virtualization

Logical naming	Standard operations	State information
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

UCSD

SDSC

Data Virtualization

Logical naming	Standard operations	State information
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

Data Virtualization

Logical naming	Standard operations	State information
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	

UCSD

SDSC

Latency Management

Logical naming	Standard operations	State information
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 Management

Logical user names	Add or delete user	User:Group:Zone
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

Logical naming	Standard operations	State information
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	

SRB Developers

Reagan Moore	- PI
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
Marcio Faerman	- SCEC Application
Sifang Lu	- ROADnet Application
Richard Marciano	- SALT persistent archives

UCSD

SDSC

For More Information

Reagan W. Moore
San Diego Supercomputer Center
moore@sdsc.edu
http://www.sdsc.edu/srb/

UCSD

SDSC