Storage Resource Manager

Don Petravick
Timur Perelmutov

Fermi National Accelerator Lab

4/15/04

SRM Motivation

Grid Architecture requires Reservation and Scheduling of the Following **Shared** Resources

- Computing Resources
- Network Resources
- <u>Storage Resources</u> (often neglected)

SRM provides Reservation and Scheduling of the Storage Resources

SRM Motivation

High Energy Physics Collaborations span multiple institutions where

- A large variety of types of Storage Systems exist
    - Single Disk / Raid
    - Robotic Tape Storage System (Enstore, HPSS)
    - Distributed Disk Cache (dCache)
    - Hierarchical Storage System (dCache - Enstore)
- Heterogeneous environments and proliferation of custom Mass Storage Systems (MSSs) exist
- User applications often need to access data at multiple institutions on multiple MSSs

SRM provides Standardized Uniform Access to Heterogeneous Storage

Access to Multiple MSS
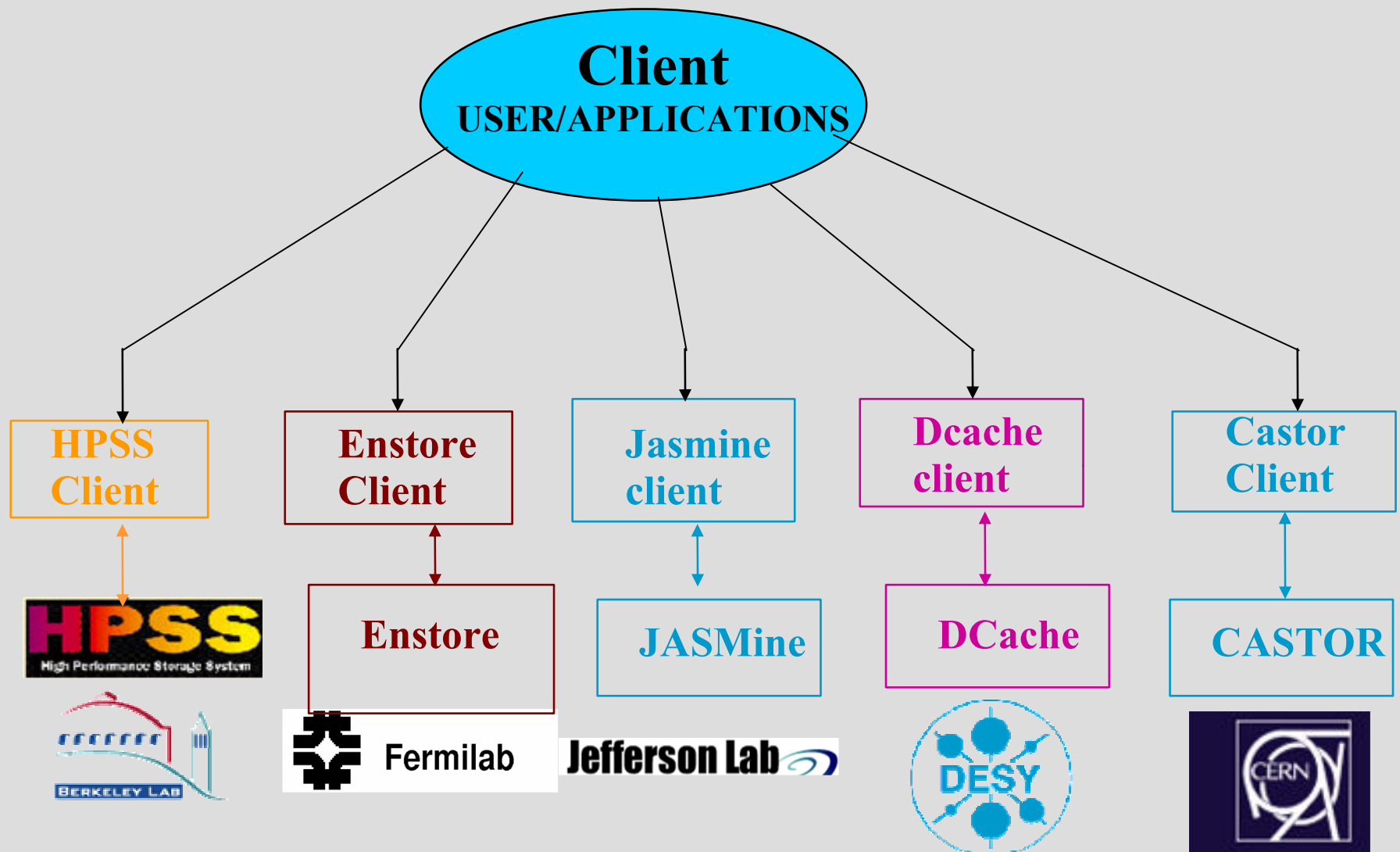
Uniform Access via SRM
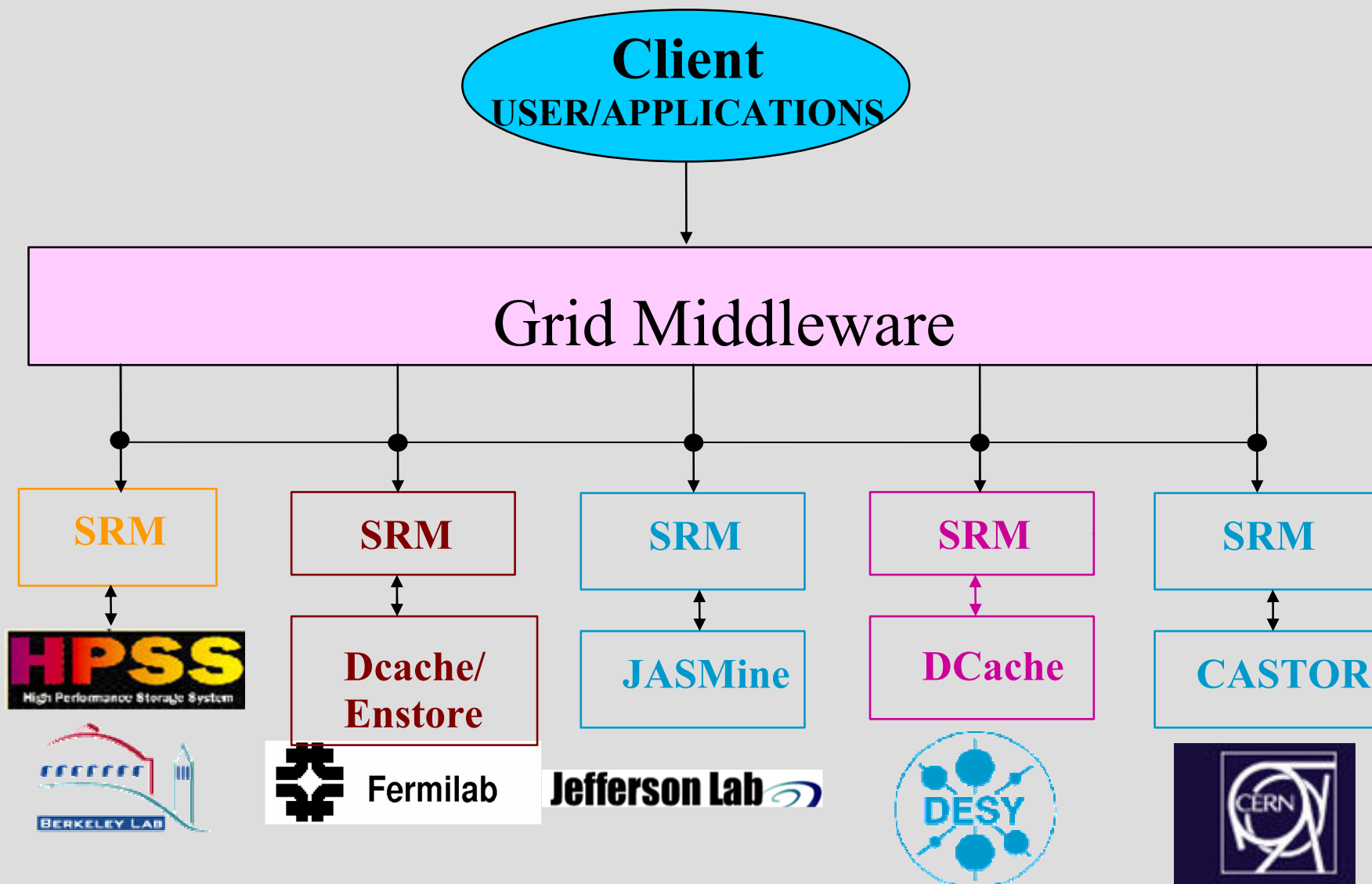
Storage Classification

Storage systems can be classified by:

- Persistence of data
  - Permanent
  - Temporary
- Data access availability
  - Data immediately available
  - Data needs advanced reservation before utilization (tapes need to be mounted, files need to migrate to disks, etc.)
- Supported transfer protocols
  - File transfer protocols
  - POSIX like access protocols

Need management interface that supports all of the above

Storage Resource Managers

SRMs Are Middleware Components That Manage Shared Storage Resources on the Grid and Provide:

✓Uniform Access to Heterogeneous Storage
✓Protocol Negotiation
✓Access to Permanent and Temporary Type of Storage
✓Advanced Space and File Reservation
✓Reliable Transfer Services

SRM
File and Space Types

Volatile   Durable   Permanent   ← **Space Types**

Volatile   Durable   Permanent   ← **File Types**
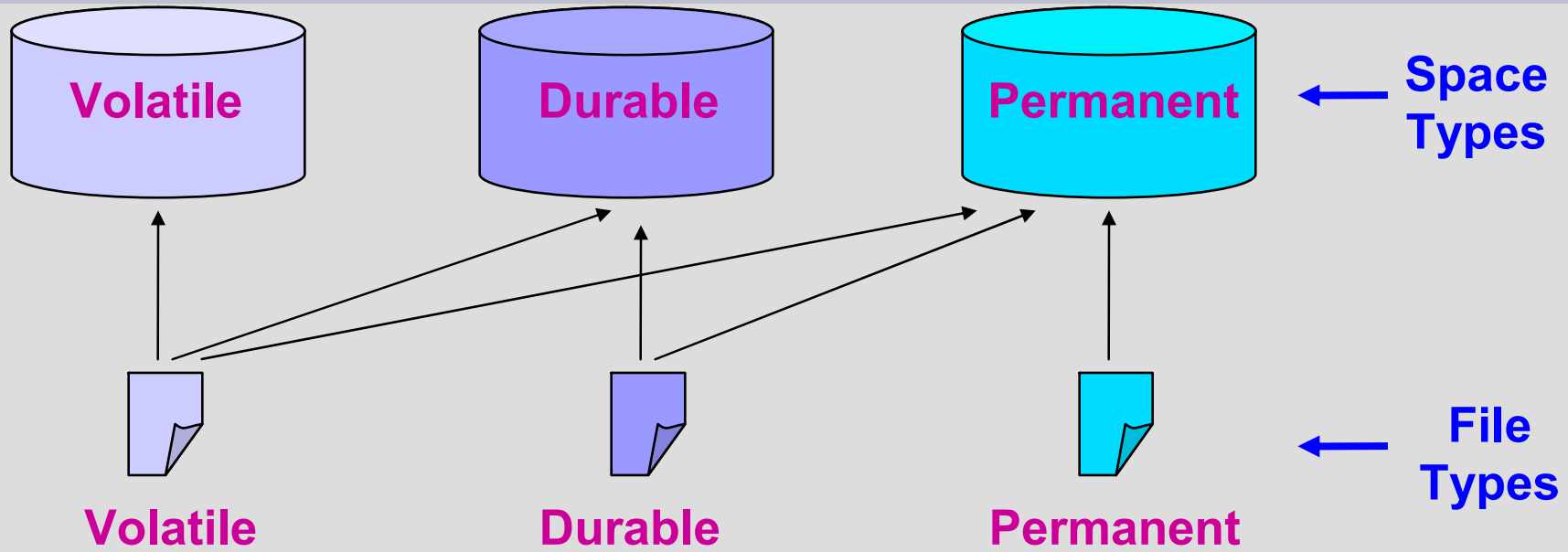
File Type / Property	Volatile	Durable	Permanent
Archived	No	No	Yes
Lifetime	Yes	Yes	No
AutoDeleted	Yes	No	No

SRM
Groups of Functions

SRM interface consists of the following groups of functions:

- Space Management Functions
- Data Transfer Functions
- Directory Functions
- Permission Functions
- Status Functions

SRM Interface Details

Space Management Functions
- SrmReserveSpace
- SrmReleaseSpace
- srmUpdateSpace
- srmCompactSpace
- srmGetSpaceMetaData
- srmChangeFileStorageType
- srmGetSpaceToken

Directory
- SrmMkdir
- srmRmdir
- srmRm
- srmLs
- srmMv

Data transfer functions
- srmPrepareToGet
- SrmPrepareToPut
- srmCopy
- SrmRemoveFiles
- srmReleaseFiles
- srmPutDone
- srmAbortRequest
- srmAbortFiles
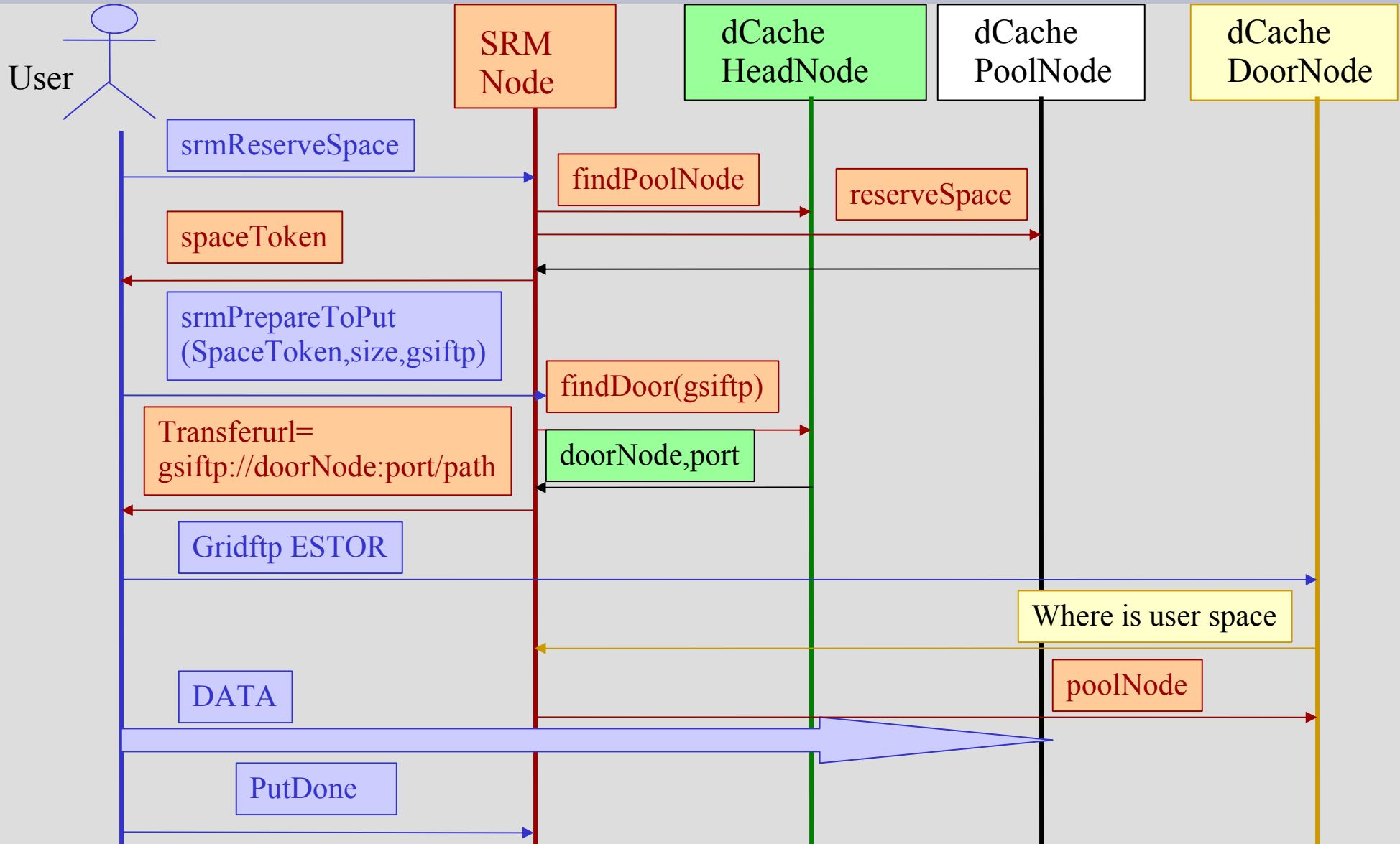- srmSuspendRequest
- srmResumeRequest

Status Functions
- srmStatusOfGetRequest
- srmStatusOfPutRequest
- srmStatusOfCopyRequest
- srmGetRequestSummary
- srmExtendFileLifeTime
- SrmGetRequestID

Permission
- srmSetPermission
- srmReassignToUser
- srmCheckPermission

Srm Example - srmPrepareToPut

Srm Example - Robust Replication

Anywhere

SRM-Client

HRM-COPY
(thousands of files)

FNAL

SRM
(performs writes)

SRM-GET (one file at a time)

GridFTP GET (pull mode)

CERN

SRM
(performs reads)

Enstore   dCache

cache   **CASTOR**

Network transfer

archive files

stage files
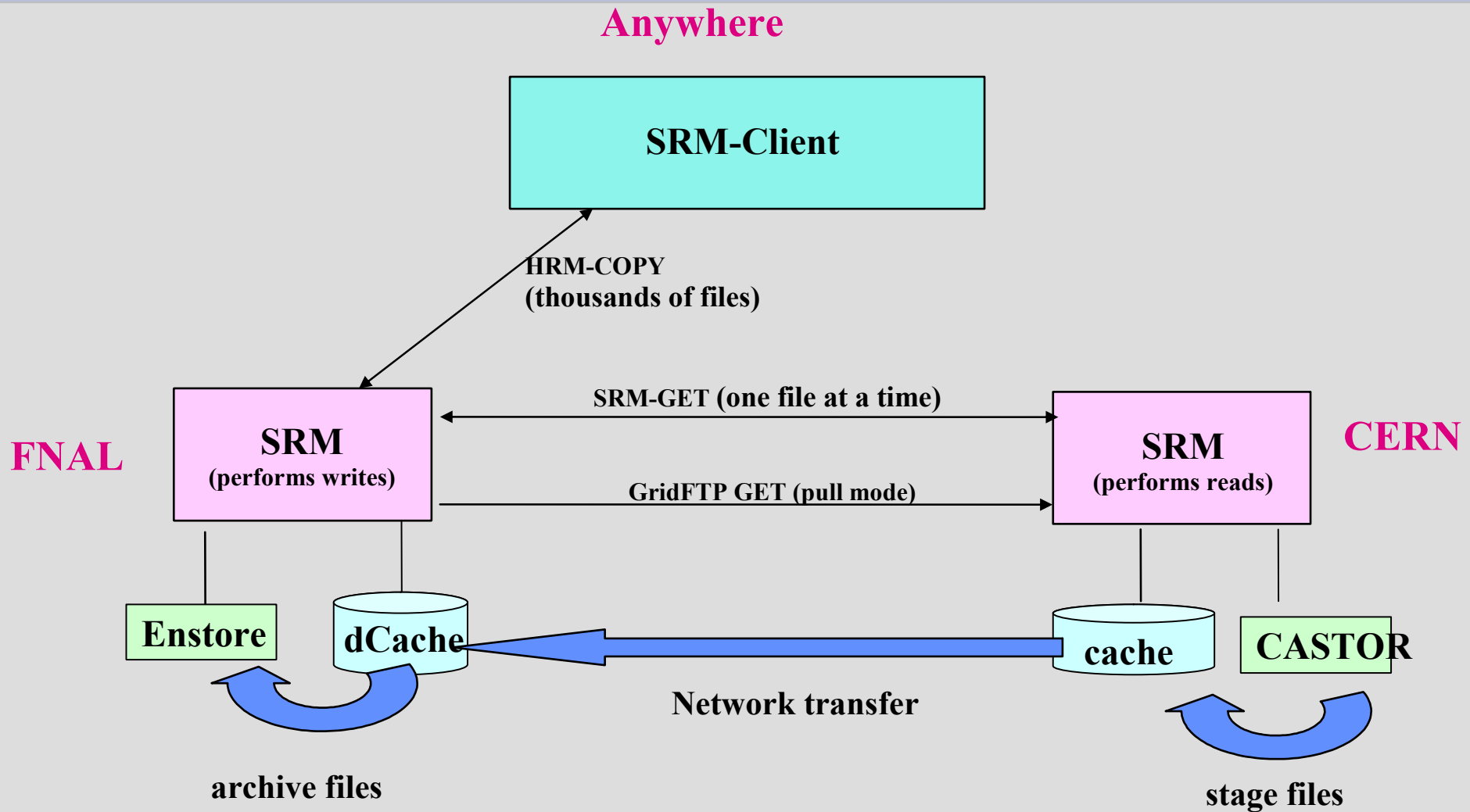
The sequence diagram of the SRM Copy Function performing "Copy srm://ServerB/file1 srm://ServerA/file1"

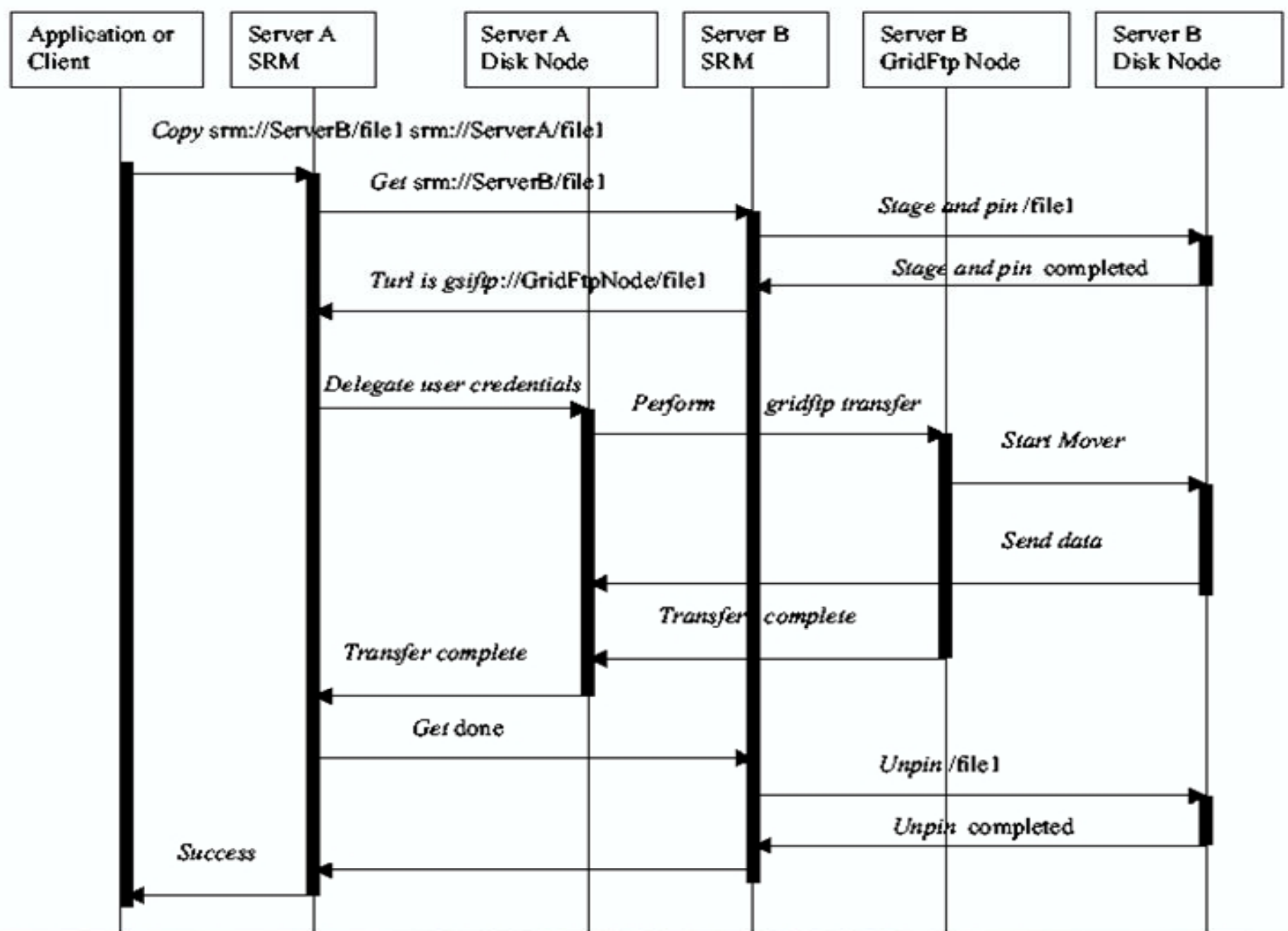# Grid File Access Library and SRM

Physics Application				
Grid File Access Library (GFAL)				
Replica Manager Client	SRM Client	Local File I/O	rfio I/O open() read() etc.	dCap I/O open() read() etc.

POSIX I/O →

RM Services — SRM Service ... rfio Service ... dCap Service

Wide Area Access ←

GridFTP Service ... MSS Service ... Local Disk

Role of SRM in Grid

COLLECTIVE 2 : SERVICES SPECIFIC TO APPLICATION DOMAIN OR VIRTUAL ORG.				
Request Interpretation and Planning Services	Workflow or Request Management Services	Application-Specific Data Discovery Services	Community Authorization Services	Consistency Services (e.g., Update Subscription, Versioning, Master Copies)

COLLECTIVE

COLLECTIVE 1: GENERAL SERVICES FOR COORDINATING MULTIPLE RESOURCES						
Data Transport Services	Data Federation Services	Data Filtering or Transformation Services	General Data Discovery Services	Storage Management (Brokering)	Compute Scheduling (Brokering)	Monitoring/ Auditing Services

RESOURCE:

File Transfer Service (GridFTP)	**Storage Resource Manager**	Data Filtering or Transformation Services	Database Management Services	Compute Resource Management	Resource Monitoring/ Auditing

CONNECTIVITY

Communication Protocols (e.g., TCP/IP stack)

Authentication and Authorization Protocols (e.g., GSI)

FABRIC

Networks

Mass Storage System (dCache/Enstore)
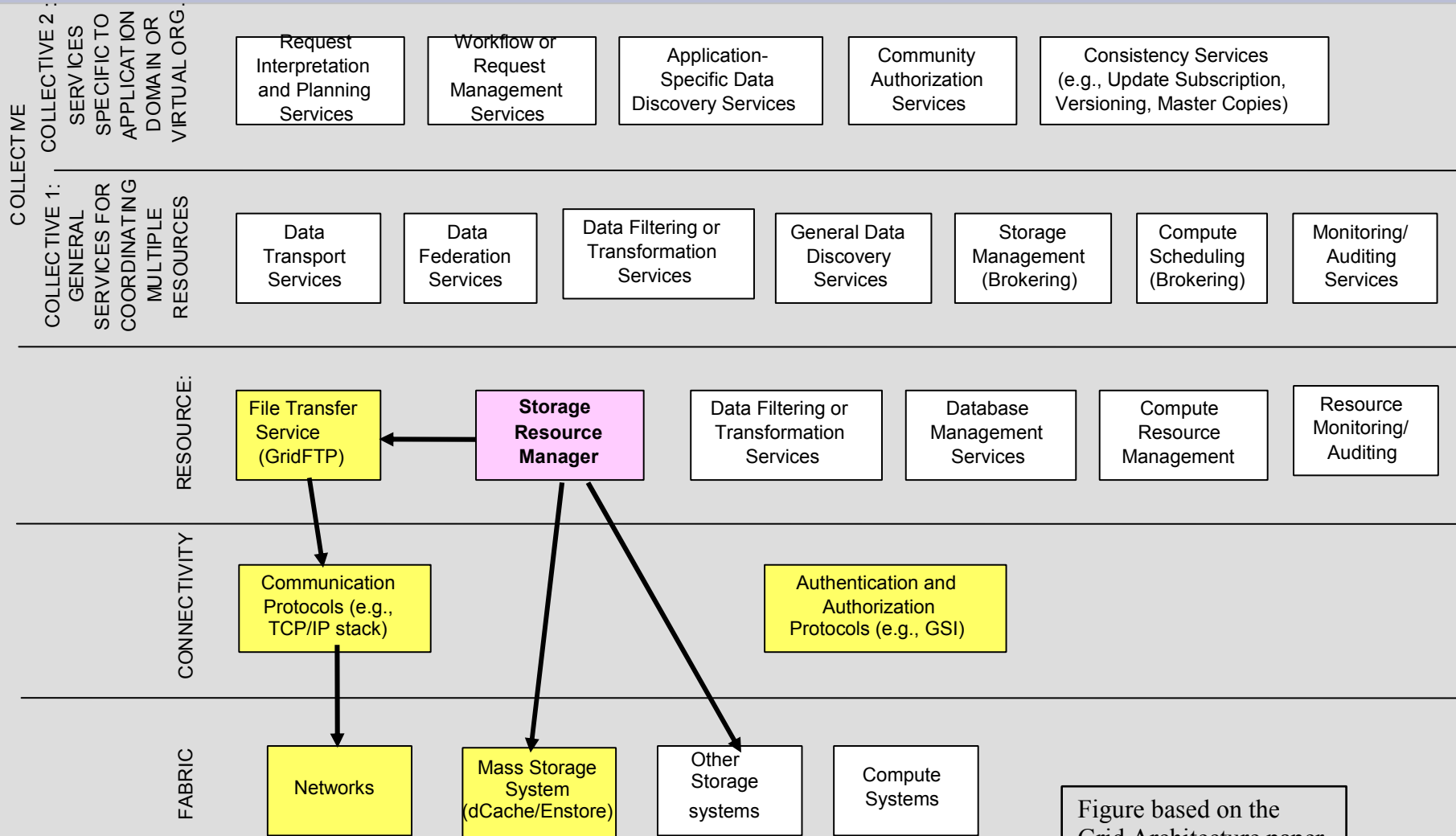
Other Storage systems

Compute Systems

Figure based on the Grid Architecture paper by Globus

US-CMS DATAGRID

Compact Muon solenoid (CMS) is an experiment at the lepton Hadron Collider (LHC) at CERN in Geneva, Switzerland.

- US-CMS is building data grid to facilitate physics data analysis at academic institutions across the united states.
- Data grid 3 tier architecture.
   - Tier 0 CERN, Geneva, Switzerland.
    - Tier 1 consists of 5 regional Centers, FERMILAB in Batavia, IL is a north American Center.
    - Tier 2 consists of 25 Centers, 5 of these are in north America.

SRM copy is used as a management protocol and reliable replication service for movement of data from tier 0 to tier 1 and from tier 1 to tier 2.

CMS SC2003 DEMO
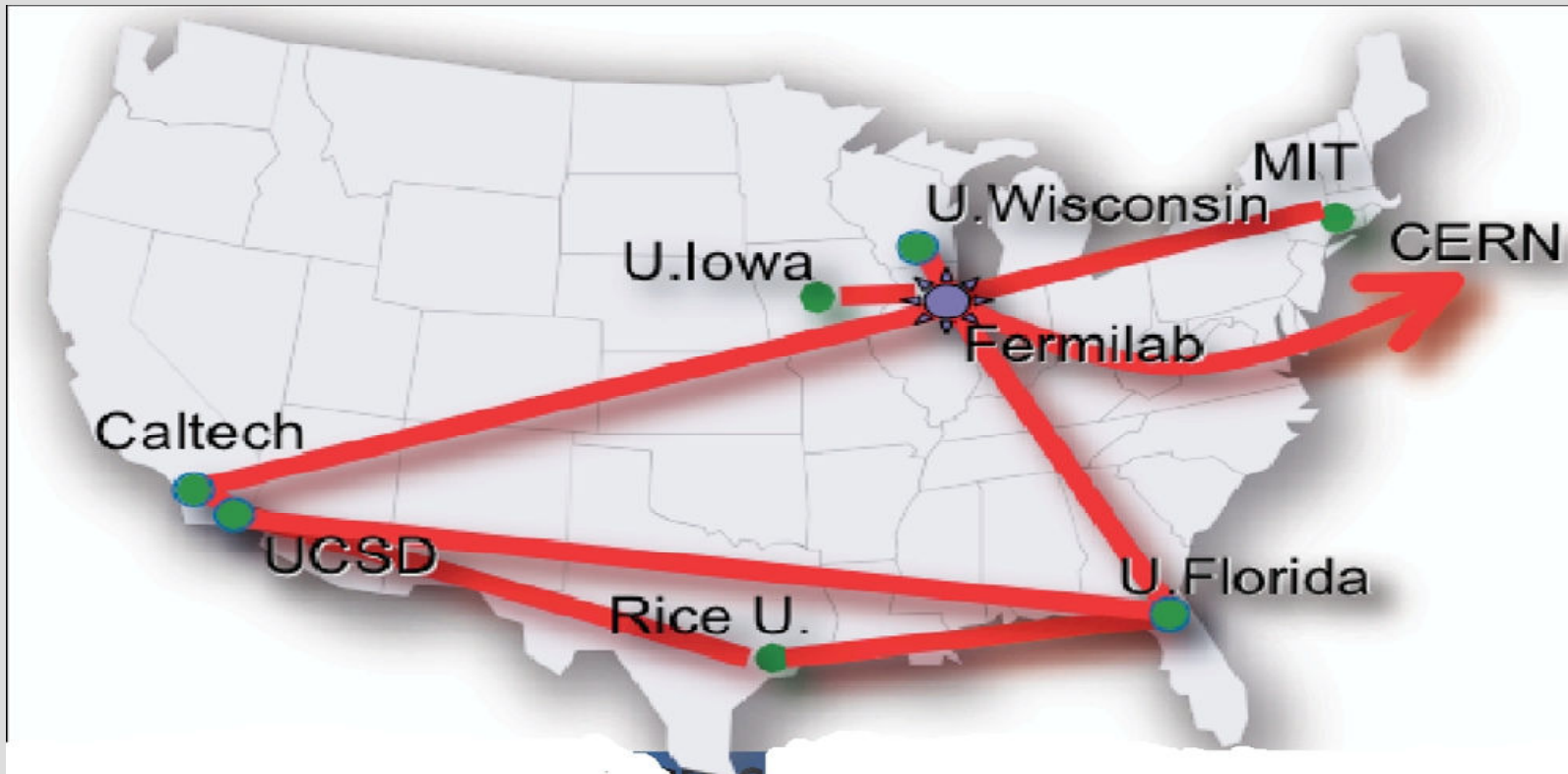

U.S. CMS Data and Storage Management (DCache, SRM)
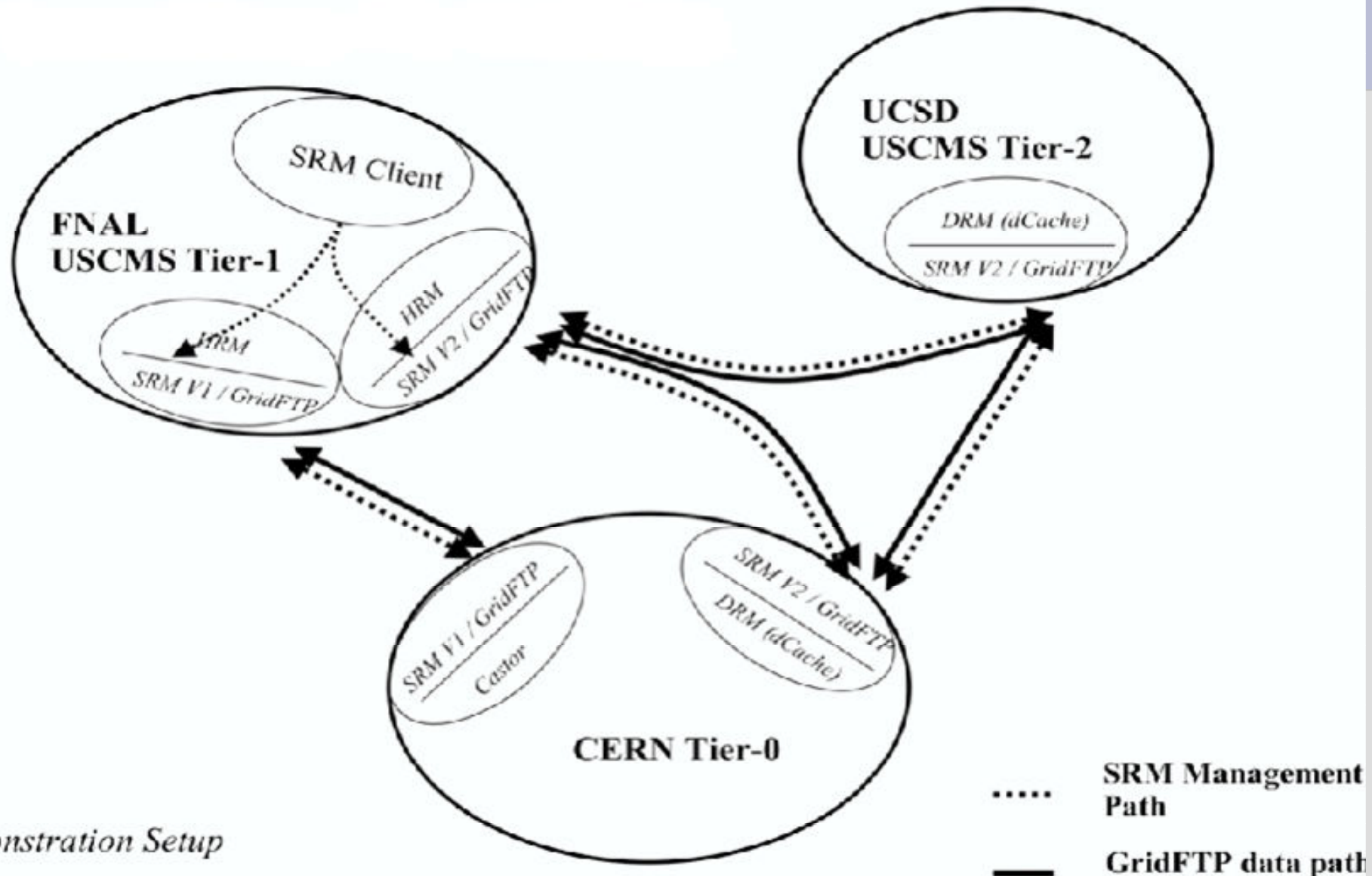
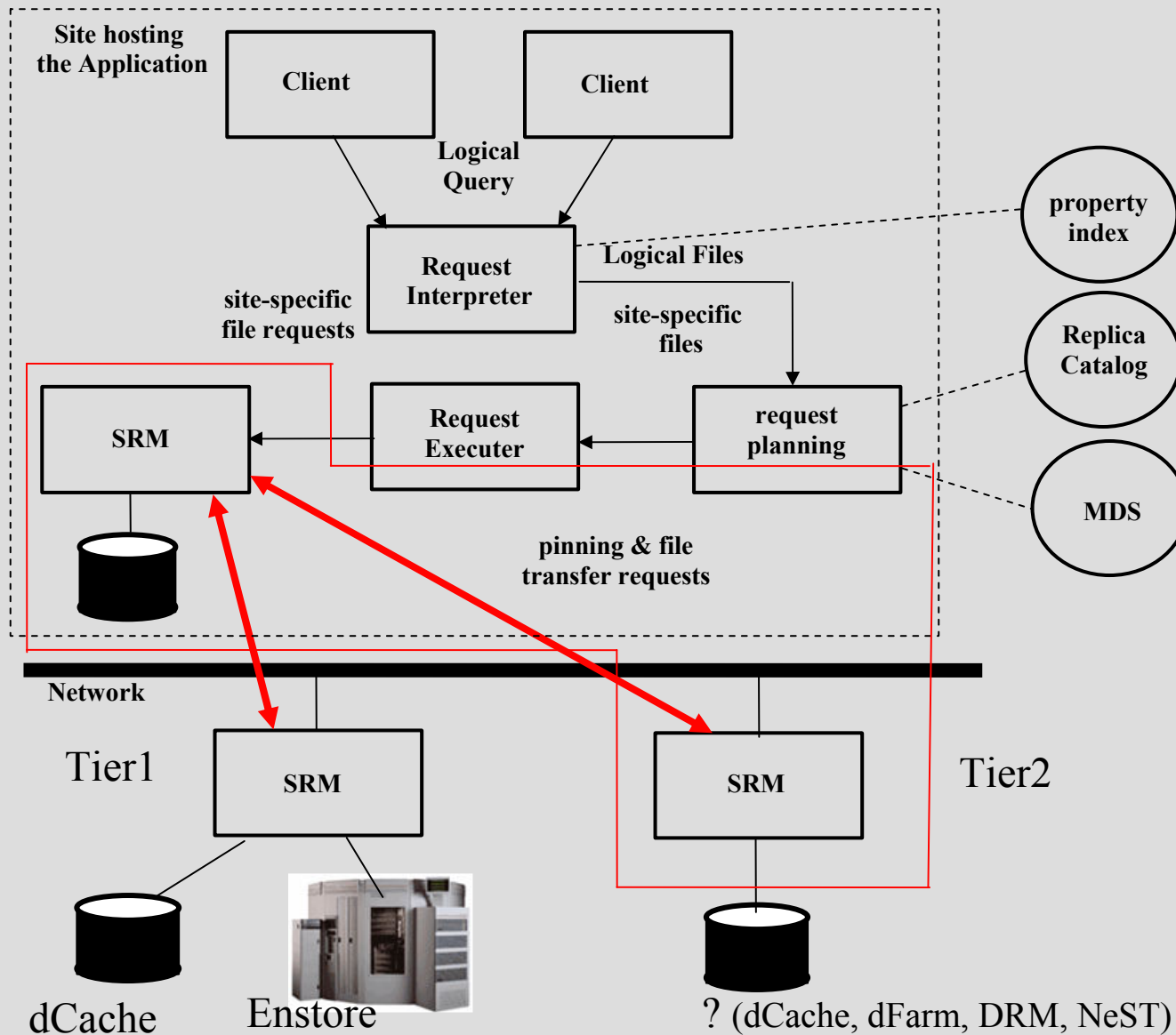Demonstration Setup

US-CMS Data Challenge 2004

- Event Reconstruction at the Tier-0 Center at 25Hz for the period of a month
- Transfer of raw and reconstructed data to distributed Tier-1 centers
- Access of events at distributed center for analysis-type applications

Client Access to CMS-DATA

Resources

- The Storage Resource Manager Collaboration Web Site http://sdm.lbl.gov/srm-wg/
- The Storage Resource Manager Interface Specification, version 2.1, Edited by Junmin Gu, Alex Sim, Arie Shoshani, LBL, http://sdm.lbl.gov/srm/documents/joint.docs/SRM.spec.v2.1.final.doc.
- Dcache documentation by Patrick Fuhrmann, DESY, Germany, www.dcache.org
- US-CMS http://www.uscms.org
- US-CMS Data Challenge 2004 http://www.uscms.org/s&c/dc04