Introduction to SRM-dCache

Ted Hesselroth

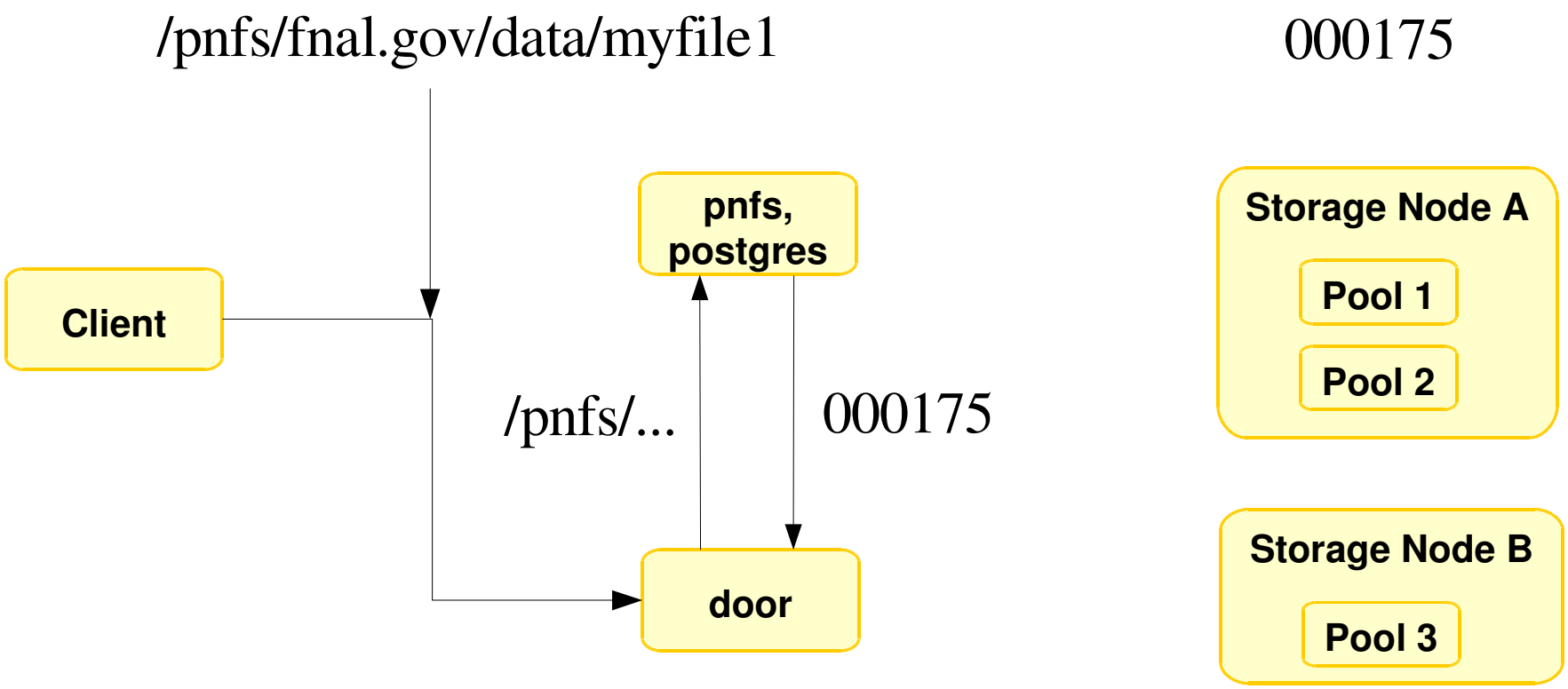
Fermilab

dCache History

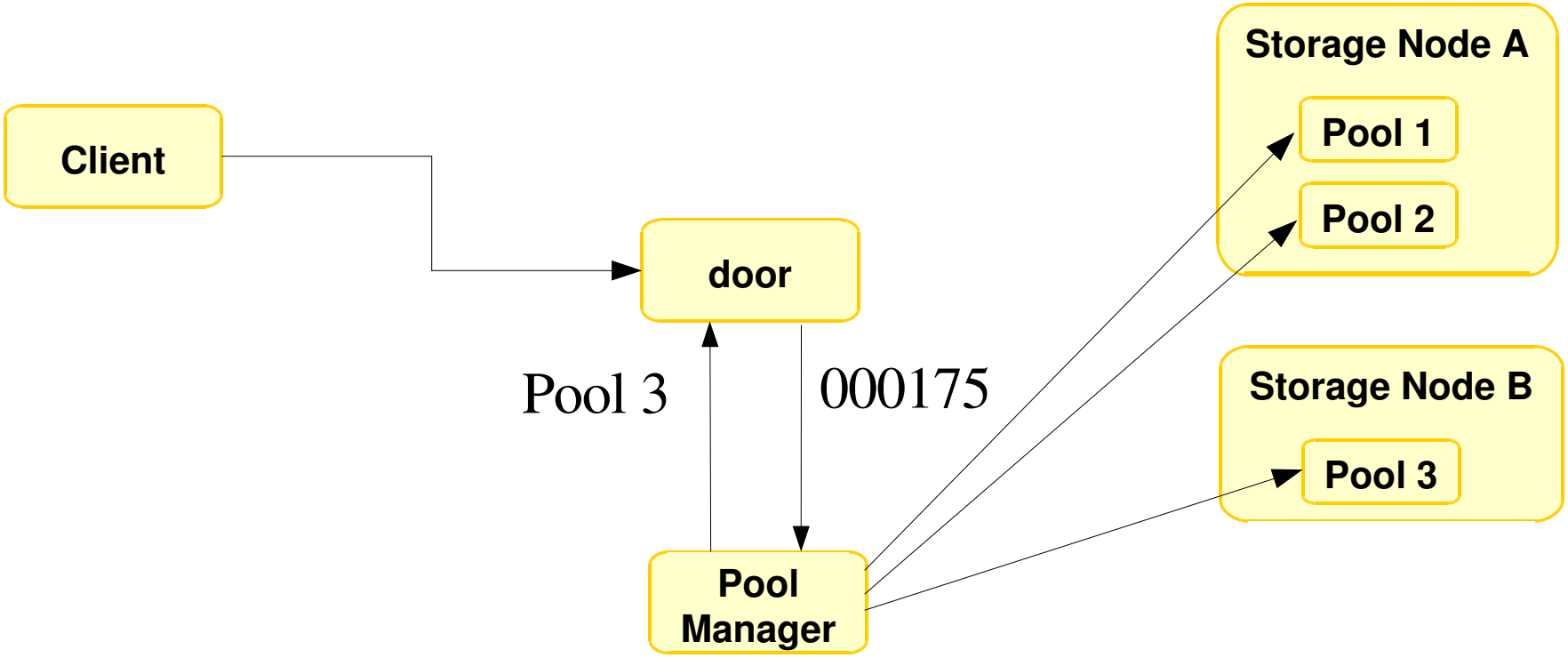
- Originally a cache for a tape system at DESY
  - Can also be used as a disk-only system
- Fermilab deployed for CDF, Dzero, Fermigrid, USCMS
  - Began collaboration in software development
  - DESY: 3 FTE, FNAL: 3 FTE + 2 FTE for VDT install
- Deployments world-wide ~40.
  - Used in many LHC sites
  - OSG Sites
- Mostly java, C library for client access

Distributed Storage

- Use of namespace in place of physical file location

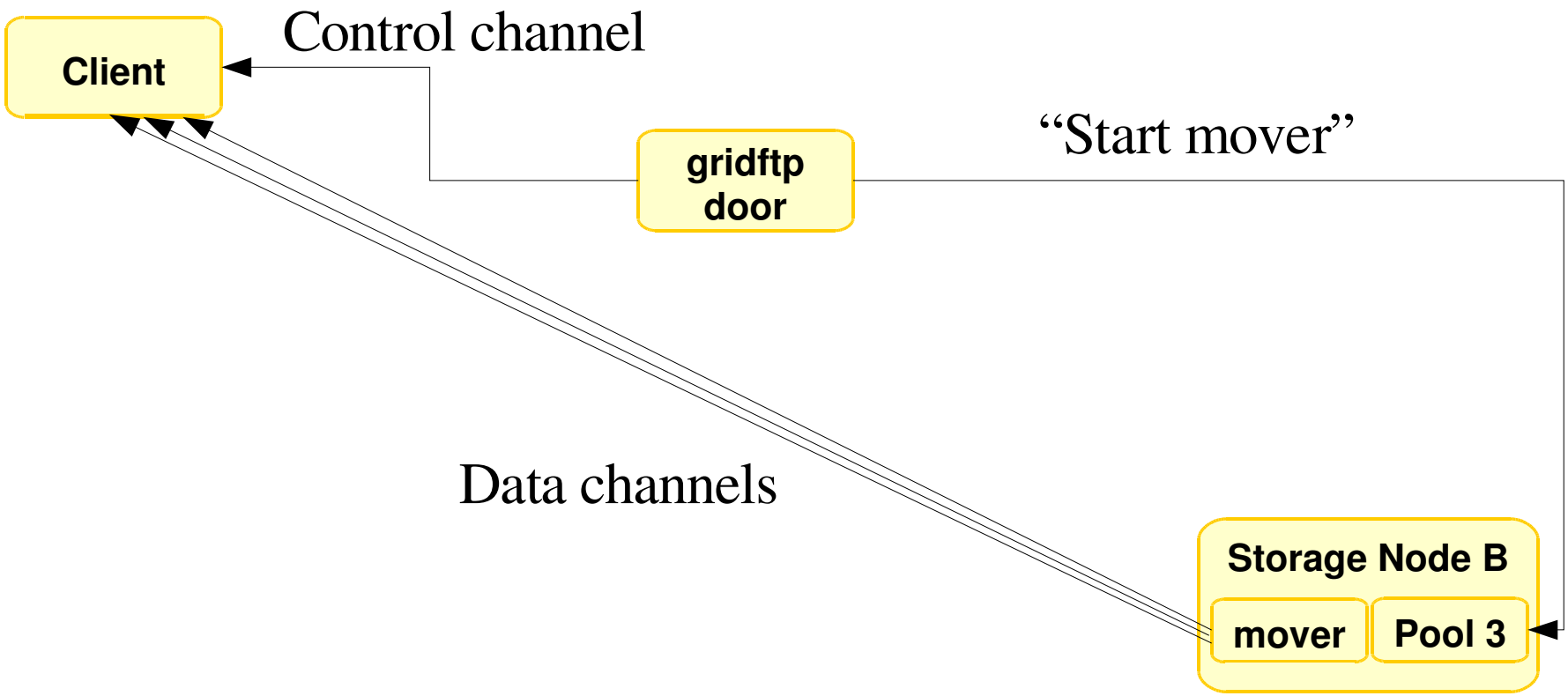


The Pool Manager



- Selects pool according to cost function
- Controls which pools are available to which users

The Gridftp Door



Other Doors

- dcap
  - Used in LAN
  - Provides posix-like IO
  - Security options: unauthenticated, x509, kerberos
- ftp
  - Security options: x509, kerberos
- xrootd

Storage Resource Manager

- Various Types of Doors
  - gridftp, dcap, gsidcap, xrootd, etc
- Need to address each service directly
- SRM is middleware between client and door
  - Web Service
- Selects among doors according to availabilty
  - Client specifies supported protocols
- Provides additional services

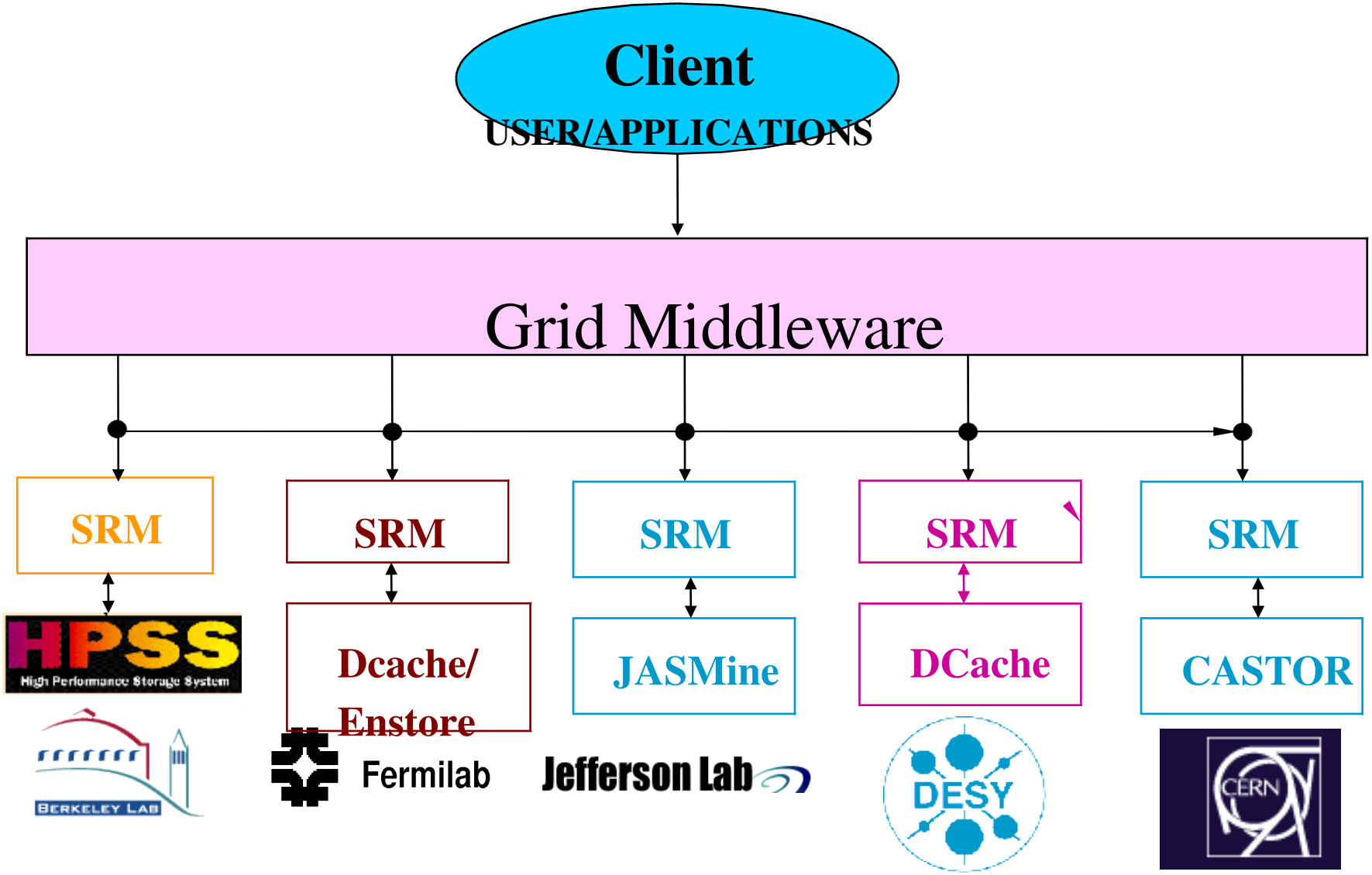
SRM Features – V2.2

- Space Reservation
- Queueing
- Status
- Monitoring
- Filesystem operations; ls and rm.
- Site-to-site transfers

SRM History

- Specification is product of SRM Working Group
  - LBNL, FNAL, CERN, JLAB, RAL, CNAF
- Defined as a specification for a storage interface
  - Common interface to various storage systems
    - FNAL – dCache
    - LBNL – DRM
    - etc
- FNAL: 2 FTE

SRM and Client Software



SRM-dCache clients

- SRM: srmcp, srmls, srmrm
- gridftp: globus-url-copy
- dcap: dccp
- Administration/Monitoring: web browser, ssh, cellspy
- Transfer services: Phedex, FTS

Fermilab SRM Clients

- srmcp srm://ligodata.uwm.edu:8443/pnfs/uwm.edu/data/testfile file:////tmp/test.tmp
- srmcp srm://storage.ligo.lsu.edu.:8443/pnfs/uwm.edu/data/file.01 srm://ligodata.uwm.edu:8443/pnfs/uwm.edu/data/testfile.01
- srmcp uses GSI x509 proxy
  - Can contain group and role attributes
- srmcp can specify for gridftp transfers
  - Port range
  - Number of streams
  - Buffer size, ...
- srmls, srmrm

The dcap library and dccp

- Provides posix-like open, create, read, write, lseek
  - int dc_open(const char *path, int oflag, /* mode_t mode */...);
  - int dc_create(const char *path, mode_t mode);
  - ssize_t dc_read(int fildes, void *buf, size_t nbytes);
  - ...
- dccp
  - dccp /pnfs/uwm.edu/data/testfile /tmp/test.tmp
  - dccp dcap://ligodata.uwm.edu/pnfs/uwm.edu/data/testfile /tmp/test.tmp

Monitoring – Disk Space

dCache ONLINE - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://dcache.rcac.purdue.edu:2288/usageInfo

TedEffort < SrmInternal < Twiki | TedWorkingNotes < SrmTroublesho... | dCache ONLINE | dCache ONLINE | SEARCH RESULTS...

Disk Space Usage

CellName	DomainName	Total Space/MB	Free Space/MB	Precious Space/MB	Layout (precious/used free)
dcache-00_1	dcache-00Domain	39936	111	34087	
dcache-00_2	dcache-00Domain	1126400	328343	787403	
dcache-00_3	dcache-00Domain	1126400	293425	802653	
dcache-00_4	dcache-00Domain	1126400	208103	896960	
dcache-00_5	dcache-00Domain	1126400	197682	907380	
dcache-01_1	dcache-01Domain	38912	22712	14180	
dcache-01_2	dcache-01Domain	1126400	802059	321706	
dcache-01_3	dcache-01Domain	1126400	404227	722172	
dcache-01_4	dcache-01Domain	1126400	771971	351422	
dcache-01_5	dcache-01Domain	1126400	643169	481972	
dcache-02_1	dcache-02Domain	38912	127	28819	
dcache-02_2	dcache-02Domain	1126400	228776	884238	
dcache-02_3	dcache-02Domain	1126400	294354	809192	
dcache-02_4	dcache-02Domain	1126400	284281	830713	
dcache-02_5	dcache-02Domain	1126400	268442	852888	
dcache-03_1	dcache-03Domain	38912	134	29645	
dcache-03_2	dcache-03Domain	1126400	305672	806907	
dcache-03_3	dcache-03Domain	1126400	218627	895758	

Find: remov Find Next Find Previous Highlight Match case

Done

Monitoring – Service Status

dCache ONLINE - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://dcache.rcac.purdue.edu:2288/cellInfo

TedEffort < SrmInternal < Twiki | TedWorkingNotes < SrmTroublesho... | dCache ONLINE | dCache ONLINE | SEARCH RESULTS...

Services

CellName	DomainName	Requests Pending	Threads	Ping	Creation Time
CopyManager	srm-dcacheDomain	0	1	208 msec	10/27 14:40:14
DCap	doorDomain	0	4	103 msec	10/27 14:22:34
DCap-gsi-dcache	gsidcap-dcacheDomain	0	4	105 msec	10/27 14:23:16
DCap-gsi-dcache-00	gsidcap-dcache-00Domain	0	4	198 msec	09/28 08:14:33
DCap-gsi-dcache-01	gsidcap-dcache-01Domain	0	4	189 msec	11/03 16:36:41
DCap-gsi-dcache-02	gsidcap-dcache-02Domain	0	4	105 msec	10/31 13:01:49
DCap-gsi-dcache-03	gsidcap-dcache-03Domain	0	4	199 msec	10/27 16:33:13
GFTP-dcache	gridftp-dcacheDomain	OFFLINE			
GFTP-dcache-00	gridftp-dcache-00Domain	0	9	103 msec	09/28 08:14:26
GFTP-dcache-01	gridftp-dcache-01Domain	0	4	188 msec	11/03 16:36:35
GFTP-dcache-02	gridftp-dcache-02Domain	0	4	149 msec	10/31 13:01:42
LoginBroker	httpdDomain	0	2	14 msec	10/27 14:22:47
PinManager	utilityDomain	0	2	208 msec	10/27 14:22:55
PnfsManager	pnfsDomain	0	51	198 msec	10/07 19:18:20
PoolManager	dCacheDomain	0	4	147 msec	10/27 14:22:19
SRM-dcache	srm-dcacheDomain	0	2	205 msec	10/27 14:23:23
SpaceManager	utilityDomain	0	2	147 msec	10/27 14:22:55
dcache-00_1	dcache-00Domain	0	10	70 msec	10/29 13:36:58

Find: remov Find Next Find Previous Highlight Match case

Done

SRM Watch – Current Transfers

Copy, Put, Get Requests Start Page - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://cmsdcam3.fnal.gov:8080/srmwatch/servlet/StartPage

TedEffort < SrmInternal < Twiki | TedWorkingNotes < SrmTroublesho... | Copy, Put, Get Requests Start Pa... | dCache ONLINE | SEARCH RESULTS...

Select Copy, Put or Get below

Copy, click here for complete active listing

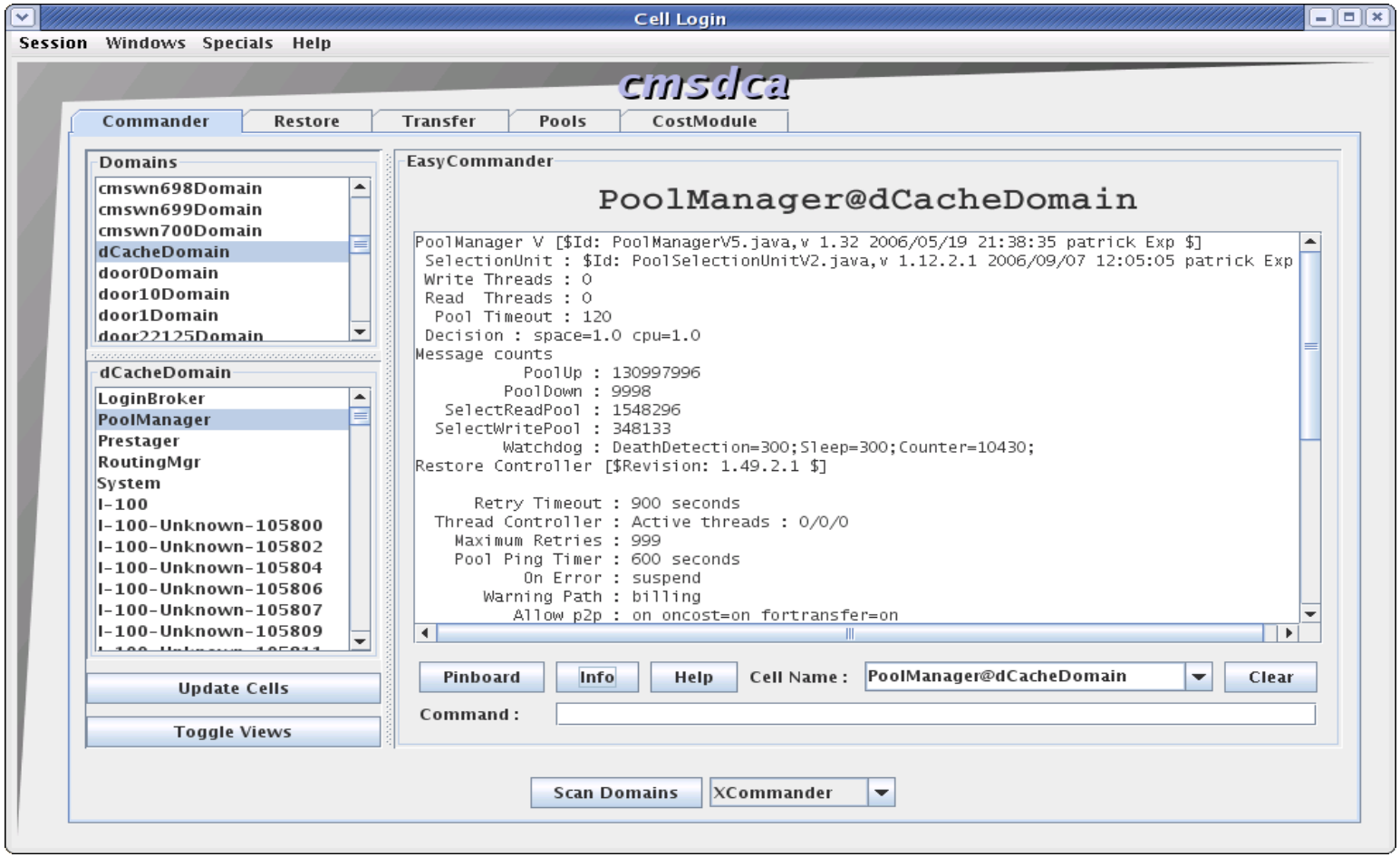
ID	Creation Time	Copyfilerequests	Current State	Credential Name
-2143016844	Fri Nov 03 16:36:11 CST 2006	1	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Bockjoo Kim 740786
-2143018621	Fri Nov 03 15:48:02 CST 2006	2	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Yujun Wu 167397
-2143240253	Fri Oct 27 11:54:33 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143242069	Fri Oct 27 10:54:32 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143242079	Fri Oct 27 10:54:30 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143242117	Fri Oct 27 10:53:50 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143242155	Fri Oct 27 10:51:45 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143244700	Fri Oct 27 09:36:40 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143244725	Fri Oct 27 09:36:13 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143244740	Fri Oct 27 09:35:52 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143246230	Fri Oct 27 08:56:56 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143254170	Fri Oct 27 03:58:45 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143254195	Fri Oct 27 03:57:15 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143255330	Fri Oct 27 02:35:42 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143255354	Fri Oct 27 02:35:01 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143255374	Fri Oct 27 02:34:37 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368
-2143257069	Fri Oct 27 00:19:54 CDT 2006	4	AsyncWait	/DC=org/DC=doegrids/OU=People/CN=Dimitri Bourilkov 367368

Find: remov Find Next Find Previous Highlight Match case

Done

Cellspy - Commander

- Select dCache cell from list
- Status and command windows



Major dCache Internal Services

Service	Function
Location Manager	Supplies ip address and port of dCache domain
dCache	Routes messages between dCache services
admin	Allows login to running services
pnfsManager	Queries pnfs service
pinManager	Prevents modification of files during transfer
replicaManager	Maintains replica count of files on pools
Billing	Maintains records and statistics of all transactions
http	Provides information on state of dCache
gPlazma	Authorizes users

Generic USCMS Tier2 Installation Example

- 1 admin node
- 1 pnfs node
- 1 srm node
- 3 dcap + grid nodes
- pools on worker nodes or RAID
- 30 TB
- ~20 MB/sec

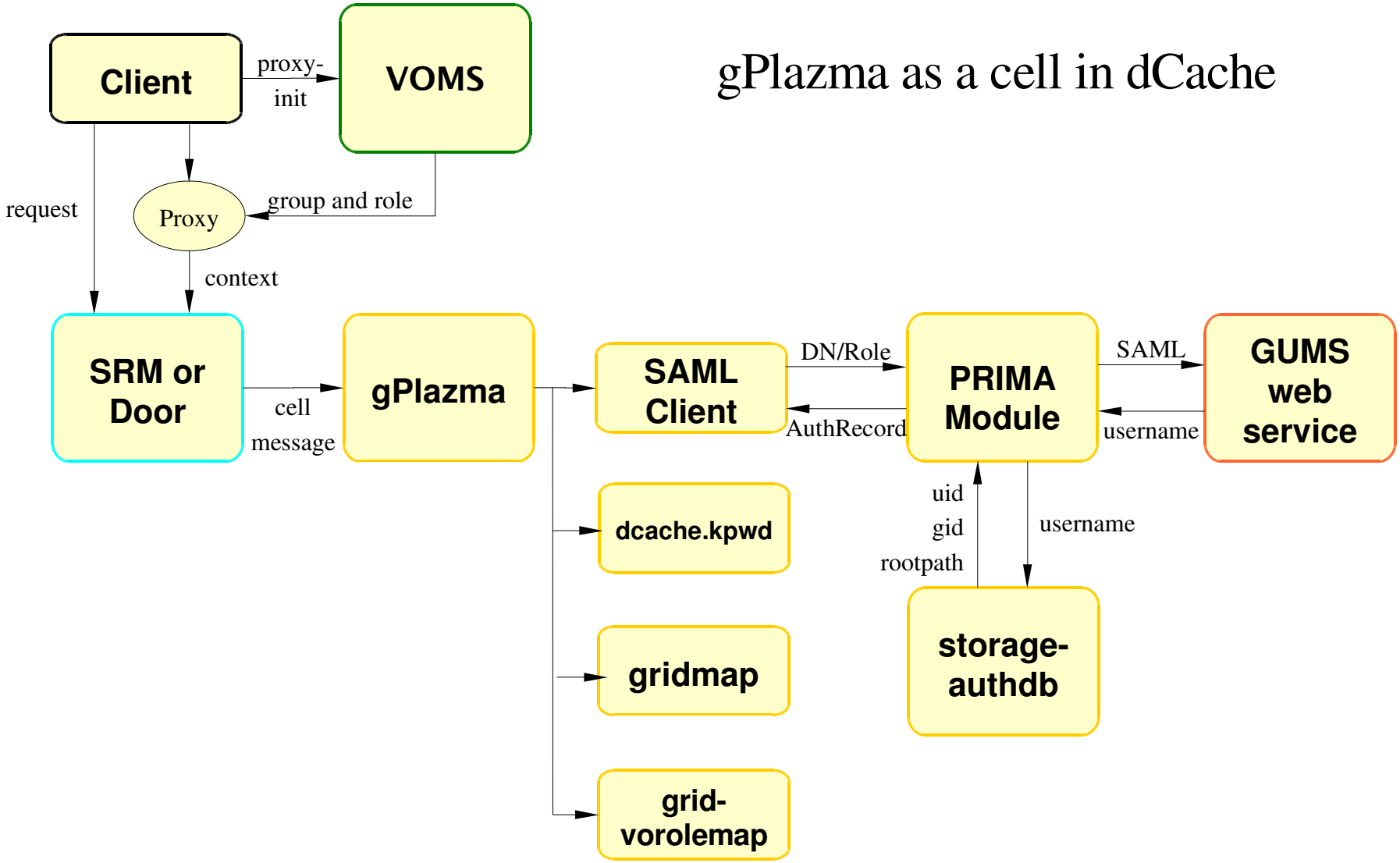
Private Networks and Firewalls

- Gridftp doors must be on public network
- Pools may be on private network
  - Gridftp Adapter for Local <-> SRM transfers
  - SRM <-> SRM may be possible
- Firewalls
  - Ports and port ranges may be specified

Authorization

- GSI and kerberos supported
- Host certificates on SRM, gridftp, gsidcap, gPlazma
- Client proxies are delegated within dCache
  - Allows SRM to act as a client
- gPlazma allows choice of authorization methods
  - Supportes role-based authorization
  - Can use GUMS server for central administration
    - Storage obligations added to GUMS permit decision

gPlazma Authorization



VDT Installation

- Install scripts and necesary RPMs available as a package
- Multinode installation
- Maximize automation
- Input
  - Number of nodes, install script suggests service distribution
  - Hostnames
  - Pool paths and sizes
- Same setup file and install script used on all nodes