

FNAL dCache Admin Notes I

Version 1.2

Created: 11 March 2005

Modified: 25 April 2005

Rob Kennedy

- 1) Terms and Actors
- 2) Diagram of a User Interaction
- 3) FNAL dCache Configurations – *(subject to change)*
 - a) CDFDCAT: overall, door ports
 - b) FNDCA: overall, door ports
 - c) CDFDCA: overall, door ports, sub-caches
 - d) CMSDCA: overall, door ports
- 4) Responsibilities
- 5) Documentation Status

Terms and Actors: Dictionary

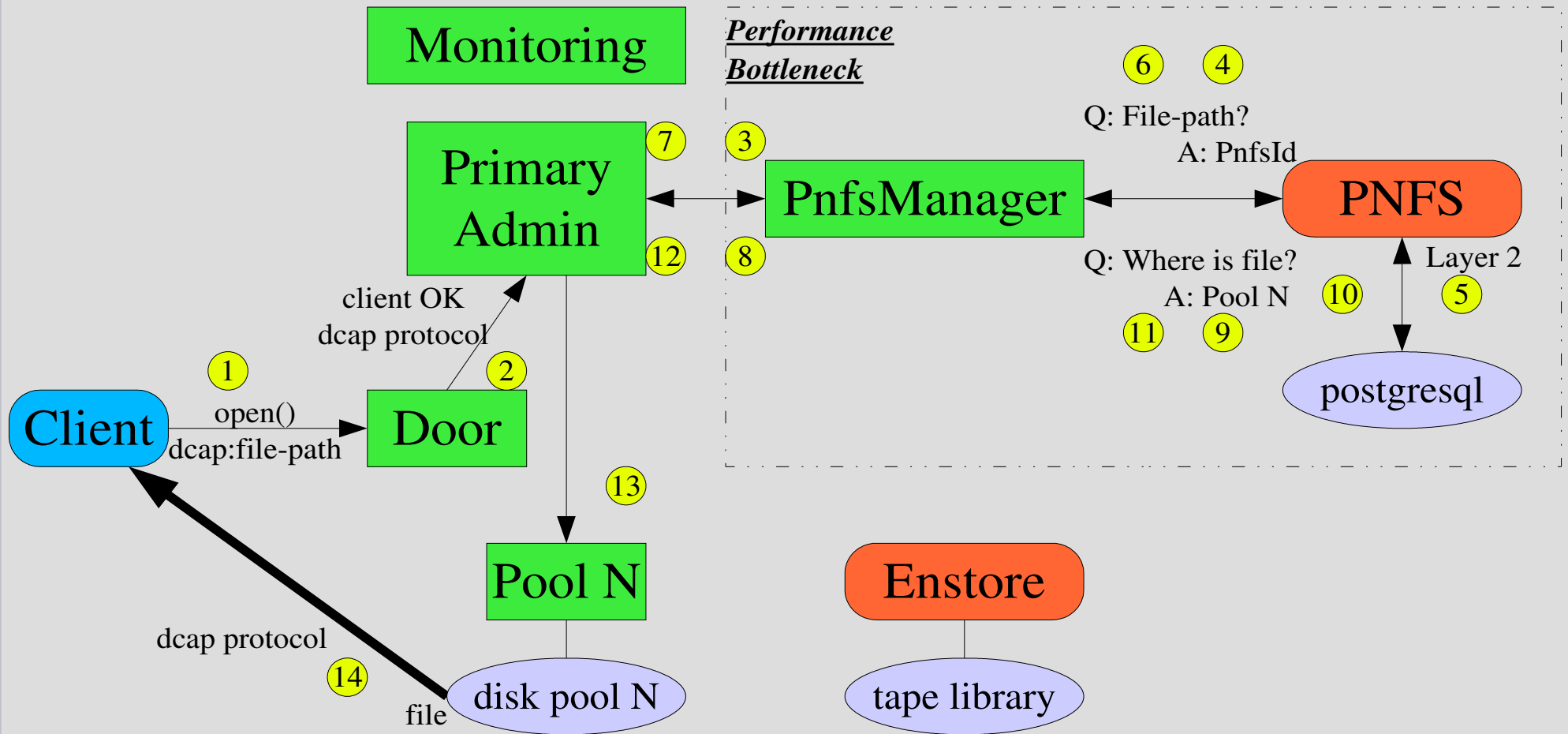
- Admin node: a host to dCache services
- Primary Admin node: SPOF host (single-instance services)
- Door node: a host to dCache “portals” or **doors** for clients
- Door: authenticate and authorize a client for a protocol
- Monitor node: a host performing various monitoring tasks
- Pool node: a host to dCache data file storage
- Pool: a virtual data partition in the dCache storage space
- Mover: process receiving or sending data to client.
- Restore: restoring data from MSS into cache.
- Store: storing data in cache to the MSS.
- P2P: pool-to-pool file transfers inside of a dCache system.
- PNFS: name space and replica catalog for dCache & Enstore

Terms and Actors: Service Structure

- Cell: smallest unit of dCache service. Each cell has the same infrastructure (cell nucleus), but supports varying functionality.
- Domain: a collection of cells running in one Java virtual machine. Cells in a domain not necessarily related in function.
- LocationManager: Singleton service. Knows about all domains, helps setup message *tunnels* amongst them. Abbreviated “lm”.
- PnfsManager: Singleton dCache *interface* to PNFS, is not the manager of PNFS. **Horribly confusing name!** dCache-PNFS interaction is known to be a performance bottleneck overall.
- SRM: Storage Resource Manager – Uniform Grid interface
- FNAL dCache: core dCache service, SRM, PNFS, FNAL-added monitoring plots, Dcap client library, and other misc.

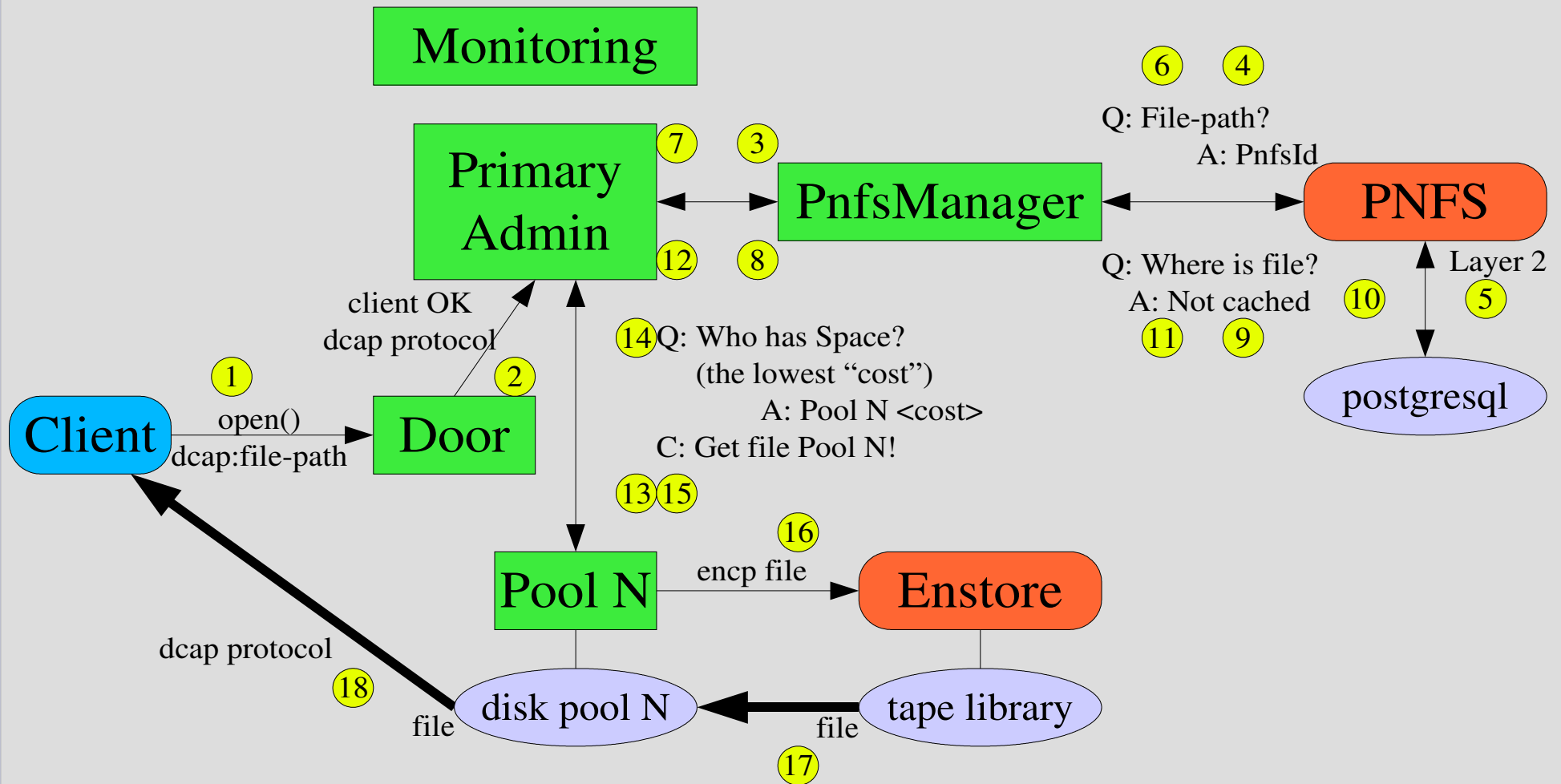
Client Reads a File in Cache

(simplified)



Client Reads a File not in Cache

(simplified)



CDFDCAT Configuration

CDF Test dCache

- cdfdcat (fcdfdcache6): Primary admin node, door node, and monitoring node. Runs PnfsManager. (Yes, it runs there. ;^)
- cdfensrv1: same PNFS service as CDF production dCache.
- Pool nodes: fcdfddata008, fcdfddata112. Each with 3 pools.
- Doors: intended to test CDF and other client interactions
 - 10 kerberized dcap (DCapK<n>) -> reg. user krb5 principle
 - 2 X509 “Grid” dcap (DCapG0,1) -> register user X509 DN
 - 1 weakly authenticated FTP (WFTP0) - read-only
 - WFTP requires use of separate registered password
 - 1 kerberized FTP (KFTP0) -> register user krb5 principle
 - 1 Grid FTP (GFTP0) -> register user X509 DN
 - 1 SRM interface (srm_1) -> register user X509 DN

CDFDCAT Doors

CDF Test dCache

- 10 kerberized dcap – cdfdcats: 26125, 26136-26144
 - DCapK<n>
- 2 Grid dcap – cdfdcats: 26525, 26536
 - DCapG<n>
- 1 weakly auth. FTP – cdfdcats: 26126
 - WFTP0
- 1 kerberized FTP – cdfdcats: 26127
 - KFTP0
- 1 Grid FTP – cdfdcats: 2811
 - GFTP0
- 1 SRM interface – cdfdcats: 8443
 - srm_1

FNDCA Configuration

STKEN/public dCache

- fndca (stkendca3a): Primary admin, door, and monitoring node.
- stkensrv1: PnfsManager (“pnfsDomain”). Interface to PNFS.
- Pool nodes: stkendca4a-12a, 5a is dead. 2-3 pools each, with some read-from, others write-into. Some pools currently only serve certain KTeV file families, esp. pools on stkendca8a.
- Doors: may add more dcap doors or logins/door soon
 - 2 unsecured dcap (DCap<n>) - read-only
 - 2 kerberized dcap (DCapK<n>)
 - 1 weakly authenticated FTP (WFTP0) - read-only
 - 1 kerberized FTP (KFTP0)
 - 1 Grid FTP (GFTP0)
 - 1 SRM interface (srm_1)

FNDCA Doors

STKEN/public dCache

- 2 unsecured dcap – fndca: 24125, 24136
 - Dcap<n>, where n=0,1
- 2 kerberized dcap – fndca: 24725, 24736
 - DcapK<n>, where n = 0,1
- 1 weakly auth. FTP – fndca: 24126
 - WFTP0 (now called WFTP, will change...)
- 1 kerberized FTP – fndca: 24127
 - KFTP0 (now called GKFTP, will change...)
- 1 Grid FTP – fndca: 2811
 - GFTP0 (now called GFTP, will change...)
- 1 SRM interface – fndca: 8443
 - srm1

CDFDCA Configuration

CDF Production dCache

- cdfdca (fcdfdcach1): Primary admin node
- cdfensrv1: PnfsManager (“pnfsDomain”) Interface to PNFS.
- cdfdca1,2,3 (fcdfdcach5,2,3): 3 door nodes
- cdfdcam (fcdfdcach4): monitor node
- Pool nodes: fcdfdata<nnn>. ~ 65 file servers, 3 pools each.
 - 7 different **sub-caches**. Total of > 150 TB data space.
- Doors: On --**each**-- of the 3 door nodes:
 - 10 kerberized dcap (DCapK00-09,10-19,20-29)
 - 400 logins max per dcap door since on dedicated hosts.
 - 1 WFTP<n>, 1 KFTP<n>, 1 GFTP<n>, where n=1,2,3
 - SRM interface active only on cdfdca1 (srm1).
 - “OFFLINE” state for other 2 SRMs to be cleaned up

CDFDCA Doors

CDF Production dCache

- 30 kerberized dcap doors total
 - DCap00-09 – cdfdca1: 25125, 25136-25144
 - DCap10-19 – cdfdca2: 25145-25154
 - DCap11-29 – cdfdca3: 25155-25164
- 3 weakly auth. FTP – cdfdca1,2,3: 26126
 - WFTP1,2,3
- 3 kerberized FTP – cdfdca1,2,3: 26127
 - KFTP1,2,3
- 3 Grid FTP – cdfdca1,2,3: 2811
 - GFTP1,2,3
- 1 SRM interface – cdfdca1: 8443
 - srm1

CDFDCA Sub-Caches

CDF Production dCache

- “Sub-caches” are portions of a dCache system that exclusively serve a predefined set of file families for read or for write.
- Golden (GoldenPools)
 - Semi-static storage of high priority Physics data sets
- Big Buffer (BigBufferPools)
 - Large data sets. Protects others from being cycled out.
- Little Buffer (LittleBufferPools)
 - Severely limits user access to deprecated data sets
- (2) Rata Data (RawStreamAPools, RawOtherStreamPools)
 - Like Big Buffer, but for raw data. Stream A is high priority.
- General Read Cache, Write Cache (readPools, writePools)
 - Everything else for reading. Everything for writing.

CMSDCA Configuration

CMS Production dCache

- Classic dCache + **“Resilient”** dCache – a new kind of dCache
- New: ReplicaManager; Speed: PnfsManagerV3, replica catalog
- cmsdca (cmsdcdr2): Primary admin node
- cmspnfs1: PnfsManager (“pnfsDomain”) Interface **to** PNFS.
- cmsdca1 (cmsdcdr1), cmsdca2 (cmsdcdr3): door nodes
- cmsdcam (cmsdcmon1): monitor node
- Pool nodes: RAID file servers **and farm worker nodes**.
- Doors: 7 unsecured dcap (DCap<n>) - read-only
 - 1 X509 “Grid” dcap (DCapG0)
 - 1 WFTP
 - 2 GFTP<n>
 - SRM interface (srm_2)

CMSDCA Doors

CMS Production dCache

- 7 unsecured dcap – cmsdca1: 24125, 24136-24141
 - DCap0-7
- 0 kerberized dcap
- 1 Grid dcap – cmsdca1: 24525
 - DCapG0
- 1 weakly auth. FTP – cmsdca1: 24126
 - WFTP
- 0 kerberized FTP
- 2 Grid FTP – cmsdca1: 2811, cmsdca2: 2811
 - GFTP1,2
- 1 SRM interface – cmsdca2: 8443
 - srm_2

Responsibilities

- CDFDCAT: Run2-sys, CDF-DH, devs. No ISA responsibility.
-
- FNDCA: ISA responsibility. Devs consult on configuration, optimization, experiment liaison work, esp. as Grid use grows.
-
- CDFDCA: Split. CDF filters out user “non-dCache” requests.
 - H/w and OS – CDF buys. Run2-sys installs, prepares, and admins.
 - dCache operations – CDF-DH helps, ISA does rest, devs consult
 - dCache config,liaison – CDF-DH helps, ISA does rest, devs consult
-
- CMSDCA: Critical service challenge has begin. CMS-T1 asks:
ISA: investigate posted encp errors, Devs: watch, restart pools.

Admin Documentation Status

- “Running dCache”: Linked to at the bottom of each dCache web page now. Was missing, but now is back everywhere.
- ISA's own: Mostly good for FNDCA (except PnfsManager site). <http://www-isd.fnal.gov/ISA/D0enManual.html> “dCache link”.
- Next Gen Dev-Admin doc: Draft to be presented in later talks: what accounts to use, what machine, what pathname, etc.
- dCache Web Pages: being updated, broken links removed, adjusting per exp't & **ISA** feed-back. CDF's first changes done.
- dCache Primary Info: schedule, #s posted like Enstore's are.
-
- Procedures: must be overhauled. Downtime risk due to errors.
- HelpDesk guide: needs to be **CREATED**. AFAIK, nothing now.