

# dCache workshop IN2P3 experience

Lionel Schwarz <[schwarz@cc.in2p3.fr](mailto:schwarz@cc.in2p3.fr)>

**CC-IN2P3**

Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules



# Overview

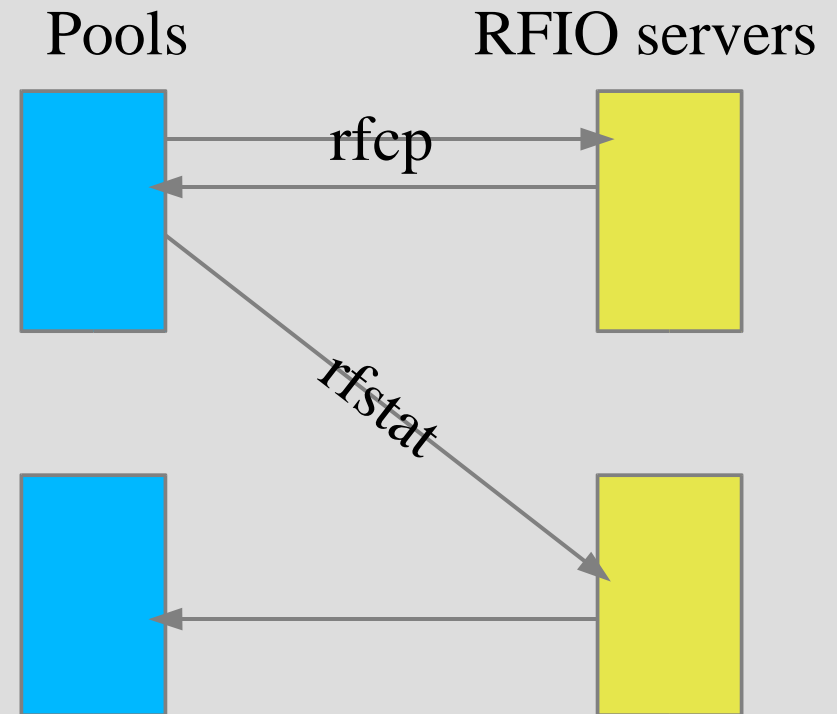
- NOT in production, but in tests since 2004
- HSM backend (HPSS)
  - HPSS link written from BNL's work
- Manpower: 1 FTE, 2 at the end of the year

# Current infrastructure

- The current dCache infrastructure is for SC only
- Installation not driven by LCG yet (october)
- Hardware
  - Head node: PIII, 1.4GHz, 1GB RAM
  - Pool nodes: bi-xeon 3GHz, 2GB RAM serving 6.5 TB disk
- Software
  - Head node: SL 3.0.3, Globus 2.4, Java 1.5.0\_01, Postgresql 7.4.6, dcache-v1.2.2-7-3
  - Pool nodes: SL 3.0.3, Globus 2.4, Java 1.5.0\_01, dcache-v1.2.2-7-3

# HPSS link

- Pools use RFIO commands to access to HPSS
  - “rfcp”: copy to/from
  - “rfstat”: stat a file
- Head node uses RFIO to delete files in HPSS when deleted in dCache
  - “rfrm”: delete a file

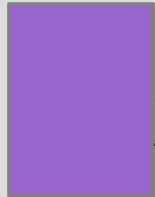


# Copy to/from HPSS

- “rfcp” and “rfstat” have to be run on behalf on the local user (known by RFIO)
  - Use “su” in the copy script
  - Issue when staging a file: the pool data directory has to be world-writable
- The StorageClass defines:
  - The RFIO server to be used with “rfcp” and “rfstat”
  - The “Class Of Service” (COS) to be used in HPSS
- All files (for a VO) are stored in HPSS in a single directory
  - Plan to use sub-directories to avoid a high number of files in a single directory

# Example : CMS user

Client



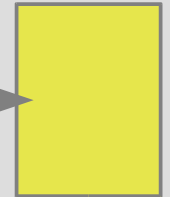
/pnfs/in2p3.fr/data/cms/tapes/

Pool



export RFCP\_HPSSCOS=18  
 rfcf /data/... cchpsscms:/dcache/cms/...

cchpsscms



```

$ cat "/pnfs/in2p3.fr/data/cms/tapes/.(tag)(OSMTemplate)"
StoreName cchpsscms
StorageGroup 18
  
```

# The PoolManager

- Pool groups can be used to separate logical LCG activity
  - “store-only” HSM-enabled pools for bulk transfers of incoming data (from T0, T2 or other T1)
  - “restore-only” HSM-enabled pools for bulk transfers of outgoing data (to T2 or other T1)
  - “rw” disk pools for disk-only storage
  - ...

# SC3 setup

Client

“store-only” pools  
running GridFTP

RFIO server



- Used COS 1 (direct access to 9940b tapes without using HPSS disk)
- Disk cache (6.5 TB) provided by 4 “store-only” pools
- SRM and Postgres DB running on the head-node



# Open issues

- Installation is OK using home-made script to automate the post-installation procedure
- SRM problem with FTS occurred first day of tape throughput tests! No understanding of the problem, no clean fix (database destroyed)
- Missing documentation on admin door commands
- Some issues with HSM link (load-balancing, timeouts...)
- dCap Call-back feature against firewalls

# Future plans

- Some more tests
  - Install a dCache test platform against the HPSS test platform to test performance and stability more deeply
  - Intensive local I/O tests (non-sequential reads) using dCap and Xrootd
  - Open dCache for non-LHC experiments as a single-root file system for disk storage facility (astro, D0)
- dCache in production in october
  - Install as a LCG SE
  - Integrate in the local warning system

# Conclusions

- We are confident in using dCache in production but aware of the amount of work in order to reach a good level of administration
- We feel that we would benefit from having the sources...
- We are ready to participate to any effort (development, tools, documentation...)
- Suggestions
  - Have a dCache admins mailing list
  - Have a regular (once a year?) meeting with dCache developers and admins (and users?)