PIC
port d'informació
científica

dCache upgrade plans at PIC

Gerard.Bernabeu@pic.es

PIC
port d'informació
científica

Index

- Current dCache deployment
- Targeted deployment
- 1.9.10 experience
  - New config files
- PIC Tier1 dCache instance upgrade plan

PIC
port d'informació
científica

Current dCache deployment

Tape dCache (Tier1/2 dCache)

- 1.9.5-23 (~4.5PB in SATA disks)
- Enstore Tape back-end (~4PB LTO3-5 in 2 robots)
- supporting 10 experiments (ATLAS, CMS and LHCb from WLCG)
- several protocols
  - HTTP with home made browsing and user/password authentication
  - [gsi]dCap, GridFTP
  - Xrootd – actually not used by any experiment
    - a few tests done by CMS failed. Not able to open a file.

- 3 dCache servers: dccore, dcip, srm (2*2Core@2.3, 8GB RAM)
- 1 pnfs (1*4core@3.5, 48GB RAM, 80GB FussionIO, 6 HDD RAID6) + pnfs RO (postgresql9 Streaming Replication)
- 4 doors (1 CPU L3406+4GB+2x1GE aggregated)

PIC
port d'informació
científica

Current dCache deployment

Disk dCache (Tier3)

- 1.9.10-5 (60TB)
  - deployment delayed because of xrootd bugs. Solved in -5
  - users worried about bug RT#6010
- No Tape back-end
- 1 dCache server & Chimera (1*4core@3.5, 48GB RAM, 6 HDD RAID6)
- 2 doors (1 CPU low power+4GB+2x1GE aggregated)
- Protocols in use: gridFTP, WebDAV/HTTP, [gsi]dCap, Xrootd

Test dCache (Tape dCache)

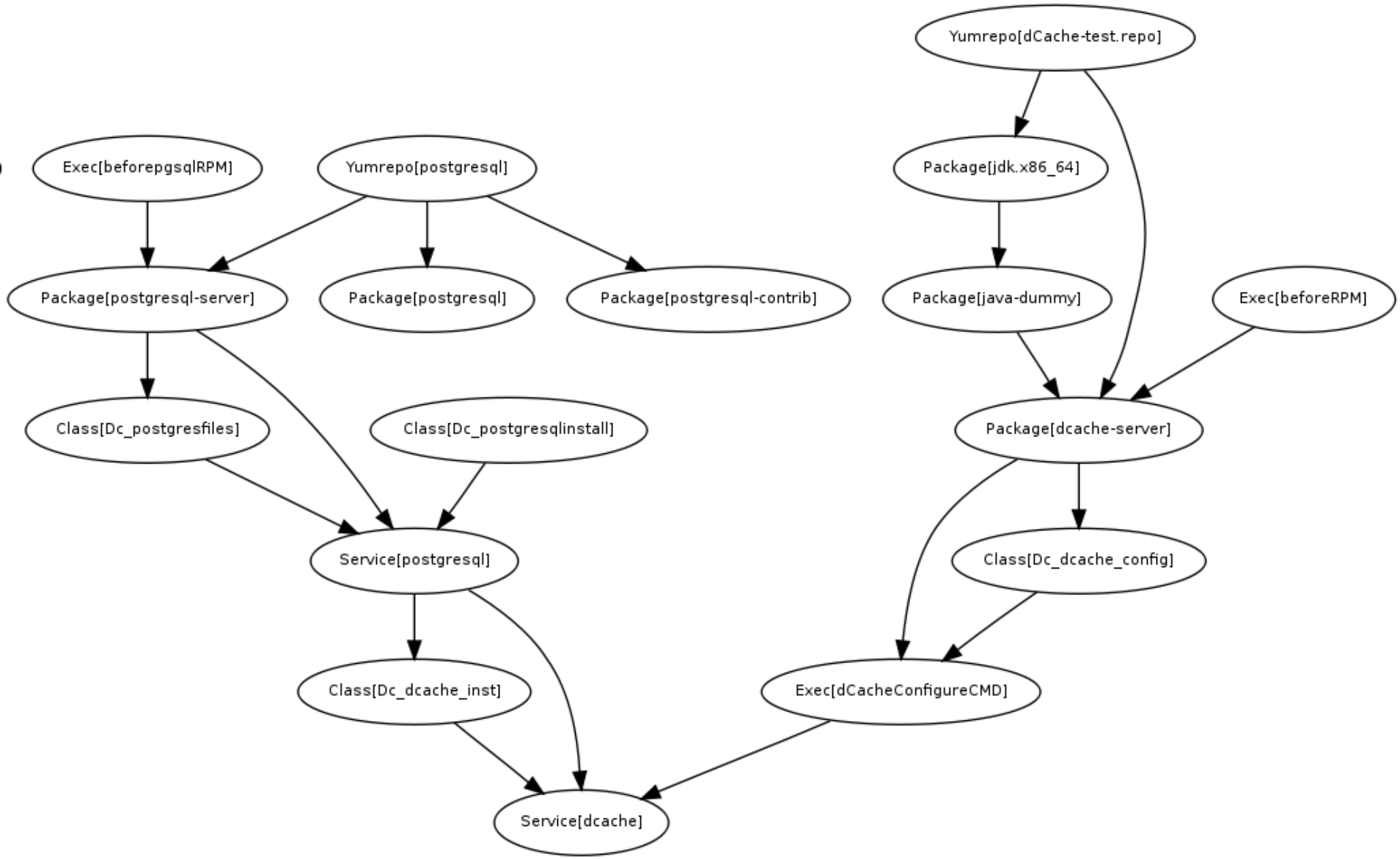
Virtual Machines, for functional tests before upgrading Tier1/2 dCache.

All servers are managed by Puppet, including PostgreSQL and dCache install&config**.**

Puppet

PIC
port d'informació
científica

Puppettized d*Cache*



PIC
port d'informació
científica

Targeted deployment

- Move Tier3 to disk-only instance
- Move Tier2 to disk-only instance
- Upgrade both dCache instances to 1.9.12 (Golden release)
  - Move the Tape dCache to Chimera
  - Start using NFSv4.1
  - Reduce the number of active dCache servers from 4 (pnfs, srm, poolmanager, info) to 3 (pnfs, srm, catch-all)

PIC
port d'informació
científica

1.9.10 experience

- NFSv4.1 is focused on reads, works but special kernel is required for SLC5.
- WebDav/HTTP is more efficient than dCap in LAN bulk data transfers (wget vs dccp).
  - Using ROOT client doesn't work. HTTP reimplemented in 1.9.11.
- Nice HTTP browsing capabilities
  - lack of simple user/pass authentication (we'll use a proxy to workaround).
  - Not customizable (ie: add file size), this will come with 1.9.12.
- Several bugs found in Xrootd, solved in 1.9.10-5
- We still need to get rid of some dependencies with the old info provider (gone after 1.9.5), like the user-view space monitoring

PIC
port d'informació
científica

1.9.10 experience

Fri Mar 11 16:35:03 CET 2011 http://dcmon.pic.es/SpaceMonitorOutput.html

atlas;;Free:540782;Total:2133993;Used:835906;Reserved:757305;Cached:248457
ATLASDATADISK;Free:86005;Total:650000;Used:563995;TokenID:11643
ATLASMCDISK;Free:295124;Total:300000;Used:4876;TokenID:11746
ATLASDATATAPE;Free:62994;Total:64000;Used:1006;TokenID:11747
ATLASMCTAPE;Free:19914;Total:20000;Used:86;TokenID:11748
ATLASdefaultToken;Free:0;Total:0;Used:0;TokenID:30730542
IFAEdefaultToken;Free:0;Total:0;Used:0;TokenID:30730551
ATLASGROUPDISK;Free:7216;Total:20000;Used:12784;TokenID:5269825
ATLASPRODDISK;Free:25000;Total:25000;Used:0;TokenID:64959145
T2ATLASLOCALGROUPDISK;Free:4566;Total:8000;Used:3434;TokenID:5278301
T2ATLASPRODDISK;Free:23208;Total:25000;Used:1792;TokenID:3517583
T2ATLASSCRATCHDISK;Free:8124;Total:28000;Used:19876;TokenID:17930260
ATLAST2spaceAccounting;Free:272752;ToAssign:128878;InTokens:389530;Total:518408;TokenID:-
ATLAST1spaceAccounting;Free:776878;ToAssign:421585;InTokens:1194000;Total:1615585;TokenID:
withoutToken;Free:540782;Total:-;Used:9687;TokenID:-
atlas;Tape-Recall;Free:1049630;Total:2133993;Used:0;Reserved:0;Cached:1084363

cms;;Free:792050;Total:938420;Used:42413;Reserved:103957;Cached:658379
CMSunmerged;Free:46123;Total:50000;Used:3877;TokenID:48064436
CMSbackfill_1;Free:7798;Total:25000;Used:17202;TokenID:48064482
CMSbackfill_test01;Free:25000;Total:25000;Used:0;TokenID:48064484
CMStemp;Free:25000;Total:25000;Used:0;TokenID:48676186 withoutToken;Free:792050;Total:-;Used:21334;TokenID:-
cms;Tape-Recall;Free:237628;Total:938420;Used:0;Reserved:0;Cached:700792

PIC
port d'informació
científica

New config files

With the new config files it is now very easy to customize services in domains (java procs).

- After spending a couple hours with the new config files I wouldn't go back.
- Very useful to go through /opt/d-cache/share/defaults (all possible properties are there!)
- When referring to variables/properties, use ${name}, it is not bash (don't miss the {})
  - In 1.9.10 the system doesn't complain if a property is not properly defined/recognized
- Would be good to have a recommended daemon-service layout based on cell relationships

PIC
port d'informació
científica

New config files

```
cat ./opt/d-cache/etc/layouts/dccore.conf.disk
#dccore
[dCacheDomain]
[dCacheDomain/poolmanager]
[dCacheDomain/dummy-prestager]
[dCacheDomain/broadcast]
[dCacheDomain/loginbroker]
[dCacheDomain/srm-loginbroker]
[dCacheDomain/topo]
[dCacheDomain/gplazma]
[dCacheDomain/gsi-pam]
[dCacheDomain/pinmanager]

[infoDomain]
[infoDomain/billing]
[infoDomain/httpd]
[infoDomain/webadmin]
[infoDomain/admin]
[infoDomain/info]
[infoDomain/statistics]


#dcns
[nfsDomain]
#Not possible to run both 4.1 and 3 in same server
[nfsDomain/nfsv3]
#[nfsDomain/nfsv41]

[namespaceDomain]
[namespaceDomain/pnfsmanager]
[namespaceDomain/cleaner]
[namespaceDomain/acl]
[namespaceDomain/dir]
```

```
#srm
[srm-${host.name}Domain]
[srm-${host.name}Domain/srm]
[srm-${host.name}Domain/spacemanager]
[srm-${host.name}Domain/transfermanagers]

#Pools
#[${host.name}Domain]
#[${host.name}Domain/pool]
#name=${host.name}_1
#path=/dcpool/vpool1
#waitForFiles=${path}/data
##queues and movers setup statically according to the pool
setup file

#Door servers
#[dcap-${host.name}Domain]
#[dcap-${host.name}Domain/dcap]

#[xrootd-${host.name}Domain]
#[xrootd-${host.name}Domain/xrootd]

#[gridftp-${host.name}Domain]
#[gridftp-${host.name}Domain/gridftp]

#[webdav-${host.name}Domain]
#[webdav-${host.name}Domain/webdav]

#[gsidcap-${host.name}Domain]
#[gsidcap-${host.name}Domain/gsidcap]

#[ftp-${host.name}Domain]
#[ftp-${host.name}Domain/ftp]
```

PIC
port d'informació
científica

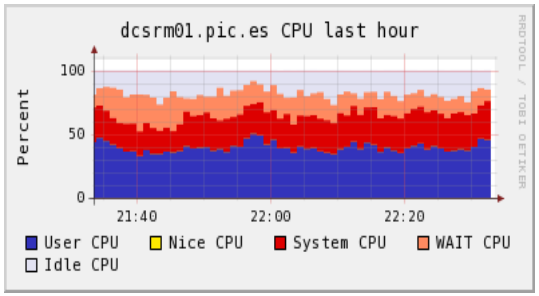
New config files

```
class dc_dCache_Config199 {

      case $node_type {
            pool: {
                  file { '/dcpool/vpool1/setup':
                      owner => 'root', ensure => 'file',
                      source => [ "puppet://ser01.pic.es/dc_dcache/199/dcpool/vpool1/setup.thumper",
                             "puppet://ser01.pic.es/dc_dcache/199/dcpool/vpool1/setup",],
                      replace => "false",     group => 'root', type => 'file', mode => '420',
                  }
            }
      }
      $serviceLocatorHost=$instance ? {
                        "disk"      => "dccore-disk.pic.es",
                        "test"      => "dccore-test.pic.es",
                        "prod"      => "dccore.pic.es",
                  }
      file { '/opt/d-cache/etc/dcache.conf':
          Owner => '0',     ensure => 'file',     group => '0',    type => 'file',        mode => '420',
          content => "
            #Custom
            dcache.layout=$node_type
            #All
            adminHistoryFile=/opt/d-cache/adminshell_history
            serviceLocatorHost=$serviceLocatorHost
            DefaultRetentionPolicy=REPLICA
            DefaultAccessLatency=ONLINE

            #No need to change this in the disk-only installation:
            #pnfsInfoExtractor=diskCacheV111.util.EnstoreInfoExtractor
            #dcache.namespace=pnfs

            #Pools
            metaDataRepository=org.dcache.pool.repository.meta.db.BerkeleyDBMetaDataRepository
            poolIoQueue=ftp,lan
```

PIC Tier1 dCache instance upgrade plan

Disk dCache Plan

1)As soon as it is available upgrade to 1.9.12



Tape dCache plan

2)Upgrade SRM hardware (8GB RAM and 2xdual core AMD 2218 is not enough anymore)

3)Get rid of old infoProvider dependencies (SRM reservations monitoring page)

4)Upgrade to 1.9.12 on the Tape dCache. On the upgrade also

    1)Join in a single server PoolManager and InfoServices servers

    2)Move all postgreSQL DBs to 9.0, migrating from PITR to streaming replication

5)Move the Tier2 to the *disk dCache*

6)Migrate to Chimera when tested and working with Enstore2 (Xmas 2011?).