dCache, an update for the

Patrick
for the dCache Team

support and funding by
Content

Project Topology
- The Team
- The Partners
- The Sub-projects
- Version Management

Work in progress
- SRM 2.2
- Chimera
- NFS 4.1

In a nutshell
- Big Picture
- Basic Feature Set
- New Features in 1.7.0

Deployment and distribution
- Automated testing procedure
- Deployment process
- Interaction with OSG/VDT
Project Topology: The Partners

[Diagrams of various organizations' logos and names, including DESY, VDT, Open Science Grid, NDGF, CERN, and others.]
Project Topology: Version Management

Now

dCache 1.8.0

JAVA 1.5

optional new code

same code

Stable 1.7.1

LCG goes java 5

April 1, 2007

JAVA 1.4

dCache 1.7.1

dCache 1.7.0

Stable 1.7.1
In a nutshell

Managed Storage

Basic Feature Set

New Features in 1.7.x
In the Nutshell

dCache

Managed Storage Controller and Protocol Engine (⇒ OSG SE)

**In the Nutshell**

**dCache**

**dCache sub-services**
- Resilient Controller
- Maintenance Module
- Admin Client Interface
- HSM Flush Subsystem

**Managed Disk Storage**

**Protocol Engines**
- **Name-space Services**
  - PNFS (chimera)
- **Information Protocols**
  - GLUE 1.2
  - SRM
  - EIS
  - Streaming Data
  - (gsi)FTP
  - http(g)
  - Posix I/O
  - xRoot
  - dCap

**Backend Tape Storage**
- OSM, Enstore, Tsm, Hpss, DMF

Patrick Fuhrmann

OSG Meeting, SDSC

March 6, 2007
Basic Feature Set

- Strict name space and data storage separation.
- Multiple internal and external copies of the same file
- Automated file replication on access hot spot detection
- HSM connectivity (enstore, osm, tsm, hpss, dmf)
- Automated HSM migration and restore.
- Handles data in Peta-byte range on 1000's of pools
- Supported protocols: (gsi)ftp, (gsi)dCap, xRoot, SRM, nfs2/3
- Supports resilient dataset management (worker-node support)
- Sophisticated command line interface and graphical interface
In the Nutshell  New Features in 1.7.0-1

- dCache partitioning for very large installations
- File hopping on
  - automated hot spot detection
  - configuration (read only, write only, stage only pools)
  - on arrival (configurable)
- gPlazma
- xRoot support (with Alice authorization)
- BUG FIX: gsiftp movers killed on idle timeout.
- Central FLUSH manager
- Maintenance module (draining pools)
- improved GUI
- Jpython interface for all kind of configuration
Deployment and distribution

Automated testing procedure

Deployment process

Suggested interactions with VDT
Automated testing process

CVS check-in

Full Compilation and RPM creation

Results on web page and e-mailed to developers

CVS Tag

Full Compilation

RPM creation

RPM from developer repository

(Regression) Test suite

In Progress

OK

Web Site

ATP repository

(sl3/4 ; 32/64 bit)

OK or Failed to developer
Deployment and feedback Process

Feedback from user community

- support @ dCache.org for bug reports
- user-forum @ dCache.org for 'users helping users'

Deployment/Announcement of new versions resp. sub-versions

* New subversions are announced at
  user-forum and announce @dcache.org
  (and RSS feed in the future)
* and are published on the dCache.ORG web page
* and are published in the 'stable' APT repository
* RPM will always have the corresponding 'change log' included
dCache - OSG / VDT interaction

Suggestion:

• dCache.org makes sure that VDT related installation and configuration scripts and procedure are integrated into the official dCache.ORG rpm. (This is actually already the case).
• dCache.ORG informs 'liaison officer' at OSG/VDT that there is a new official dCache (sub)version.
• VDT takes the original dCache.ORG rpm and performs the necessary field tests at ???
• Results on these test runs feeds back into support @ dCache.org
• OSG takes over first level support for dCache from VDT installations and forwards real issues to 'support @ dCache.org'
Third party contributions

NDGF

OSG/VDT
The NDGF story

- The goal is to have a single dCache instance distributed among all Nordic countries. (Norway, Sweden, Denmark, Finland and maybe Island)
- Head-node including SRM will be in Copenhagen.
- Pool nodes are located at all NDGF partner countries.
- This requires to have current FTP version-1 deficiencies overcome.
- NDGF and dCache.ORG and globus working on an implementation of Protocol Version II (not to be mixed up with globus gsiftp version II).
- NDGF changed globus c-client code to be protocol version II compliant and globus agreed to take this back into their repository. Now they are working on the dCache server protocol version II code.
- To our current knowledge, the FTS guys will integrate the gsiftp protocol version II into FTS if globus goes along this line.
**Easy Installation**

1) VDT Integration: I leave this to Neha

2) dCache in 10 minutes (funded by D-Grid)

**Extensions**

1) xRoot (by D-Grid): see previous chapter

2) gPlazma (by OSG)
   - Steven will report on this
   - currently working on secondary group support
Ongoing Development

SRM 2.2

Chimera

NFS 4.1
**SRM 2.2**  Main features

### Storage Classes

Administrator determines 'retention policy' and 'access latency'

- **Retention policy** REPLICA, CUSTODIAL
- **Access Policy** ONLINE, NEARLINE

Tape1-Disk0 : NEARLINE + CUSTODIAL  
Tape1-Disk1 : ONLINE + CUSTODIAL  
Tape0-Disk1 : ONLINE + REPLICA

Storage Class Transitions foreseen (not high priority)

### Space Tokens

To guarantee space for incoming transfers.  
Later maybe for 'restores from tape' as well.
**SRM 2.2 Milestones**

Jamie Shiers (WLCG)

Services are required for testing in Q2 (two) in preparation for the Dress Rehearsals in Q3 (and the LHC pilot run in Q4)...

- **1st April 2007** - target date for the needed services to be in place at the sites
- **1st June 2007** Ruth wants to have SRM 2.2 stable
- **1st July 2007** - start date of Dress Rehearsals (also the date when the WLCG service is commissioned)

**dCache**

- **1st April 2007** – beta version of dCache 1.8 (with SRM 2.2) available at dCache.ORG. Some sites already committed to have hardware for field testing available at that point.
- **3/4 July 2007** – SRM 2.2 'First Results' workshop at DESY. **It would be great if OSG would contribute to the workshop.**
Basic WLCG MoU functionality
- Missing 0 out of 25

WLCG MoU functionality due end of 2007
- Missing 2 out of 4

Non MoU functionality
- Missing 6 out of 12

Extended use cases
- Missing 6 out of 40

Up to date information from Flavia's 'test page'
http://grid-deployment.web.cern.ch/grid-deployment/flavia/
In case we are running out of time, read this, otherwise skip this page.

Chimera : replacement for pnfs
ready for field test : mid of march

NFS 4.1 : read data access through NFS
collaboration with citi.umich.edu
really cool stuff
Chimera

Expected Improvements compared to PNFS

• Performance scales with back-end database implementation
  ➣ Small to medium sites with mysql/postgres
  ➣ Really huge sites with oracle cluster (planned for DESY)
• Enables protection against misuse
  ➣ Different 'chimera users' (e.g. nfs, dCache, enstore) may get difference doors with different priorities if back-end db allows.
• Simplifies maintenance resp. monitoring tasks
  ➣ By using SQL database
  ➣ Easy to add customized web interfaces.
• Allows ACL plug-ins
  ➣ ACL sub-project started beginning of 2007
Current status

- Functional and performance tests in progress
- Ready for testing by external sites: mid of March
- Setting up pnfs -> chimera (de-)migration scenarios
- Production time-line: depends on results of tests; otherwise as fast as human resources allow.
NFS 4.1

Highlights

- Standardized interface to dCache name-space and data
- 4.1 extension makes use of highly distributed data
- Security (e.g. certificates) is part of spec.
- Clients are provided by OS maintainer(s)

citi.umich.edu is pushing to have the dCache server ready soon
Further reading

www.dCache.ORG