dCache.ORG

dCache.ORG

dCache NDGF workshop introduction

Patrick
for the dCache Team

support and funding by

Content

dCache.ORG

Project Topology

- The Team
- The Partners
- The Activities
- Version Management
- Testing and Deployment
- NDGF Requirements
- Version Management

In a nutshell

- Big Picture
- Basic Feature Set
- New Features in 1.7.0

Work in progress

- SRM 2.2
  - Main features
  - Milestones
  - Status
  - SRM version interoperability issues
  - SRM evaluation deployment plan
- Chimera
- NFS 4.1

Deployment and distribution

- Automated testing procedure
- Deployment process

dCache.ORG

Project Topology

- The Team
- The Partners
- The Activities
- Testing and deployment
- Special NDGF requirements
- Version Management

Project Topology : The Team

dCache.ORG

dCache.ORG

Head of dCache.ORG

Patrick Fuhrmann

Core Team (Desy and Fermi)

Forrest Christian
Ted Hesselroth
Alex Kulyavtsev
Birgit Lewendel
Dmitri Litvintsev
Dirk Pleiter
David Melkumyan
Martin Radicke
Owen Synge
Neha Sharma
Vladimir Podstavkov

Head of Development FNAL :

Timur Perelmutov

Head of Development DESY :

Tigran Mkrtchyan

External Development

Gerd Behrmann, NDGF
Abhishek Singh Rana, SDSC
Brookhaven in preparation

Support and Help

Greig Cowan, gridPP (monAmi)
Stijn De Weirdt (Quattor)
Maarten Lithmaath, CERN
Flavia Donno, CERN

Project Topology : The Partners

dCache.ORG

DESY

NDGF
Nordic DataGrid Facility

Open Science Grid
VDT
Virtual Data Toolkit

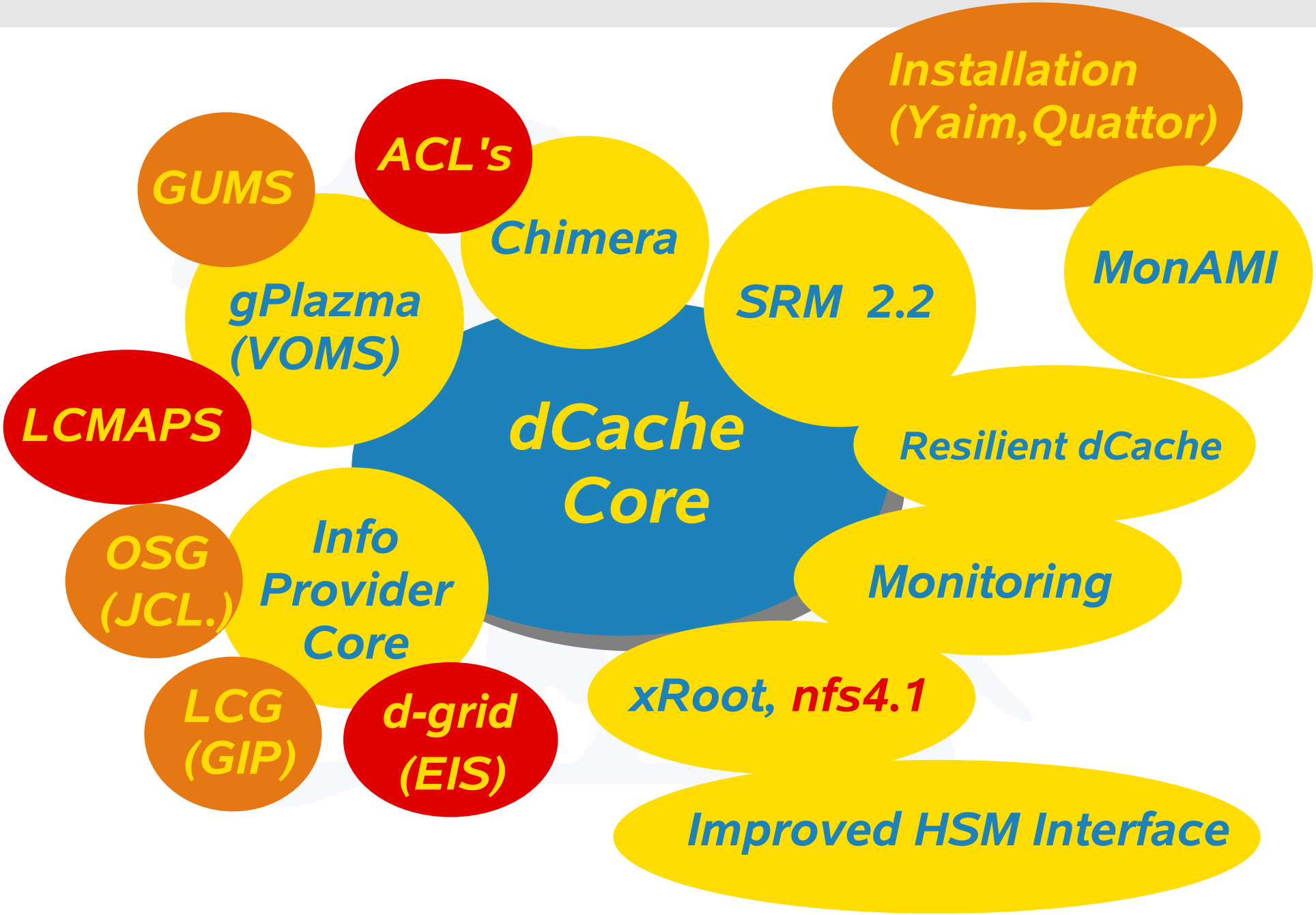
CERN

Bundesministerium
für Bildung
und Forschung

HEP CG
D-GRID

Project Topology : The Subprojects

dCache.ORG



Project Topology : The Subprojects

SRM in general

Resource Requirement footprint needs to be significantly reduced !
Timur et al. are working on this.

SRM 2.2

Timur (Interface and DB) and Tigran (dCache core) plan to have the first WLCG compliant SRM 2.2 dCache ready mid of march.

xRootd integration

- Protocol plus 'non standard Alice' authorization done.
- Gsi Authentication planned.

gPlazma

- All protocols except xRoot integrated.
- GUMS integrated.
- Scripting workaround for non-GUMS (LCG) sites.

Chimera (pnfs replacement)

- First phase of development done (Tigran)
- Performance evaluation and code review in progress (Vladimir)
- Evaluating pnfs <-> chimera migration scenarios.
- ACL sub project started end of December. (David, Dirk)

Project Topology : The Subprojects

dCache.ORG

gsiFtp improvements (Nordic Data Grid Facility)

- NDGF plans for single dCache instance spanning multiple countries.
- Need to improve current dCache gsiFtp implementation to avoid long data path.

Resilient dCache module (Alex)

- In great demand.
- Second, improved version in preparation.

Improved Monitoring

- SRM watch (Dimitry)
- dCache monitoring plots (Vladimir)
- much more is needed. (N.N)

nfs 4.1 (Tigran)

- nfs access to name-space and data.
- nfs4.1 supports distributed data locations (as dCache does)
- nfs clients will come for free.

HSM interface improvement (Radicke, Tigran)

- very important for Tier I's.
- First version of central flush manager ready.

Project Topology : (testing and deployment)

*** Fully automated *code to product* chain [Owen,Vladimir].**

- checking out CVS archive
- code compilation
- RPM production
- running test suite
- publishing on web page and APT repository

*** Slogan : dCache in 10 minutes (fast installation and configuration) [Owen,Ted]**

*** Adjusting dCache packaging to VDT needs in progress [Ted].**

*** Goal : only one set of RPMs for all distributions (**dCache.org, CERN apt, VTD**)**

*** CERN and dCache**

- production dCache in CERN repository
- dCache certification done by CERN staff against dCache instance at DESY

dCache.ORG

Project Topology : special NDGF requirements

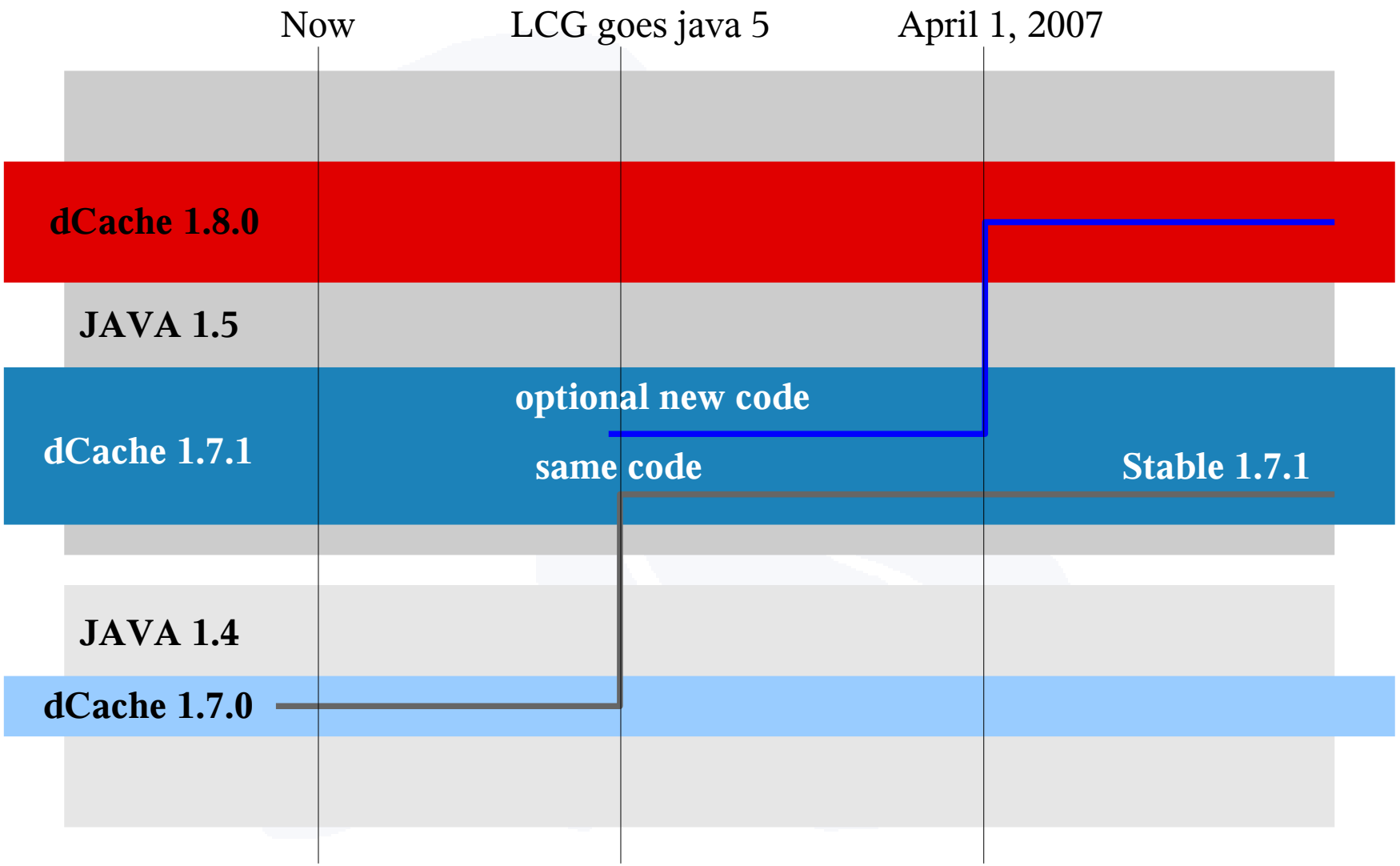
dCache.ORG

NDGF (multi site) requirements

- Secure Location Broker
- Secure Cell Communication
- FTP protocol version II
- Fine grained ACLs for cell commands
- Satellite sites independent of central system (puh...)

Project Topology : Version Management

dCache.ORG



dCache.ORG

dCache.ORG

In a nutshell

Managed Storage

Basic Feature Set

New Features in 1.7.x

In the Nutshell

Managed Storage Controller and Protocol Engine

dCache.ORG

dCache.ORG

dCache

Backend Tape Storage

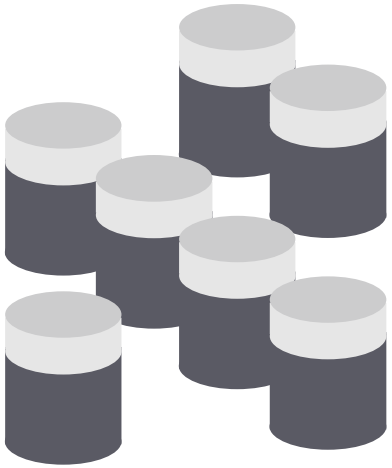
OSM, Enstore,
Tsm, Hpss, DMF

HSM Interface

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

Storage Control

SRM
EIS

Streaming Data

(gsi)FTP
http(g)

Posix I/O

xRoot
dCap

In the Nutshell Basic Feature Set

dCache.ORG

- 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.ORG

- 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

dCache.ORG

dCache.ORG

Deployment and distribution

Automated testing procedure

Deployment process

Automated testing process

dCache.ORG

CVS check-in

Full Compilation and RPM creation

Results on web page and e-mailed to developers

CVS Tag

Full Compilation
RPM creation

In Progress

RPM from developer repository

(Regression) Test suite

OK

Web Site
ATP repository
(sl3/4 ; 32/64 bit)

OK or Failed to developer

Deployment and feedback Process

dCache.ORG

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

Ongoing Development

SRM 2.2

- Main features
- Milestones
- Status
- SRM version interoperability issues
- SRM evaluation deployment plan

Chimera

NFS 4.1

dCache.ORG

SRM 2.2 Main features

dCache.ORG

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 (OSG) 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.ORG

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.

SRM 2.2 Status

dCache.ORG

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 Flavias 'test page'

http://grid-deployment.web.cern.ch/grid-deployment/flavia/

SRM version interoperability issues

* File name limitation <= 199 characters (full path has no limit)
  - |- Caused by the underlying file system engine (pnfs)
  - - Will be fixed with upcoming new engine (chimera)
  - - Production date for chimera is not decided yet (~ end of the year)

* PUT overwrite
  - - very likely fixed within the next week
  - - will be certainly fixed in dCache 1.8.0

* SURL doesn't exist on PUT before data transfer starts
  - - Our opinion is that this part of the spec. may introduce usability problems.
  - - If client doesn't clean up in case of a problem, subsequent PUT on the same filename will fail until a timeout expires. Except for overwrite=ALLWAYS

dCache.ORG

SRM version interoperability (minor issues)

dCache.ORG

* prepare to put followed by a 'Done' without data transfer should fail in dCache, but doesn't. *Remember* : dCache doesn't create the SURL on PUT before data transfer has started.

  - Agreed in general, but still discussion on details.

* Trying to extend lifetime of expired TURL succeeds but shouldn't.

  - Agreed, and will be fixed soon.

SRM version interoperability (details)

dCache.ORG

- The initial dCache version with SRM 2.2 included, is **dCache 1.8.0.**
- **dCache 1.8.0** and higher will support **SRM 1.1 and SRM 2.2** at the same time on the same TCP Port.
- Both SRM protocol versions will run in the same dCache instance, using just one file system instance. (pnfs)
- Both SRM versions will have access to the **same file name space.**
- Files written with 1.1 can be accessed via 2.2 and vice versa.

SRM evaluation deployment plan (Agreement)

- Sites agreed to deploy dCache 1.8 (SRM2.2) in April :
  - FERMILab, DESY
  - BNL
  - gridKa
  - IN2P3
- For those sites we will closely watch the installation and the behavior.
- Systems will have 1-2 head nodes and >= 10 TBytes of disk storage.
- Systems will be connected to a Tape Back-end to support all possible storage classes.

dCache.ORG

SRM evaluation deployment plan (restrictions)

dCache.ORG

- Full upgrade to 1.8.0 is a prerequisite for the SRM 2.2 activation.
- There is no way to have dCache versions prior to 1.8 running with SRM 2.2
- The following restrictions apply concerning the agreed test systems :
  - It will be a special dCache evaluation instance, and **not part of the production system**.
  - The service is not part of the production monitoring and may be **shut down at any time**, without further notice.
  - All **data** should be regarded as **'not persistent'** and should be copied to the production system in order to become permanent.

SRM evaluation deployment plan (timing)

dCache.ORG

April

1. Week : FERMI – DESY transfers
2. Week : Installation at BNL
3. Week : Installation at gridKa
4. Week : Installation at IN2P3

NDGF should join mid of April

May

RPM and Installation instructions at dCache.ORG

SRM evaluation deployment plan (timing)

dCache.ORG

Further steps depend on the success of the procedures described previously.

Just fair to say :

> Although it's certainly our goal to be in production shape in July, we can't yet give advice on whether or not to use dCache SRM 2.2 during the Dress Rehearsal.

dCache.ORG

dCache.ORG

Chimera

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

dCache.ORG

Chimera (cont.)

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.

see Tigrans talk

dCache.ORG

NFS 4.1

dCache.ORG

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

see Tigrans talk

dCache.ORG

dCache.ORG

Further reading

www.dCache.ORG