



dCache.ORG

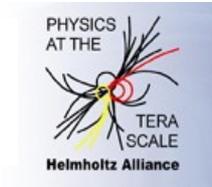
dCache.ORG

dCache, selected topics (LCG MB)

Patrick Fuhrmann



additional funding, support or contributions by



d-grid
DGI II





Content



Chimera and the Pnfs to Chimera Migration

Background

Chimera advantages and deployment

Risk assessment (Chimera , Pnfs)

dCache release policy and the Golden Release

Time bases releases

Golden Releases

dCache.ORG

dCache.ORG



Pnfs to Chimera migration



Background

- ★ The dCache file system engine is one of the most important dCache components. All other services rely on that functionality
- ★ The performance of the file system engine determines the performance of the entire system.
- ★ Pnfs is the initial implementation of the file system engine and has been introduced a decade ago.
- ★ The technology Pnfs is build on doesn't allow the system to be improved any more.
- ★ Main flaws :
 - ★ All Pnfs accesses have to go through the NFS 2/3 interface (even dCache)
 - ★ A single lock protects an entire Pnfs database, which means :
 - ★ The Pnfs performance doesn't benefit from multiple cores.
 - ★ Data is stored in BLOB which doesn't allow SQL queries.
- ★ Pnfs is no solution for the Petabyte storage domain.



Pnfs to Chimera migration



Solution : Chimera

- ★ dCache.org designed and implemented a successor : Chimera
- ★ Chimera is based on standard database technology
- ★ Advantages :
 - ★ There is no Chimera server, which allows Chimera to scale with the number of CPU/cores of the name space host.
 - ★ Standard SQL allows easy monitoring and maintenance.
 - ★ Further dCache features and speed improvements assume that the name space is based on Chimera.
 - ★ e.g. : ACL inheritance.



Pnfs to Chimera migration



Chimera deployment

- ★ Chimera is in production-use at some sites since Nov 2008
- ★ The NDGF Tier I migrated to chimera end of March '09.
- ★ In the meantime most German Tier II sites migrated to Chimera as well as sites in the US, Scandinavia and Italy. Other Tier II's will follow soon.
- ★ gridKa/KIT will build the new CMS instance on Chimera

dCache.ORG

dCache.ORG



Pnfs to Chimera migration



The migration process

- ★ Migration is composed of 3 steps
 - ★ Dumping Pnfs into the SQL format.
 - ★ Injecting the SQL into the new database.
 - ★ Verifying Pnfs against Chimera (md5 checksums)
- ★ For NDGF (8 million files) this took (11 + 3.5 + 11 hours)
- ★ Different Pnfs databases can be converted concurrently

dCache.ORG

dCache.ORG



Pnfs to Chimera migration



The migration process

- ★ The migration process can be tested without interference with the production system.
- ★ The Pnfs dump step does a very strict Pnfs consistency check.
- ★ If Pnfs inconsistencies are found, they can be fixed in advance.
- ★ This allows to run an uninterrupted 'real' conversion and it allows a good estimate on the time it takes.
- ★ There is no way back from Chimera to Pnfs unless you accept to immolate all newly create files.
- ★ dCache will treat chimera issues (if there are any) with highest priority, assuming fixing the potential problem will take less time than converting back to Pnfs.



Pnfs to Chimera migration



dCache.ORG

Risk assessment

Risk assessment has to be done by the site manager.

dCache.org tries to provide as much information as possible.

dCache.ORG



Pnfs to Chimera migration



dCache.ORG

dCache.ORG

Risk assessment

The conversion process

- ★ There is no risk associated with the migration process itself, as one can fall back to Pnfs as long as no file has been written into Chimera.
- ★ The migration process can be tested many times in advance to find the optimal setup, without interfering with the production system.



Pnfs to Chimera migration



Risk assessment

Risk running Chimera

- ★ Chimera is in production for about 8 months.
- ★ Pnfs is in production for about a decade.
- ★ However Chimera has been used by one Tier I and many Tier II's during the STEP 09 exercise and no problems were reported.
- ★ In case of Pnfs inconsistencies not many people would be able to fix the system, while Chimera is based on standard SQL tables and somebody with reasonable database experience would be able to investigate. (The structure of the tables are self-describing and no server is needed to extract the file system information)
- ★ There is no way back from Chimera, w/o writing own back-conversion scripts, which can be done as the source code and the design are available.



Pnfs to Chimera migration



Risk assessment

Risk running Pnfs

- ★ If performance issues are found during data talking we can't speed up Pnfs. One would have schedule an upgrade to Chimera.
- ★ Running faster hardware is only moderately beneficial as Pnfs can not scale with the number of cores/CPU's.
- ★ The STEP 09 exercise can be seen as a hint to judge Pnfs behaviour if one believes that this has been a reasonable test.
- ★ If in the past the PnfsManager has shown loaded queues over a long period of time (your local sysadmin knows what that means) a migration to Chimera would be recommended.



Pnfs to Chimera migration



Risk assessment

Inconveniences running Pnfs

- ★ Pnfs doesn't support ACL inheritance.
 - ★ which is required by the SRM auto directory creation.
- ★ Future dCache improvements will assume Chimera to be the file system back end.
- ★ Reasonable monitoring, maintenance and accounting can only be done with Chimera. e.g.
 - ★ Amount of data by user or group
 - ★ Amount of data on tape or pending to go to tape
- ★ The NFS 4.1 interface is only available with Chimera.



dCache release policy



Release policy

- ★ dCache.org is moving towards time bases releases.
- ★ Time bases release a common practice in large software projects.
- ★ Time of release is know but not exactly the features
- ★ Advantages
 - ★ Easier to synchronise with distributors
 - ★ No waiting for a release because the implementation of a feature is delayed.
 - ★ Sites can properly schedule updates.

dCache.ORG

dCache.ORG



dCache release policy



Golden release

- ★ Golden releases are supported much longer than regular ones.
- ★ The last Golden Release has been the CCRC branch which was a great success.
- ★ The next Golden Release will be support throughout the entire first LCG run period.
- ★ The Golden Release will be dCache 1.9.5
- ★ No new features will be added to that release, only bugs will be fixed.
- ★ There will be of course new regular releases being published with shorter lifetime and new features.
- ★ Sites are free to choose the Golden or the Feature Releases.
- ★ Questions from Andrea S. :
 - ★ 1.9.2 would no longer be supported with releasing 1.9.5
 - ★ 1.9.3 and 1.9.4 would become unsupported with releasing 1.9.6 etc.



Further reading

www.dCache.ORG

dCache.ORG

dCache.ORG