Fermilab DCache Project Status Report

Robert D. Kennedy

09 May 2006

Modified: 10 May 2006

CD-doc-1443

Basic Reference Material

- FNAL Project page: http://computing.fnal.gov/ccf/projects/dCache
  - material is dated, will be moved/linked to FNAL Dcache plone site
- FNAL Plone site: http://dcache.fnal.gov
  - new, hosted by central FNAL plone server, with virtual address hosted elshwhere
- FNAL GDH Group page: http://computing.fnal.gov/ccf/gdh.html
  - overlaps with project since project/line personnel not heavily matrixed
- Global Dcache Project page: http://www.dcache.org
  - details are more applicable to LCG, OSG-oriented sites
- Master schedule/drivers: See CD-doc-990, CD-doc-1286
  - continuously evolving with schedules and requirements of customers

Project Manager's Summary – April/May 2006

The focus of the FNAL DCache project for the last six months has been to maintain and expand Fermilab DCache systems, pursue a broader more effective development base for DCache (coordination with DESY, engaging BNL), meet the storage requirements of US-CMS-T1, and to prepare DCache to become a more generally deployed Storage Element on the OSG. Some achievements in this area include:

- FNDCA – maintained stable operations
  - Expand FNDCA capacity (with IA), re-organize pool structure, add Fermigrid pools
  - Host raw data read/write pools separately from general write pools
  - Added a test stand (FNDCAT) for developer testing and admin procedure practice
- CDFDCA – maintained stable operations
  - Successfully transferred liaisonship within group
- Broadening developer base – very modest, slow, but visible progress in this area
- US-CMS-T1 – help maintain operations, prepare and support for service challenges
  - Feature development (FTP progress markers), scaling investigations (billing cell)
  - Resilient dCache work on-going to achieve near-zero administration
  - SC4 successful, in progress. Bandwidth contention between WAN and LAN traffic.
- US-CMS-T2 – resilient dCache deployed and operational at seven sites, beyond original plan.
- Other Development Work
  - Dcache in a production branch-tagged release, continuous integration, bug fixes
  - Evaluation and initial testing of the PNFS replacement from DESY, Chimera
  - FTP feature enhancements, with performance improvements and test harness on the way
  - New monitoring plots system created, to be packaged for dcache.org deployment

For an activity break-down of status, please see Table 1 at the end of this report.

Resources Used

Resources used in the FNAL Dcache project include members of the CCF-GSS-GDM group: Rob Kennedy, Alex Kulyavtsev, and Vladimir Podstavkov. Also materially contributing to development, support, and/or documentation were: members of the IA group, Timur Perelmutov, Dmitry Litvintsev, and more recently Ted Hesselroth. Since the master resource-loaded schedule was constrained by a predetermined number of people, by definition we have not used more resources than expected. We have reduced scope and pushed out delivery schedules where appropriate.

Schedule/Milestones

The master schedule for FNAL DCache project states the FNDCA configuration expansion would be completed in October 2005. Due to security-required operating system upgrades (mostly sapping IA's time), hardware problems, and developer time commitments, this milestone has slipped and is still not yet completed. Since security comes first, and staff levels are fixed, this was unavoidable. The highest priority of this expansion have been achieved nonetheless. The admin node has been split into several admin nodes allowing for simpler expansion of the system. Addition of Fermigrid volatile space is in progress. Plans are ready to complete operating system upgrades, split up pools for finer-granularity to allow pool groups split across hardware, and then restore KTeV and Minos to full space allotment per hardware contribution.

After several days of discussions and demonstrations on a recent visit to DESY, I judged the Chimera product from DESY to be an acceptable replacement to PNFS for improved scalability of service, as opposed to developing an in-house alternative from scratch. Moderate scale testing of Chimera will begin here shortly, with deployment (if all goes well) to all FNAL systems before the end of the year. Plans are underway to better align and coordinate development efforts at DESY and FNAL to make this and other planned sub-system overhauls more effective with the same limited resources.

Significant Changes since Previous Report

Corrected post-meeting: The CMS SC4 phase in April required more effort than I had anticipated in recent project planning. This did not cause significant disruption to our project plans since this was offset by the additional effort from Ted Hesselroth. A revised quarterly project schedule is in discussion which today happens to be fitted to the CSA06 dates. Priorities within that schedule may change, but the overall schedule itself needs to be maintained in order to coordinate longer-term efforts.

We have participated in the submission of several funding proposals which either directly involve the FNAL Dcache project or may indirectly contribute to components of Dcache. These include: a SciDac2 proposal to work on Dcache components most likely to be impacted in the next 4x service scale increase for LHC, a SciDac2 proposal to work on component selection within a distributed system, a proposal for FNAL to work on FTP issues in Globus, and a follow-on proposal to continue work on the SRM. There has been no response yet on the status or disposition of these proposals.

There have been significant discussions between FNAL and DESY concerning SciDac2 requirements, Dcache software licensing, and a potential MOU to provide a firmer formal foundation for

the multi-institutional Dcache project. There have been expressions of willingness to reach an agreement consistent with funding requirements, but there has not yet been a formal expression of terms for such an agreement.

Future Plans

We will have to address some issues long deferred in order for the Dcache service to make it a more easily managed service as well as a more smoothly evolving service. In addition to continuing to transfer some administrative functions to IA to free up developer time, we will consider development to permit service quality to be better monitored by non-developers (Billed DCache "errors" are usually just delays that do not lead to client service failures), as well as developing realistic test procedures to reduce the effort to validate production releases and deploy bug fixes. Test harnesses are needed to test functionality more systematically and thoroughly. A means to test the many potential configuration variations in DCache, and to simulate realistic loads and use-patterns to test functionality at scale, should be considered. We need to work to avoid the long-standing pattern of ultimately testing functionality "in production".

Presentations/Reports/Documentation

Please see the FNAL Dcache web page and the soon-to-be filled our plone site for other presentations, reports, and documentation. In the past 6 months, talks and papers were submitted to CHEP06, and several presentations were given to share our experience with BNL and other interested parties.

Project Activities	Task State	Commentary
Support	active	Investigations, consulting, some tasks to IA
Doc & Packaging	to be active	Review user, admin doc on FNAL web
Feature Development	active	bugs, scaling in "other" Dcache components
GridFTP	active	feature additions, performance, error handling
Resilient Dcache	active	driven by US-CMS-T1,T2. Reduce administration
Tapeless Data Path	*closed*	*done*
Pin Manager	*closed*	*done*
VO Auth Mod Integr.	active	work left for both DESY, FNAL – close to complete
Dcap support	low activity	packaging for kits, some bug investigations
Collaboration	active	DESY visit, tele-conf, how to coordinate better
Doc & Packaging 2	to be active	insure dcache.org rpms usable for FNAL systems
Features 2	to be active	build procedures, test harness, FSL workgroup
Pnfs	active	Chimera tests, Pnfs/postgresql support
Other	inactive	scheduling, cost model, XFS, etc.
WAN	inactive	work done under GridFTP task or by SRM project
OSG Participation	low activity	Re-engage in summer outside of US-CMS-T2
LCG Participation	low activity	mostly DESY and SRM involved. Some CMS issues
Management	active	SciDac2 proposals, DESY discussions... to MOU?

Table 1: Overview of activities under CD/CCF/DMS/Upper Storage/Dcache