T3 Steering Board Meeting 04

Zurich, ETH Hönggerberg HPK D 32 – Monday, 2/11/13 14:00 - 16:00

Present: *Derek Feichtinger* (chair, T3 admin), *Christoph Graph* (CHIPP Computing Board chair, representing ETHZ), *Clemens Lange* (representing UniZ, via Vidyo), *Urs Langenegger* (representing PSI), *Fabio Martinelli* (T3 admin), *Frank Meier* (representing UniZ), *Daniel Meister* (representing ETHZ, T3 admin), *Silvia Taroni* (representing UniZ)

Absent: —

Changes in Membership

- *Silvia* joined the SB representing UniZ

- *Clemens* will join the SB representing UniZ

- *Frank* will be leaving PSI; need to locate a successor

- *Daniel* replaces *Leonardo Sala* representing ETHZ

Status Update by Fabio

CPU

- CPU usage was ok; we almost always have free slots for fast turn-around jobs during working hours

- Cluster is not undersized; policies (quotas for long jobs vs. short jobs) can be adapted easily if decided so by the SB

- **Conclusion:** We don't need more CPUs for 2013

`/shome` storage

- We have now up to 12TB available for home directory storage; with that we can easily fulfill the quota of 150GB per user (assuming around 50 users)

- There is an offer from Oracle extending the warranties on these machines for a reasonable price that would give us the flexibility to wait for another year looking for a new solution that is comparable in terms of features and capabilities.

- **Conclusion:** We would like to take this offer for at least a subset of the SUN HW.

dCache storage

- Planned additions to the dCache based mass storage for 2013 consist of two new boxes giving a net increase of 264TB

- We can operate this cluster totalling 750TB for 2013 but need to start replacing old HW by 2014

- **Conclusion:** Storage is (almost) used up to the limits these days; this is the most critical extension for 2013

User Interface Machines

- Fabio has installed 9 new user interfaces which now have a total scratch space of 2.3TB

- **Conclusion:** If there are funds available the users could profit a lot from high-end user interfac machines.

Discussion Topics

Leaving Users

- We need a standard procedure to handle users leaving the institutes (and maybe also the CMS collaboration)

- **Conclusion:** The institute contact is responsible to notify the admins of leaving users; the admins will regularly send a reminder with some metric describing how active each user was. All notices of leaving users will trigger the same standard procedure that is to be proposed by the admins and agreed to by the SB.

SE Quotas / Group Folders

- The SE is almost full at the moment and it is very difficult to enforce proper cleanup by the users.

- The space usage is very unequal; mostly due to the distribution of work within the groups (i.e. one user producing files that are then used by many others).

- There are some unofficial group folders created by the users themselves; they should be promoted if still active.

- Users are naturally grouped by working on the same resp. on similar analyses. These groups are usually subgroups of the actual physics groups and a user can very well belong to more then one group.

- Quotas are not (and will not be) enforced technically (this is not possible with the underlying storage technology); offending users are informed via automated e-mails.

- **Conclusion:** We agree that creating group folders is the best approach to entangle the situation. The storage used by a group will be deducted equipartitionally from the personal quota of all the users in the group (due to the expected number of active users and available storage the personal storage is set to 10TB per user). The admins will react as flexible as possible to special cases and demands.

- **Conclusion:** We need a *real* cleaning campaign involving all active users; for this the admins will prepare a detailed overview of used storage (e.g. also including time of creation, time of last access, etc.). Cases of users not reacting / not participating will be escalated to the SB.

Tools for organizing Files on the SE

- There are a number of options available with the new dCache version. (1) Provide read-write access over NFS, (2) provide read-only access over NFS, (3) use available tools that allow for better user interaction (e.g. `uberftp`).

- **Conclusion:** For the moment we would prefer the solution with a read-only mount and third-party tools for modifying the files.

NX access to the User Interface Machines

- The OpenSource alternative we use to provide fast X-window access the the UIs has some incompatibilities with the newest commercial version for Mac OS X.

- Testing a new work-around is ongoing.

- **Conclusion:** First determine the actual need for this access (maybe inform users about alternatives); if many Mac OS X user actually use it, find an appropriate solution.

Upgrade Plans for 2013

- The funding requested from SNF for the new hardware is not at all guaranteed (info should be available end of March / beginning of April).

- New UI hardware (cost for a reasonable set of machines around 25 kCHF) could be added depending on available funds. Old HW machines prevent us from profiting from falling prices of e.g. harddisks for scratch space.

- Need to determine how much funding would be available from the institutes in a worst-case scenario (i.e. nothing from SNF).

- **Conclusion:** The top priority of 2013 is the upgrade of the SE HW (approximate costs are 100 kCHF).

GNU gold Issue

- Some of the newer CMSSW releases exhibit a strange bug when used with our Solaris NFS server (compiled executables are filled with NUL-bytes due to a bug in the new linker).

- The bug has been identified and fixed but due to time (and in the near future technical) constraints the admins cannot patch every CMSSW release.

- The new Oracle contract would allow us to upgrade the NFS on the server side and hopefully to get rid of this issue.

- **Conclusion:** The users can work around this issue but it is a big inconvenience for the them (especially because it is sometimes hard to detect that one is affected by this bug) and we should find a better solution than requireing the users to compile everything on scratch and copy it back.

Size of Home Directories

- The current home directory quota (150GB per user) is a bit large given the fact that one should not store huge analysis/data files there but on the SE.

- **Conclusion:** The SB could not find a convincing use case but because the system can handle these quotas at the moment without any problems and there is no impact on other users, we decided not to change anything.

Next Meeting

The next meeting of the SB will take place at the beginning of 2014 or at any time if there are urgent topics to discuss.

Action Items

- **on the T3 admins:** define the exact procedure for leaving users (e.g. who gets warning e-mails, what is the grace period to clean up and/or move the data, etc.)

- **on the institute contacts:** check list of current users to find accounts that are no longer needed

- **on the T3 admins:** find the best way to support group folders and develop tools that take care of the accounting on the SE

- **on the T3 admins:** create a list of unofficially created groups and send it around to the SB to decide which of those should be promoted to official groups

- **on the T3 admins:** prepare a careful analysis of the files currently stored on the SE to start a next cleaning campaign

- **on everybody:** encourage users to clean up their data

- **on the T3 admins:** run last performance checks over the NFS read-only mount of the SE before making it available to the users

- **on the T3 admins:** determine the need for NX access and find an appropriate solution

- **on the institute contacts:** get a clear idea of what kind of funding would be available from your institute if we don't get the SNF funds for the new HW

- **on the T3 admins:** check whether the NFS server upgrade solves the GNU gold issue and if not look for a better solution