RCF Computing issues

Some comments and considerations
-Linux farms and RCF plans
-Disk usage
-DST tape archiving.

RCF

- At this point we have a considerable capacity
  - CRS ~ 39 machines
  - CAS ~ 34 machines
  - ~14 Tbyte central disk + several Gb distributed
- For next years FY purchase several old machines will go. We should replace these.
- Apart from this it is hard to argue for a large increase in cpu capacity – the farms are mainly idle ! (~20% usage)
- RCF wants to migrate all LSF to 5.1; installed on new cas nodes.
  - Requires stopping and emptying queues.
  - Will do one experiment at a time.
- Upgrade of AFS system – likely to Kerebos 5, will require cell name change – and more or less simultaneous upgrade of clients at your home institution. It will also give one time password access to RCF. This summer.

CRS software

- The underlying crs_software is being rewritten. At least initial the user interface will be the same, but the underlying file-transfer an job distribution will be based on Condor I.e. a first step in moving towards standardized GRID technology.
- The person in charge is Tomasz Wlodek.
- Time scale is several month. Initial testing with realistic Brahms scripts and files has started.
- Progress report will be given Monday June 9.

Redhat upgrade.

- It is the plan to upgrade the Linux farm software before the next run.
- The choice is at this point
  - RH 7.3 + gcc 3.2 (ATLAS requirement)
  - RH 8.x + gcc 3.2
  - RH 9.0 + gcc 3.2
- There is obvious an advantage to move forward with systems that we would also employ on institutional clusters.
- Test machines is being put up
- Kris H. has compile root + brat 2.9.4 on RH 8.0
  - Issue so far **autotools** (dependencies)
  - Moving to modern iostream usage.
  - Another issue is Cernlib, compile usage – **brag**.
- A RH 9.0 will be made available soon.
- In discussions I have advocated our pref as 9.0 (to stay on the curve and not behind).
- Comments, considerations??

Disk Usage

- My last comments go to our usage of central disk.
- It has ONLY worked so far because disk amounts been increasing and is now near 14 Tb.
  - There is absolutely no coordination on these.
  - Nomenclature is somewhat arbitrary, and cleanup is usually only done when some analysis cannot be done.
  - Lots of duplication of files. Too methods of generating file are usually also unclear.
- Need to move forward on this. Discussion has been had with Christian and Ian. Some initial ideas.
  - All non-approved official files should reside in /brahms/dataxx/scratch, which will be cleaned up automatically after some agreed upon time (3 mo?). One should know after such time if files are worthy to be promoted to official one.
  - Official pass files e.g. local tracking files, global tracking, acceptance files, should be in well defined directories; Only one place and copied to HPSS dst file class. Owner should be bramreco.
  - One of the files systems has as /brahms/u diskquota's assigned. This is the perfect place for having larger amount of output from intermediate analysis ,(Gb per active user).
  - Simulation files (breg, cdat) needs some more considerations.
  - Some file catalog is also needed for this to work I.e. when data become official-version numbering too.

File system

The RCF central disk is a collaboration shared resource with priority for such activities, not for tasks that more properly belongs to institutional computers.

Cleanup of users areas and store files should be performed when students finishes and leave collaboration.

Some work has to be done to develop a file catalog, and scripts to help with maintaining files systems. All files should be group readable at this point and once we implement a management system all scratch directory files should have group read/write access.