LHCb Requirements for the National Analysis Facility

Draft, 9.11.2008

Abstract

The National Analysis Facility (NAF) could be a very useful complement to large-scale batch job processing using the grid. With interactive login nodes, and the LHCb software environment, it would be extremely suitable for end-user analysis. Moreover, this allows for an effective and short analysis cycle, thus providing important debugging possibilities. The common platform could also significantly simplify simultaneous collaborative development of physics analyses and the corresponding software. Finally, with sufficient resources, the NAF will prove to be an important asset for the German groups to produce physics results. This document describes how the German LHCb groups could use the NAF and what requirements that imposes on the available resources.

1 LHCb Dataflow and Computing Model overview

The LHCb offline-computing model[1] describes how the collaboration expects to generate simulated events, and how to process simulated and the real data produced by the experiment. The data processing is split in several different steps for which the procedure and the required resources are specified:

- Real data (RAW), produced by the spectrometer are filtered for interesting events and those are subsequently transferred to the CERN Tier-0 centre. There, the data are distributed among Tier-1 centres for archiving and further processing. Master and replicated copies are saved on Tape only storage system.
- Simulated data (SIM) are generated at Tier-2 centres. Usually, these are digitized and reconstructed at the same site as where these are produced. The results (DST, fraction of SIM and DIGI) are transferred to Tier-1 Storage Elements (SE).
- Data are reconstructed in two steps. In the first step (reconstruction), information is produced for event stream selection (rDST). In the second step (stripping), the data are filtered for particular event-type criteria, resulting in the final data representation (DST). The data, suitable for individual future analysis, are distributed to all Tier-1 disk based SEs.
- Principally, analyses are performed on Tier-1 centres using those DSTs. Some Tier-2 centres with sufficient storage resources can also be involved.

All steps use the Grid infrastructure ([2]), which is under control of LHCb's Distributed Infrastructure with Remote Agent Control (DIRAC, [3]). DIRAC provides monitoring and bookkeeping as well.

Simulation and reconstruction are organized and operated by central LHCb Grid Team.

2 Analysis schemas

Several different analyses schemas are available for end-users.

2.2 DST data analysis

Input DST/SIM data are processed using the LHCb analysis framework, called DaVinci. Input data can be on any supported media (POSIX, d-cache, xrootd, etc.) and access is transparent for the end-user. The LHCb computing model foresees executing analyses jobs on the Tier-1 centres under

DIRAC control with GANGA ([4]) as the user interface to DIRAC. With GANGA, the end-user will be able to submit jobs to either the local batch system, or to the grid in general. Job splitting is supported by the GANGA front-end.

Results from the processing are ROOT files with histograms and/or ntuples, which are saved on media with POSIX access.

2.2 Ntuple data analysis

Sometimes processing general data directly is inefficient. In this case, data of interest are first saved into ntuples and than processed by separate ROOT jobs. German LHCb groups are currently actively developing a separate framework for that schema, with the possibility to share generated ntuples between users. Media with POSIX access is preferred for that schema, since there is no transparent way to use other methods.

2.3 Other schemas

For some specific studies (alignments, trigger studies, etc.) in which several members of the German LHCB groups are involved, suitable dataflow and job submission models are hard to define. Therefore, the analysis environment must be flexible and support generic operations.

3 Service requirements for the NAF

1. Running analysis jobs on Tier-1 centres, in parallel with production jobs, from LHCb and other experiments, puts strict requirements on content and quality of such jobs:
    1. NAF provides an environment for developing and testing analysis software.
    2. NAF is a facility very suitable for user customized jobs, contrary to the grid, which supports only standard analysis jobs.
2. End-user requirements can be in disagreement with current common experiment environment:
    1. NAF provides user controllable software installation
    2. NAF provides sufficient shared and individual resources for user's custom needs
3. Personal support and guaranteed resource availability:
    1. Due to the local nature of the NAF, the required granular support is readily feasable
    2. Dedicated national resources help free resources at Tier-1 centres and in turn

3 Resource requirements

To fulfil the specified service requirements, the NAF should provide sufficient amount of resources for the on average 20 users from the German LHCb groups. The operation of the experiment is expected to occur in two phases; during the first phase, the detector will be commissioned, here it is expected that in addition to physics analyses, the NAF facility will also be used for commissioning the track reconstruction and various detector calibrations. For this purpose, it is expected that significant RAW data will be imported and reconstructed multiple times. Therefore, the ramp-up of resources is steeper than may be expected. After the detector has been commissioned, under normal conditions, LHCb expect 140 TB of Data Summary Tapes (DST) per nominal year[1]. In Table 1, the requirements on the NAF concerning disk space is summarized. The disk space is divided in two classes; mass storage for DST files, and user space.

- Shared software area and home directories (AFS). Basic LHCb software installation, users' private installations, important output files (listed as User space in Table 1).
- General LHCb Data space (SE, d-cache). Data replicated from Tier-1 SEs by users, special datasets, see the label “Data space” in Table 1.
- User specific data (User space) and output. Here, users will store for instance their ntuples, or output files their jobs generated, included in “User space” line of Table 1.

Resource	2009	2010	2011
Data space (TB)	60	180	240
User space (TB)	40	60	60
Total:	100	240	300

Table 1: Resource table with expected mass storage space requirements,for each year individually.

- Interactive servers. Users must have possibility to login, compile the code, debug, run jobs interactively or on a batch system and inspect the results. It is imagined that 1 such server per user should be sufficient. See row 2 in Table 2.
- The number of working cores is based on the number of expected users. Based on our current experience, for each user, 4 batch nodes should suffice. (1.5GB RAM per core).
- The required network bandwidth between the NAF and the home-institutes of the participating groups, and the NAF and the LHCb T1 centres.

Resource	2009	2010	2011
# users	15	30	30
# cores (login)	20	40	40
# cores (batch)	60	120	120
Total:	80	160	160

Table 2: Resource table indicating the number of cores expected for each year.

References

[1] LHCb Computing Technical Design Report CERN-LHCC-2005-019, LHCb-TDR-011

[2] WLCG Memorandum of Understanding, CERN-C-RRB-2005-01/Rev., 15 June 2007

[3] Andrei Tsaregorodsev ’DIRAC, community grid solution’, Victoria: CHEP 2007 (id 189)

[4] J.Elmsheuser et. al, Distributed Analysis within the LHC computing, Proceedings of GES 2007, Baden Baden, May 2007, http://edoc.mpg.de/316514