



30 January 2008, Harry Renshall

CCRC'08 – Highlevel Overview for Sysadmins

Since several years the four LHC experiments have been running periodic stress tests of their planned computing operations when LHC data arrives.

They have written design documents which describe their models for how they will acquire, distribute and process their data. They differ in details but have a common base in a model of levels or 'Tiers' of resources.

Master copies of the raw data are kept on tape at CERN, the Tier0 site. CERN also does a first reconstruction pass on the raw data using approximate calibration constants. The final ones are not ready for weeks or even months but will be used for a second reconstruction which is done entirely at the next level, the Tier1 sites. A complete copy of the raw data is also exported to the Tier1 sites, where it is also put on tape. There are 11 such sites but they do not support all the experiments.

10 Tier1s support ATLAS, the largest experiment in data rates and volumes, 7 support CMS, 7 support ALICE and 6 support LHCb, the smallest experiment. The Tier1 resources are not evenly distributed e.g. BNL (near New York) will provide 28% of ATLAS Tier1 resources while CNAF in Bologna will provide only 5%.

Raw data is sent from the experiment pit to disks in the CERN computer centre. There it is migrated to tape under the Castor system, reconstructed by batch jobs using LSF and sent to the experiment Tier1 sites using the Grid File Transfer service. At the Tier 1 it is copied to their local mass storage systems (the most common is called dcache) and when the calibration is ready it is recalled from tape and reconstructed using grid submitted batch jobs. The output data is then ready for physics analysis.

There is a third level, the Tier2 sites, which then get a copy of the analysis files for their physicists. These are about 250 smaller computer centres, mostly at universities. They also generate simulated (Monte Carlo) events that are needed by the experiments and which are stored on tape at Tier1 sites and, usually, also at CERN.

Over the years the experiments, sponsored by CERN-IT, have exercised the various components of their computing models. For example last year ATLAS succeeded over several days in sending simulated raw data to the computer centre at the required 320 MB/sec, processing it with several thousand concurrent batch jobs and exporting the raw and processed data to their 10 Tier1 sites at about 1000 MB/sec. Some of the Tier1 put the required data to tape but not all. When all experiments take data together the aggregate rate out of CERN will be 1600 MB/sec and we will be storing about 40TB of data (about 100 full tape cartridges) into Castor per day.

The plan is to run for 100 days in 2008, from July to December.

A weakness of all the tests done is that they have never overlapped all the experiments together so never reached the full rates expected at CERN and the Tier sites and never had to cope with the lack of homogeneity of having the 4 different experiment data management and batch processing models being busy at the same time. The CCRC08 (Common Computing Readiness Challenge) tests address this by deliberately exercising all experiments computing at the target rates at the same time.



30 January 2008, Harry Renshall

The plan is to run in two phases. The first is to test all the functionality involved and attempt to build up to the target rates as far as resources allow (the full 2008 resources at the remote sites will not be available till 1 April and some sites will be late) and this will run from 4 to 28 of February. We expect to find many problems, software, hardware and organisational, in this period but be able to correct them over March and April for a second run during the whole of May which is planned to rapidly reach the full 2008 data rates over the whole distributed processing and then run stably for many weeks. The May exercise is hence a full-scale dress rehearsal for the accelerator run due to start in July.

The accelerator should then be in action 24 by 7 for many months with periodic stoppages of several days each. The current operational plans mean that most physics data from the experiments will arrive during a 12 hour period overnight but the batch processing, both local and over the grid, and data export will be a continuous operation. The sysadmin team will play a vital role in ensuring these continuous operations.

A Twiki giving many details of CCRC08 is to be found at <https://twiki.cern.ch/twiki/bin/view/LCG/WLCGCommonComputingReadinessChallenges> and an overview of the whole LHC Computing Grid (LCG) project is at: <http://lcg.web.cern.ch/LCG/>