Offline Computing Shifts

Dario Barberis, 5 September 2005

During data-taking, and for some of the activities also during the rest of the year, there will be the need to run offline computing shifts. This document is meant to raise the discussion within the Collaboration, so that we can put in place an organisation that is sufficient to cover the operation needs, as specified by the Computing Model (see the Computing TDR).

We can distinguish a few different types of operations:

1. Tier-0 operation (in quasi-real time)
2. Re-processing at Tier-1s
3. Simulation productions (mainly at Tier-2s)
4. Organised group analysis mass productions

Tier-0 operations

Tier-0 operations have to take place in quasi-real time (except for Heavy Ions runs, when massive processing takes place in the shutdown period of 3-4 months after data-taking).

The sequence of actions is as follows:

- The end of run (or more likely the end of spill) triggers the submission of a set of calibration jobs (done by the offline shift at the Pit)
- Calibration experts check the results of the jobs, validate the calibration for the given run(s) and trigger the insertion of the calibration constants into the conditions database
- The offline shift submits reconstruction jobs
- Reconstruction experts check the reconstruction results
- The offline shift distributes reconstructed data to Tier-1s and Tier-2s

All these actions are interleaved, as we assume that the full processing of a given run can take up to 48 hours. On any given day, calibrations are run on the data of the day, reconstruction on the data of the day before, etc.

Data movement is also a substantial task. RAW data files have to be imported from the online system, copied to Castor tape and distributed to Tier-1s. ESD/AOD/TAG data has also to be copied to Castor tape and to Tier-1s (and Tier-2s for AOD and TAG). Checking that all files are transferred correctly and registered in the data catalogues is part of the shift duties.

The nominal data rate is 200 Hz for 50k seconds/day, i.e. 10^7 events/day in 10k files. The number of jobs to run every day is 10k for reconstruction (4-5 hours/job), plus probably a few 1000's for calibration.

The number of file transfers is 30k for RAW (once in, twice out), 30k for ESD (to tape and to two Tier-1s), and at least 10 times the number of AOD+TAG files (to be defined) for streamed reconstructed objects.

The operation model should foresee two groups of people: “experts” and “shift workers”. We should define a small group of 4-5 “experts” who can be on call to sort out serious problems. Some of the experts can be the designers and developers of the Tier-0 production and data management system.

There should be 3 people on the offline shift:

- a shift leader, who has already some experience (has had some training and has run shifts before) and can sort out trivial problems
- a person responsible for data management operations
- a person responsible for job submission and monitoring

In addition, experts of calibrations and reconstruction (appointed by the Data Preparation area?) must be on call and present when needed to check job outputs.
We expect the offline shift people to be placed close to the control room at Point 1.

Re-processing at Tier-1s

Re-processing work is similar to Tier-0 operations, i.e. the same jobs steps have to be executed. Every Tier-1 will re-process the RAW data it has in custody, save all output files to local tape storage, and distribute the output ESD files to another Tier-1 and the AOD+TAG files to all Tier-1s and Tier-2s.

The number of jobs to be run per day will be in the first couple of years similar to Tier-0 operations (a few 10^4), but it will increase in subsequent years, as older datasets will be reprocessed with newer software.

We can imagine having again two teams of people, “experts” and “shift workers”. Experts would be on call and intervene if and when needed. Shift workers, as for Tier-0, would submit and monitor jobs and take care that the data management system works correctly. Shift work can be run remotely but people have to realise that it is a real commitment (as if they were at CERN running online shifts). Three people/shift can take care of ~10^4 RAW data files (or equivalent) per day. In addition, there must be an ATLAS person on call for each Tier-1, to guarantee prompt trouble-shooting.

Simulation Productions

Simulation productions will take place mainly at Tier-2s. Tier-1s can help in case they have spare capacity; the same is true of Tier-3s. We foresee a centralised production organisation, parallel and similar to Tier-1 re-processing productions.

The current model foresees 20% simulated events with respect to the number of RAW events collected. The event generation step should be run independently (using the distributed production system) by the physics groups that need the simulations. The following steps (detector simulation, pileup and digitisation, reconstruction) will consist of a number of jobs that is similar to real data re-processing. The same kind of organisation, with a core group of experts on call and a shift of 3 people round the clock, should be able to cope with this workload.

Experts from the groups that requested the simulation will have to be available to check and validate each job step.

Organised group analysis mass productions

It is difficult at this point in time to estimate the number of jobs and the magnitude of data management work for organised physics analysis productions. We assume here that each activity group will have a small group of production managers who can take care of job submission and data management. Probably the same core group of experts as for simulation productions can act as consultants in case of trouble.

Summary of shift needs

- 3 people on shift for Tier-0 operations
- 3 people on shift for Tier-1 re-processing
- 3 people on shift for simulation productions
- Production managers of physics groups run their own productions

- Group of experts on call for Tier-0 operations
- Group of experts on call for world-wide operations (Tier-1s and Tier-2)
- Contact people on call (at least) at every Tier-1