ATLAS TIER-0 SCALING TEST

L.Goossens, A.Nairz, G.Poulard, M.Branco, D.Cameron

CERN, Geneva, Switzerland

Abstract

To validate its computing model, ATLAS, one of the four LHC experiments, conducted in the period December 2005 - January 2006 a first Tier-0 scaling test. The Tier-0 is mainly responsible for prompt reconstruction of the data coming from the Event Filter, and for the distribution of this data and the results of prompt reconstruction to the Tier-1s. Handling the unprecedented data rates and volumes poses a huge challenge on the computing infrastructure. In this paper we report on our experiences in a first attempt to scale up to nominal operation over a period of three months.

INTRODUCTION

The ATLAS Computing Model [1] states that the Tier-0 (T0) is responsible for the following operations:

- first pass reconstruction of the data coming directly from the on-line event filter farm (RAW data), producing respectively Event Summary Data (ESD), Analysis Object Data (AOD) and tag data (TAG), with decreasing size and detail,
- processing of the events selected for the calibration and alignment stream into calibration and alignment (time-dependent) data,
- archiving of all produced data products on tape,
- uploading of TAG data into the T0 tag database,
- uploading of calibration/alignment data into the T0 conditions database,
- distribution of all data products to the ATLAS Tier-1 (T1) sites.

The Computing Model foresees to distribute one copy of the RAW data to one T1, two copies of the ESD to two T1s and one copy of AOD, TAG and calib/align data to all 10 T1 sites.

In the following sections we report on our experiences in a first attempt to scale up to nominal operation over a period of three months. The first two sections describe the architecture of the ATLAS T0 and the scope, scale and plan of the scaling test. Next, two sections describe in more detail the test results. A final section presents our conclusions and future plans.

ATLAS TIER-0 ARCHITECTURE

Figure 1 shows the ATLAS T0 architecture. In the middle is the actual T0 which consists of the following components:

- a Grid-enabled hierarchical mass storage system (Castor2 [2]) providing all required storage means
- a Grid-enabled local file catalog (LFC [3])
- a CPU-farm (LSF [4])
- a conditions database
- a tag database
- a T0 manager (TOM) and its associated database
- a dedicated instance of the ATLAS Distributed Data Management system, called DQ2 [5,8]

On the left of the diagram the output bufer of the Event Filter farm is shown. An asynchronous process copies the produced RAW files into the T0 hierarchical mass storage system and registers the files in the local file catalog and the additional DQ2 catalogues.

Based on the RAW data arriving, TOM will define and orchestrate all necessary jobs: reconstruction jobs, AOD merging jobs, calib/align jobs, tag uploading jobs, etc. The jobs are run on a dedicated LSF CPU farm.

TOM will also enter in the DQ2 instance all necessary T0 to T1 export tasks taking into account the agreed data volumes for each T1.

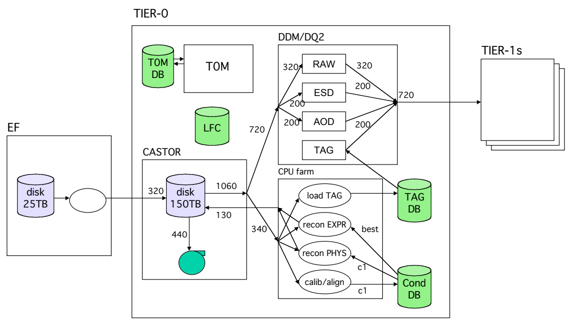

Figure 1: ATLAS Tier-0 architecture.

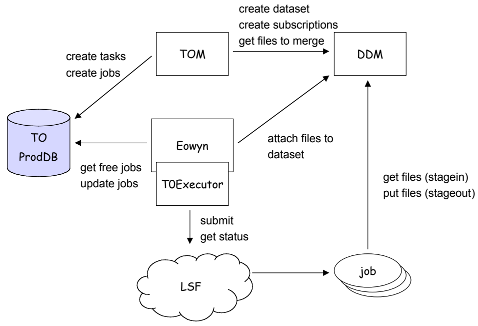

Figure 2: TOM architecture.

Figure 2 shows the internal TOM architecture. The T0 manager is actually a light-weight layer on top of the regular ATLAS production system (ProdSys) [6]. TOM only needs to define jobs in the T0 ProdDB based on the data arriving (as recorded in DQ2), ProdSys will take care of running the jobs and interacting further with DQ2, e.g. to register the produced outputs.

Note that ProdSys features a 'facility'-neutral supervisor component and 'facility'-specific plug-ins called executors. The standard ProdSys supervisor called Eowyn was used, whereas a new custom executor interfacing to LSF was developed.

Figure 3 shows the internals of the DQ2 [8] instance. DQ2 is the ATLAS Distributed Data Management system, responsible for handling file-resident data from raw data export and archiving, through global managed production and analysis, to individual physics analysis at home institutes. The design layers over a foundation of basic file handling Grid middleware, a set of loosely coupled components that provide logical organization at the dataset (hierarchical, versioned file collections) level, supporting in a flexible and scalable way the data aggregations by which data is replicated, discovered and analyzed. A combination of central services, distributed site services and agents handle data transfer, bookkeeping and monitoring. DQ2 uses LCG LFC [9] as the Grid local replica catalog for LCG sites. For the export of T0 data to the T1s gLite FTS [7] is used as the file transfer agent. Within the T0, DQ2 was integrated directly with the Castor storage at CERN, and uses RFIO as the transport protocol. This provides for efficient transfer of data from the mass storage system to the worker nodes doing reconstruction.

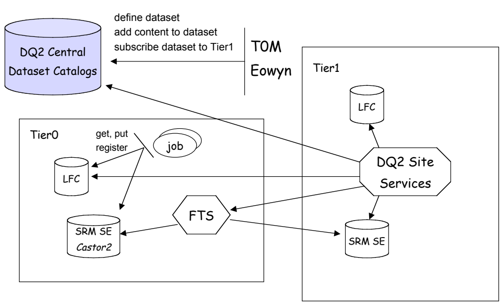

Figure 3: T0-DQ2 architecture.

SCOPE, SCALE AND PLANNING

It was known from the beginning that at the time of this first test, ATLAS would not yet be in a position to test the calib/align processing. In the end, we also did not exercise the conditions and tag database data flows. However, this still left the bulk of the data flow and operations: EF writing into Castor, ESD/AOD production on the reconstruction farm, archiving of RAW/ESD/AOD to tape, export of RAW/ESD/AOD to T1s.

At nominal rate the ATLAS T0 will run approximately 3000 reconstruction jobs in parallel each lasting ~15k seconds. This makes in total O(10k) jobs per day, producing O(10k) permanent files and O(100k) intermediate temporary files (un-merged AOD). The aggregated Castor disk IO rate is 460 MB/s for writing and 1500 MB/s for reading. Tape writing rate is 440 MB/s. Figure 4 illustrates the main data flows and rates, giving some more details.

The initial plan was to start in week 42 (Oct) and gradually ramp up to nominal rate in week 51 (Dec).

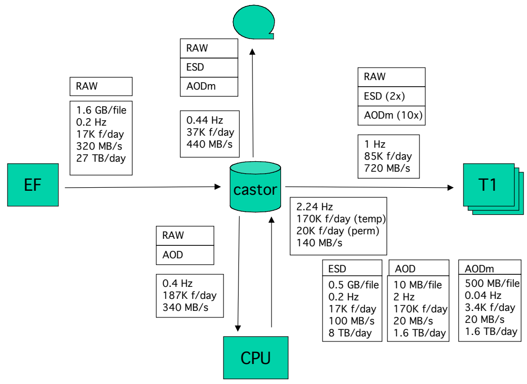

Figure 4: T0 scale (I).

OCT-DEC 2005 TEST

For the T0 tests we developed the following software components:

- a dedicated instance of ProdSys, integrating Eowyn, the T0 executor and DQ2,
- an instance of TOM, with functionality to handle both reconstruction and AOD merging jobs ,
- an Event Filter emulation (Jerry).

We also implemented two versions of 'reconstruction' software:

- the first one generating, handling and manipulating only fake (random) data files of nominal size, requiring almost no CPU time; 'reconstruction' processing was replaced by appropriately adjusted sleep times,
- the second one based on a trimmed-down version of real ATLAS reconstruction software, adjusted in a way to achieve nominal performance with respect to timing and output data sizes; input to reconstruction was specially prepared simulated RAW data.

As the main focus of the tests was on exercising the T0 dataflow, mostly the fake version was used.

The size of the test-bed gradually ramped up until in the last weeks of December we had 10 (EF farm) + 120 (reco farm) nodes available. The number of disk servers evolved from 4 to 8 in the course of the Oct-Dec test. At the start of the test a small number of tape drives was used. This limited tape-writing capacity was however quickly insufficient to keep up with the disk writing and tape writing was switched off altogether.

The ramp-up of the data rates proceeded on schedule until the beginning of December. The main milestones reached were:

- week 21 Nov: reached about 10% of nominal T0 internal rates with CPU intensive 'real' reconstruction jobs,
- week 5 Dec: reached 273 MB/s aggregated read/write rate inside T0 (>30%); rates like this were reached several times in the course of the test: see figure 6 as an illustration (from 19-20 Dec; note that only writing rates are shown),
- week 12 Dec: reached peak rate of 220 MB/s (30%) in T0 to T1 export, involving up to 7 T1s simultaneously and 8 T1s in total (see figure 7)

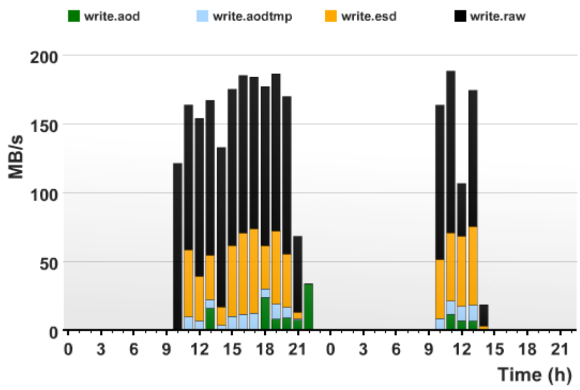

Figure 6: T0 internal data flow (19-20 Dec 2005).

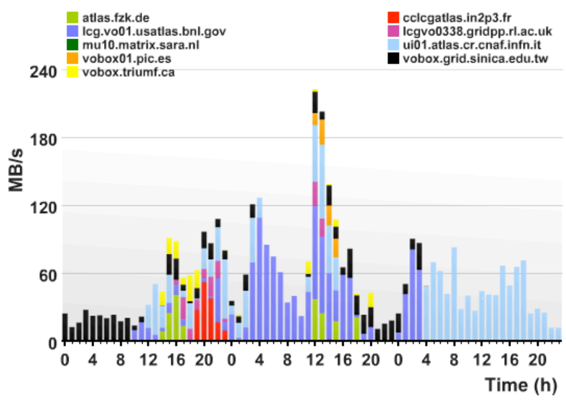

Figure 7: T0 → T1 export (15-18 Dec 2005).

Attempts to reach higher data rates failed and were traced back to the network configuration, which simply did not allow for higher rates without extensive reconfiguration efforts. Given that a major network upgrade was planned during the end of year period, it was decided to stop the T0 test.

Other major problems encountered were:

- a bug in the Castor2 file purging algorithm,
- slow deployment of DQ2 onto T1s,
- instabilities on T1 storage elements,
- general difficulty in monitoring Grid data transfers.

Given that we had reached rather easily the limits of the network configuration, preventing in fact other components to be pushed to *their* limits, we negotiated with CERN IT an additional test period in January 2006, after major hardware upgrades to the test-bed would have been done.

JAN 2006 TEST

For the January test period our hardware configuration consisted of:

- 20 (EF farm) + 200 (reco farm) LSF nodes,
- 10 Gb network connectivity,
- 24 disk servers;
- 12 tape servers.

Contrary to the Oct-Dec 2005 test, T0 to T1 export could not be resumed in January 2006, due to resource conflicts with the LCG Service Challenge 3 (SC3) re-run taking place at the same time. The test therefore concentrated on exercising T0 internal data flows: reading from and writing to the disk servers and, included again, writing to tape.

With the new hardware configuration we reached without major problems nominal rates on all internal T0 data flows, first with about 100 parallel jobs only generating the nominal data flow, and subsequently with 3000 parallel jobs sleeping to last in total the expected nominal wall-time (15k secs). Figures 8 and 9 show monitoring snapshots taken for the run on Jan 28-29, 2006.

The main problem encountered during the January 2006 test was that at peak operation, the CERN LSF system had difficulty in coping with our additional load of 3000 jobs. We shared the system with all other CERN users, who usually run about the same number of parallel jobs *in total*. Response times for both job submission and job state polling became too long to sustain the required job submission rate and hence total number of jobs (see Fig. 9, top graph between 22:00 and 04:00).

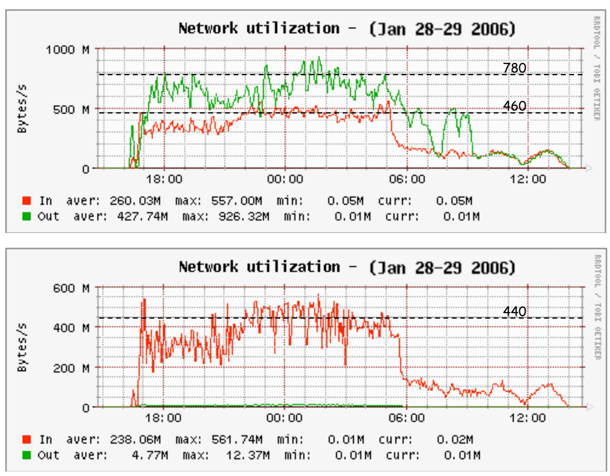

Figure 8: Castor2 reading/writing.

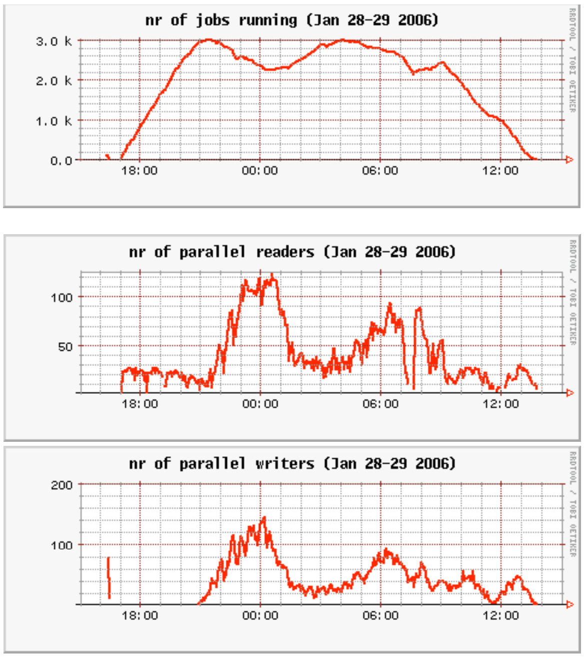

Figure 9: parallel readers/writers/jobs.

CONCLUSIONS AND FUTURE PLANS

At the end of the January test we reached full nominal rates on all major T0 data flows except the T1 export. This is a major success and above expectations. It is a pity that the T1 export could not be exercised due to incompatibility with the SC3 re-run, as our collected data suggests that also on this data path very high, perhaps even nominal, rates could have been achieved.

However, a large fraction of the time something was not working, leaving few and short time-windows to test. Indeed, we depend on many external services and *all* of them need to be working. We count on major improvements in this area for the next round of testing.

Many problems were found and addressed. Castor2 and the network configuration seem to be OK now. The LSF batch system may require some more tuning effort.

Three more T0 tests are planned for 2006. The most important one will take place in June. Its additional goals are: exercising the ATLAS pit to Castor connection, realistic conditions and tag DB data flows, comprehensive real-time monitoring in place, automatic and continuous operation for O(week), and exercising of induced error scenarios. An earlier intermediate test is planned for April and a later one in September, to consolidate the results achieved in June.

ACKNOWLEDGEMENTS

We would like to thank our colleagues from CERN IT/FIO for their contribution to making this first T0 test a success.

REFERENCES

[1] ATLAS Computing Technical Design Report (TDR), CERN-LHCC-2005-022, July 2005, http://atlas-proj-computing-tdr.web.cern.ch/atlas-proj-computing-tdr/Html/contents.html.

[2] CERN Advanced STORage Manager 2 (CASTOR2), http://castor.web.cern.ch/castor/DOCUMENTATION/USERGUIDE/.

[3] The LCG File Catalog (LFC), http://spi.cern.ch/extsoft/packages.php?pkg=lfc

[4] CERN Batch Services (LSF), http://batch.web.cern.ch/batch.

[5] ATLAS Distributed Data Management (DDM), http://uimon.cern.ch/twiki/bin/view/Atlas/DistributedDataManagement.

[6] ATLAS Production System (ProdSys), http://uimon.cern.ch/twiki/bin/view/Atlas/ProdSys.

[7] P. Kunszt, P. Badino, G. McCance and R. Brito da Rocha, The gLite File Transfer Service: Middleware lessons learned from the Service Challenges, CHEP 2006, Mumbai.

[8] M. Branco, D. Cameron. and T. Wenaus, A Scalable Distributed Data Management System For ATLAS, CHEP 2006, Mumbai.

[9] J-P. Baud, J. Casey , S. Lemaitre, C. Nicholson, G. Stewart, LCG Data Management: From EDG to EGEE, UK e-Science All Hands Conference, September 2005