ATLAS detector calibration model – preliminary subdetector requirements

Richard Hawkings, Fabiola Gianotti

Final version (28/2/05)

1 Introduction

This document gives an overview and summary of subdetector calibration requirements, from online calibration performed in the readout drivers (RODs) and high level trigger (HLT), through calibration steps needed before prompt reconstruction, and subsequent offline calibration. The information is extracted from subdetector responses to a calibration questionnaire circulated in December 2004. This requested the subdetectors to indicate their calibration plans at various stages in the data processing, including use of dedicated calibration triggers (e.g. test pulsers), special calibration streams, offline access to RAW, ESD and AOD data, and associated computing resource requirements. At the time of writing, all detectors have responded at some level, though the amount of detail varies, and some subdetectors are much further on than others in their level of understanding of calibration requirements. Some of the information in this note is therefore incomplete, and this note should be regarded as giving a snapshot of the situation in February 2005.

The results of the survey are summarised in the form of tables and associated text. The assumptions of the computing model document [1] for event sizes, luminosity, *etc.* have been taken where necessary, and CPU resources have been quoted in kSI2k units, assuming the canonical CPU farm node in 2007/8 to be a dual 8 GHz CPU unit delivering a total of 5 kSI2k of computing power.

2 ROD calibration

The subdetector plans for ROD calibration are summarised in Table 1. For calibration runs that take place outside of physics datataking, the duration of one calibration run and the approximate frequency with which they are expected to be repeated are given. Other calibrations take place during physics datataking with special triggers at the given frequency, or by spying on normal triggers; for these triggers the data is processed inside the RODs and not passed up the TDAQ dataflow chain. The CPU requirements give the amount of external processing power required, in addition to that installed inside the subdetector ROD crates (which in some cases is considerable).

System	Duration	Frequency	CPU/ kSI2k	Function
Pixel	<1 hour	1-3 per week	None	Pixel threshold/noise scan
Pixel	8 hours	Fortnightly	None	Threshold, charge calibration, leakage current
SCT	1 hour	Daily	None	Frontend electronics: thresholds and delays
TRT	physics		None	Dead/noisy straws, t0 and drift monitor only
LAr	15 mins	Between fills	5	Ramp runs – pulse all channels
LAr	~4 hours	Monthly	5	Delay runs – record calibration pulse shapes
LAr	physics		5	Pedestal, noise, autocorrelation
TileCal	10 hours	Monthly	None	Cesium source calibration
TileCal	30 mins	Daily-monthly	EF part	Laser, charge injection and pedestal run
TileCal	physics		None	ROD-level monitoring of minimum bias
MDT	physics			ROD-level monitoring (e.g. dead/noisy channels)
RPC	<½ hour	Daily	None	Pulser and random triggers (dead/noisy channels)

TGC	physics	Normal trig	none	t_0 calibration, trigger and chamber efficiency
CSC	< 1hour	Daily	none	Pulser for inter-channel calibration
L1calo	?	Mainly at start	EFpart	Calibration & timing with test pulses

Table 1: ROD-level subdetector calibration requirements.

Notes relating to particular subdetectors:

- The pixel threshold and calibration scans will be performed using dedicated CPU power in the DSPs and SBCs in the ROD crates, so no significant external CPU power will be needed.
- The SCT has now decided to install dedicated CPU power in the ROD crates to handle online calibration runs.
- The TRT performs considerable monitoring at ROD level, but no calibration.
- The LAr performs several types of dedicated calibration runs outside physics – in steady state running, the ramp runs will be performed routinely between fills, delay runs monthly during accelerator downtime and pedestal/noise/autocorrelation runs using random triggers in physics datataking. Eventually, all the processing will be done online and in the partition master (hence the allocation of only ~1 CPU box for each task), but initially large amounts of data will be written offline (200 GB at up to 2 GB/second for ramp runs, and 25 TB at a similar rate for delay runs) to cross-check the procedures. This will require considerably temporary computing resources.
- The special Tilecal runs outside of physics (cesium, laser, charge injection and pedestal run) require data readout from the ROD via the ROS and event filter. One event filter subfarm should be dedicated to the TileCal during these periods.
- The TGC performs ROD-level t_0 calibration, and requires the bunch structure of the LHC (location of empty bunches) as input. Some trigger and chamber efficiency checks can be performed at this level, providing data can be written directly from RODs.
- The CSC performs ROD-level channel intercalibration using charge injection with a pulser system outside of physics runs. The data can be processed offline or in the RODs, where adequate CPU power is available.
- The level 1 calorimeter trigger will require significant calibration and timing-in work with test pulses generated by the calorimeters. This will take place mainly during startup, with subsequent emphasis on monitoring. The possible overlap with calorimeter test pulse (ramp) runs needs to be understood. Similarly, the CTP will require initial timing alignment with many dedicated runs at startup.

The total CPU power required so far identified is of the order of 15 kSI2k units, corresponding to 3 dual 8 GHz farm nodes, and not requiring significant dedicated resources beyond subdetector workstations. Some event filter processing is also required for ROD-based calibration runs outside of physics datataking.

3 HLT calibration

The subdetector plans for calibration in the HLT system are summarised in Table 2, which lists the event type (standard physics event or special type of event), the stage of LVL2 or event filter processing, the additional CPU requirements over and above that required for standard HLT processing, and the function of each type of calibration.

System	Event type	Stage	CPU	Function
Pixel	Physics	EF, after track reconstruction	Small	Output alignment info for selected high p_T tracks (common ID task)
SCT	Physics	As for pixel	Small	As for pixel
TRT	Physics	As for pixel	Small	As for pixel
TRT	Physics	EF, after track reconstruction	Small	Monitoring of RODs, maybe output calibration updates
TileCal				

MDT	LVL1 μ	End of LVL2 μ trigger algorithm	0.2ms overhead LVL2	Output information in μ RoI (also used by RPC, TGC, CSC)
RPC	LVL1 μ	As for MDT		As for MDT (for level 1 trigger calibration)
TGC	LVL1 μ	As for MDT		As for MDT
CSC	LVL1 μ	As for MDT		As for MDT
MDT	LVL1 μ	After special LVL2 trigger	4 kSI2k	Alignment of MDT small chambers
L1calo	Calo pulser	EF, partial built events	Small	Monitoring of calorimeter and trigger response to test pulses

Table 2: Subdetector calibration requirements during HLT processing.

Notes relating to particular subdetectors:

- The ID subdetectors use a common stream, which is produced by post-processing the track fit information from the event filter (see Section 4).
- LAr plans to perform Z→ ee calibration at the event filter level, but this needs to be further clarified.
- L1 calo needs calorimeter test pulses and partially built events – it is unclear whether this will be done in physics as well as during dedicated runs without beam.
- The MDTs require a high statistics sample (O(1 kHz)) of muon hit information in a region of interest around a muon candidate selected at level 1, for t_0 and chamber autocalibration. The most attractive option appears to be to write this stream as part of the muon level 2 trigger processing, since the necessary hit collection and processing is already done as part of the level 2 algorithm. Alternatively, it could be written directly using a special stream after level 1, with dedicated processors doing partial event building – this option is less attractive, as the processing done in level 2 to obtain the selected hits and fitted tracks then needs to be repeated offline. The stream needs to include both MDT and RPC/TGC hit information, the latter to measure the second coordinate.
- The CSCs also require a similar calibration stream based on TGC triggers, primarily to determine the CSC alignment constants. It could be output from the muon level 2 trigger processors in the same way.
- The same muon sample can also be used to perform some aspects of RPC and TGC calibration, in a very similar manner.
- Alignment of the MDT small chambers requires a special stream selected at level 2 containing high p_T muons passing through the overlap between small and large chambers (only the large chambers are aligned optically). This is processed using a small amount of dedicated resources (4 kSI2k) in the event filter, with enough rate to redetermine the chamber alignment every hour.

4 Calibration before prompt reconstruction

A significant amount of calibration processing is expected to be performed before prompt reconstruction of the bulk physics data sample in Tier 0, based primarily on the calibration streams output from the event filter during and after HLT processing. This section lists the calibration streams foreseen, together with the CPU requirements for processing them. The calibration latency before prompt reconstruction can begin is also discussed.

4.1 Calibration streams

The currently foreseen calibration streams are summarised in Table 3, which shows the system(s) making use of the data, the physics type, the fraction of the detector readout, the event rate x size to storage, and the data source. Some streams include data from only a single detector, and some require data from a restricted region of interest around e.g. a lepton candidate. This implies writing out partial events from the HLT system, depending on the trigger type and subsequent event filter processing. For all calibration streams listed except the muon LVL1-selected stream, the need to output calibration data is only apparent after event filter processing has been completed or at least started. This implies that all event filter nodes are sources of calibration stream data, and each of the overall calibration streams will have to be assembled from fragments written to all event filter SFO units. Similarly, any LVL2 processors running LVL2 muon algorithms can contribute to the muon LVL1-selected stream.

System	Stream	Readout	Rate×size Hz × kB	Source	Comments/data type
ID	Generic tracks p_T>2 GeV	ID ROI	100 × 40	EF all	Custom trackfit+hit data
LAr	Electrons p_T>20 GeV	EM ROI	50 × 50	EF all	5-sample EM RAW data
Muon	Muon p_T> 6 to 20 GeV LVL1	MDT/CSC /RPC/TGC ROI	1000 x 1	LVL2 all	Custom trackfit using trigger and precision hits
MDT	High p_T muons in large/ small chamber overlap	MDT/RPC ROI	6 x 5	EF special	MDT/RPC hits only in large / small chamber overlap regions
HAD calos	isolated hadron p_T > ~ 20 GeV	ROI (here 0.4 x 0.4)	~ 5 x 400 ?	All	Single-prong tau trigger (needs more study)
All	Inclusive e/μ p_T>20 GeV	Full event	20 x 1600	EF all	Duplicate interesting events
All	di-leptons ($Z \rightarrow ll$)	Full event	1x1600	EF all	Duplicate $Z \rightarrow ll$ events
All	Prescaled minimum bias	Full event	1x1600?	EF special	Duplicate minimum bias events

Table 3: Calibration streams output from event filter to Tier 0.

Notes on particular streams:

- The LAr dedicated electron stream with all 5 samples is required mainly for EM calorimeter timing, understanding the signal pulse shape and determining optimal filtering coefficients. The threshold value needs to be tuned – it would also be useful to have some electrons of lower p_T with a prescale factor applied (or would this be more useful for the inclusive lepton stream discussed below?). This stream will only be required during initial running, with somewhere between 5 and 10 million events being required to fully calibrate all cells.
- The LVL1 muon stream is generated either as a by-product of the muon level 2 trigger processing, or directly from dedicated partial event building, as discussed above. The nominal 1 kHz rate requires some prescaling of the LVL1 triggers, which may depend on (η,ϕ) to ensure the best use of the available bandwidth. The threshold may not be as low as 6 GeV, depending on the available trigger processor resources and luminosity. However, the TGC trigger efficiency calibration does require data down to 6 GeV, at least with some prescaling factor.
- The MDT large/small chamber overlap stream is processed directly in the event filter, but is also output for further study and refined processing.
- The isolated hadron stream is requested by both the TRT (for e/π separation studies) and the calorimeters (hadronic response studies and comparison to testbeam data). One way to provide this would be to select high p_T taus using the τ trigger algorithms at level 1 and in the event filter – these look for narrow isolated clusters with one or three associated tracks. A selection of p_T>35 GeV and 45 GeV missing E_T gives rise to around 5 Hz trigger rate at low luminosity. However, 1-prong taus will have considerable EM contribution from π^0 ($\tau \rightarrow \pi^+\pi^0\nu$), which would have to be understood. A selection based purely on isolated tracks might also be useful.
- For ease of access to the data, some subdetectors have requested a separate $Z \rightarrow ee$, $\mu\mu$ stream. These events are also available in the inclusive e/μ p_T>20 GeV stream and in the express stream.

- The minimum bias stream will be generally useful, so it is worth having a dedicated calibration stream independently of any prescaled trigger in the bulk physics stream. A rate of 0.5-1 Hz seems enough – in principle all minimum bias triggers could be directed to specific event filter nodes for processing, removing the need to collect stream fragments from SFOs.
- The calibration streams currently sum to 40-50 MB/s (but the high p_T isolated hadron stream requires more study), i.e. about 15% of the total output to storage.

4.2 Processing requirements

The CPU resources requested by the subdetectors for calibration before prompt reconstruction are summarised in Table 4 – note that the requirements of the calorimeters have not yet been assessed.

System	CPU/ KSI2k	Process
SCT+Pix	50	Derivation of silicon alignment constants
TRT	20?	Derivation of TRT alignment and calibration constants (R,t etc)
LAr	?	Fast analysis of calibration stream, resources not yet assessed
TileCal		
MDT	130	Derivation of t_0 and autocalibration parameters
RPC	10?	RPC level 1 trigger calibration, assuming results of MDT muon fits available
TGC	50	Trigger efficiency determination, chamber alignment, momentum calibration
CSC	?	Alignment stability checks
L1 calo		Monitoring of calibration and derivation of new constants

Table 4: Processing requirements for calibration before prompt reconstruction.

Notes relating to particular subdetectors:

- The inner detector aims to derive an updated set of calibration and alignment constants after every fill, which entails first processing the ID calibration stream of p_T>2 GeV tracks to accumulate residuals and other histogrammed quantities. Most of the CPU power is then used by the silicon global χ^2 minimisation algorithm, whereas the TRT requires less CPU power per iteration but has to iterate a few times. The constants are then verified by re-reconstructing on an independent part of the calibration stream (included in the overall calibration stream bandwidth estimate) within 12 hours; this also checks for any updates in the silicon-TRT relative alignment. The CPU estimates need to be verified, but seem reasonable to first order.
- The processing requirements for MDT assume a worst case of performing 10 fit iterations to determine the calibrations from scratch, and are dominated by reconstruction and database access. This assumes that the data is already pre-processed in the level 2 trigger as discussed above – if not, some additional CPU resources would be required.
- The figure for RPC is a rough estimate and assumes that the results of the initial fits for the MDT can also be used for subsequent RPC calibration.

The processing requirements at this stage sum to 260 kSI2k units, not including those from LAr and Tilecal. The computing model allocates a total 500 kSI2k to calibration activities at Tier-0 [1], so this suggests most of these resources will have to be used for initial calibration activities in preparation for prompt reconstruction, with subsequent offline calibration being done elsewhere. The disk space requirements of these activities have yet to be assessed in detail.

4.3 Latency before prompt reconstruction

The ATLAS computing model assumes a latency of approximately 24 hours between the end of a physics fill and the start of ‘prompt’ reconstruction of the bulk physics sample from this fill at Tier 0. The intervening time is foreseen for subdetectors to perform initial processing of the calibration stream samples, generate calibration constants, verify their correctness (including human cross-checking if required) and collect them ready for reprocessing. The express stream may well be processed faster, using coarser calibrations e.g. from the previous fill.

All subdetectors feel that 24 hours is sufficient to perform initial calibration, and no subdetector would gain much by waiting more time, provided that the required statistics of calibration events (corresponding to the rates given in Table 3) and the needed CPU resources are available. Refined calibration and alignment procedures aiming at the ultimate detector performance (see Section 5) require more time and access to the full reconstructed physics data sample, and the results will only be available for subsequent later reconstruction passes.

On the other hand, in the initial phase of the LHC operation a longer latency is expected, since the detectors, the calibration algorithms, the data treatment and management and the computing infrastructure will all need to be understood and exercised for the first time. It is therefore likely that the calibration streams will be processed several times, and many human checks will be required before acceptable calibration constants can be propagated to the physics streams. The various procedures will become more and more automatic with time, and the latency will eventually reach the 24-hour target.

4.4 Remote calibration

No specific requests have been made for remote processing of calibration samples before the prompt reconstruction. However, it is recognised (e.g. by the TGC groups and the LVL1 calorimeter community) that this is an interesting possibility, should CPU resources at the Tier 0 and human resources accessing the CERN infrastructure be insufficient.

In general, it is felt more likely that calibration tasks to be performed after the prompt reconstruction, i.e. those aiming at extracting the final results for the data re-processing, be geographically distributed. This would also allow the remote community to be involved in this important activity. The amount of data to be dispatched and the network requirements still need to be understood.

5 Offline calibration

The calibration streams and procedures described in Section 4 are aimed at providing calibration and alignment constants for the first-pass reconstruction of the physics data.

More refined calibrations are needed to achieve the ultimate detector performance (e.g. ID alignment to better than 5 μm, EM calorimeter uniformity at the few per-mil level, etc.). This requires more time to accumulate enough event statistics and to perform more sophisticated studies. In general, this step of the calibration uses well-known physics channels.

A non-exhaustive list of samples, i.e. those identified so far from the subdetector answers to the questionnaire, is shown in Table 5. The required data format (RAW, ESD, AOD) is also indicated, since this has an impact on the computing resources (data distribution organization, ESD and AOD definition, CPU, etc.). This list is presently rather incomplete.

System	Streams	Data format	Comments
ID	Inclusive e/μ p_T>20 GeV Z, $J/\psi \rightarrow ll$ $W \rightarrow \tau\nu$	ESD, some RAW for patrec.	Refine alignment, TRT e/π separation, track-calo matching
LAr	Electrons p_T>20 GeV	ESD,RAW	Refine calibration and E-scale
	$Z \rightarrow$ ee, eeγ, $\mu\mu$, $\mu\mu\gamma$	ESD	Intercalibration, photon scale
	γ/Z + jet	ESD	Jet energy calib via p_T balance
	$W \rightarrow \tau\nu$	ESD	
TGC	Low p_T muons from tilecal	RAW?	Muon trigger threshold curve calib.

Table 5: Examples of physics samples for offline calibration.

Notes relating to particular subdetectors:

- The LAr calorimeter expects to be able to intercalibrate the 440 towers of 0.2x0.4 granularity to better than 0.3% RMS with 50k Z→ee events.

Although not explicitly stated in the questionnaire responses, it is obvious that other samples, like Z→ μμ, tt events, etc., will be needed for offline calibration purposes. An important aspect to clarify as soon as possible is what fraction of the RAW data needs to be repeatedly accessed and processed at this stage of the calibration, because of the potential implications on the computing resources. This is well illustrated by the ID case above. In order to verify the impact of the improved ID alignment on the pattern recognition, hits, and therefore (sub-sets of) RAW data, need to be accessed. In order to minimise the number of RAW data reprocessing passes, the possibility of including track "nearby" hits in the ESD should be explored. Ultimately, the ESD event size will have to be balanced against the CPU needed to reprocess the RAW data.

Finally we note that the above samples can be accessed from various sources: the express stream, the calibration streams and the main physics stream. The time required to determine improved calibration and alignment constants is expected to be in most cases compatible with the time (2-3 months) between the prompt reconstruction and the first re-processing.

6 Conclusions

An initial survey of subdetector calibration requirements has been carried out and the results described. Although the responses to the survey are somewhat incomplete, and the level of detail varies a lot between the subdetectors, some general conclusions can be drawn. The subdetector plans are generally in line with what is foreseen in the ATLAS computing model, and no major surprises have been found. The most significant resource needs are for auxiliary processing power to aid in ROD-level calibrations, and initial calibration processing at Tier-0 before prompt reconstruction begins. Relatively little CPU-intensive calibration work is planned for the event filter itself, although it should be noted that many subdetectors plan significant monitoring at this stage, which may well require additional resources. The offline calibration requirements are much less understood at this point, as is common for other resource aspects of ATLAS offline analysis computing.

The following open issues should be followed up.

- Many of the calibration streams involve partial event building and event readout – either of only some subdetectors, or a region of interest within a subdetector, determined either by the trigger type, or by processing in the level 2 trigger or event filter. The TDAQ architecture needs to be sufficiently flexible to allow this. Related to this, the splitting of calibration streams will result in each SFO writing many smaller event output files for express and calibrations streams, as well as the main physics sample on a single file. These smaller auxiliary stream files have to be transferred to Tier-0, catalogued and processed in an efficient manner (perhaps via file merging?)
- Some subdetectors want to write significant quantities of data directly from the RODs, without going through the full event building architecture. The requirements (including special calibration runs, with a single subdetector or several subdetectors together) need to be understood and taken into account by the DAQ dataflow.
- Some subdetectors want to read out extra RAW data at startup (e.g. LAr would like to read all 5 samples for all events in physics datataking for the first months)– since the bandwidth between the event filter and Tier-0 is finite, will this lead to a reduction in event rate? How will this be handled?
- The selections for the various calibration streams are somewhat conceptual – more work needs to be done to study the thresholds, rates and purities with a realistic simulated event sample. Similarly, the contents of the express stream needs to be better defined.
- The single isolated track (tau trigger) sample needs special attention to understand the requirements and possibilities.

- The discussion in this note has concentrated mainly on the requirements for calibration during ‘steady-state’ physics running. It is clear that much more work is needed to understand the initial start-up phase, in particular for timing-in and initial calibration of subdetectors both with collisions and in previous commissioning phases. Some initial discussion of this subject can be found in [2].

Acknowledgments

The authors would like to thank all those in the ATLA S subdetector communities who have contributed their thoughts and expertise to this note.

---

[1] The ATLAS computing model, D. Adams et al, ATL-SOFT-2004-007.

[2] Setting up the timing of ATLAS, N. Ellis, P. Farthouat, K. Nagano, T. Wengler, EDMS document ATL-DA-ON-0001.