CERN-2012-022
LHCC-I-023
December, 2012

ATLAS

Letter of Intent

Phase-II Upgrade

Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment

ATLAS Collaboration

ABSTRACT:
This Letter of Intent presents a plan for preserving and improving the current detection capabilities of the ATLAS detector to meet the challenges and take advantage of operating at the High Luminosity LHC (HL-LHC). From 2024, the HL-LHC will provide unprecedented pp luminosities to ATLAS, resulting in an additional integrated luminosity of around 2500 fb^{-1} over ten years. This will present a unique opportunity to substantially extend the mass reach in searches for many signatures of new physics, in several cases well into the multi-TeV region, and to significantly extend the study of the properties of the Higgs boson.
The increased luminosity and the accumulated radiation damage will render the current Inner Tracker no longer suitable for long term operations. It will need to be replaced with a new all-silicon tracker to maintain tracking performance in the high occupancy environment and to cope with the increase of approximately a factor of ten in the total radiation fluence. New technologies are used to ensure that the system can survive this harsh radiation environment and to optimise the material distribution, while the new readout scheme allows the implementation of a track trigger contributing to the improvements in the ATLAS triggering capabilities.
The very high luminosities also present significant challenges to the operation and performance of the rest of the detector systems as well as the trigger; the consequent high number of collisions per crossing will degrade the performance of ATLAS unless the LAr and Tile calorimeters and the Muon Spectrometer readout systems are upgraded. A new trigger architecture will be implemented exploiting the upgrades of the detector readout systems that will maintain and improve the event selection.
The increased luminosity may also degrade the performance of the forward calorimetry. Options for upgrading the hadronic endcap calorimeter readout electronics and the forward calorimeter detector design are being investigated to address the performance degradation, if this proves to be required.
Finally, the computing and software must be upgraded to meet the challenges of the increased luminosity and changes in computer architectures.

KEYWORDS: ATLAS, LHC, HL-LHC, Upgrade, CERN.

Contents

1. Executive Summary

The High Luminosity LHC (HL-LHC) will begin collisions around 2024 and will provide a levelled instantaneous luminosity of 5×10^{34} cm^{-2} s^{-1} with the aim of delivering an additional 2500 fb^{-1} to ATLAS over ten years [1]. The HL-LHC will remain the proton-proton discovery machine at the high energy frontier, and provide a large dataset for precision measurements. The upgrades of the LHC will first see an intermediate stage: Phase-I, when the peak instantaneous luminosity will increase to 2.2×10^{34} cm^{-2} s^{-1} (around twice the nominal) and delivering around 300-400 fb^{-1} integrated luminosity by 2022. The proposed ATLAS Upgrade programme for Phase-I operation is described in [2].

This Letter of Intent describes the proposed Phase-II upgrades of the ATLAS detector that will allow operation at five times the nominal LHC luminosity, 5×10^{34} cm^{-2} s^{-1}, and the full exploitation of the physics accessible with a total integrated luminosity of up to 3000 fb^{-1}, with approximately 2500 fb^{-1} collected at the HL-LHC.

The large luminosity extends the energy scales that can be studied in high energy boson-boson scattering, to study the EWSB mechanism, and to probe for signatures of new physics predicted by models such as SUSY and extra dimensions well into the multi-TeV region. The range of new physics signatures includes: searches for high mass gauge bosons (requiring good lepton momentum resolution up to high transverse momenta); reconstruction of complex SUSY cascade decays (requiring triggering and reconstruction of low p_T leptons and identification of heavy flavours); and searches for resonances in $t\bar{t}$-pairs (requiring reconstruction of leptons and heavy flavours in highly boosted topologies).

The large data sample will allow significant improvements in the precision of the measurements of the Higgs couplings, achieving precisions in a range between 5% and 30%. This will also allow rare channels to be measured: $H \to \mu\mu$, vector boson fusion production of $H \to \gamma\gamma$ and $H \to \tau\tau$, and associated production with a top-pair: $t\bar{t}H$ with $H \to \gamma\gamma$. These additional channels increase the precision with which the fermion couplings can be measured and improve the limits on new physics that can be set from loops in, for example, $gg \to H$ and $H \to \gamma\gamma$. The full luminosity will also allow the Higgs self-coupling to be studied for the first time in channels such as $HH \to \tau\tau bb$ and $HH \to \gamma\gamma bb$ channels.

The increased instantaneous luminosity at the HL-LHC results in the expected mean number of interactions per bunch crossing increasing from $\sim$55 at 2×10^{34} cm^{-2} s^{-1} to $\sim$140 at 5×10^{34} cm^{-2} s^{-1} (assuming a bunch crossing time of 25 ns) and the consequent increase in the integrated luminosity requires a detector able to operate after exposure to large particle fluences. These impact the experiment through increased detector occupancy and radiation damage. To allow for some safety margin, the design studies for the proposed upgrades assume a maximum instantaneous luminosity of 7×10^{34} cm^{-2} s^{-1}, 200 pile-up events, and an integrated luminosity of 3000 fb^{-1}over ten years where appropriate.

At the end of the current LHC programme the ATLAS detector will be running with components that are 15-20 years old. After the intermediate Phase-I upgrade, the silicon tracking systems will be approaching the end of their lifetimes: 700 fb^{-1} for the strip system and 400 fb^{-1} for the pixel system. Moreover, the higher luminosity will increase significantly the occupancies in both the silicon detectors, and the occupancy in the straw tube transition radiation tracker (TRT) will

reach $\sim 100\%$ at 5×10^{34} cm^{-2} s^{-1}, severely compromising the tracking performance. Therefore the need for good performance in vertex and track reconstruction, lepton identification and heavy flavour tagging, even in the high occupancy and radiation fluence environment of the HL-LHC, require a complete replacement of the current tracking system. Based on the experience gained with the current tracker, a new all-silicon tracker design has been developed. The new layout consists of pixel layers to provide pattern recognition and precision position measurements close to the vertex and which are complemented by a strip system that provides accurate tracking at larger radii. Simulations of the performance of the new layout show that the basic parameters of the tracking performance (impact parameters in $r\phi$ and z, and momentum resolution) can be preserved and even improved in processes with 200 pile-up events, and that the physics performance, such as the b-tagging performance and the mass resolution, is maintained in this high pile-up environment. The system design has been optimised to minimise the material by using modern light carbon fibre-based engineering materials as support structures, and by a careful design of the services and of their routing. Developments of silicon sensor technologies have been shown to provide the required radiation hardness, complemented by the use of modern ASIC technologies to develop front-end electronics with the necessary radiation tolerance. The readout of the tracker must be upgraded to provide the increased bandwidth required to cope with the higher occupancies, and to accommodate the proposed TDAQ architecture and provide input for the new Level-1 track trigger.

With the large variety of opportunities for discoveries and precision measurements, it is essential that ATLAS maintains or improves its performance in triggering and precisely reconstructing the full range of physics objects (e, μ, τ, γ, jets, heavy flavour jets, E_T^{miss}) over as large an acceptance as possible at the HL-LHC luminosity. Considering the recent discovery of a relatively low mass Higgs-like particle, it is envisaged, for example, to maintain p_T-thresholds at around 20 GeV for single lepton triggers, to preserve the acceptance of key signatures such as W and Z bosons and $t\bar{t}$- pairs. Extrapolating Phase-I trigger rates to HL-LHC luminosity, while maintaining low single-lepton thresholds, would require a total Level-1 rate of at least 500 kHz, which exceeds the trigger rate of 100 kHz that can be supported by the ATLAS detector in Phase I. Several architectures were considered to address this issue; the proposed baseline architecture uses a two-step first level hardware trigger, in which the first step, based on the Phase-I Level-1 trigger architecture and hardware, provides a Level-0 accept at a rate of at least 500 kHz within a latency of 6 μs. In the second step, a new Level-1 system reduces the rate to 200 kHz, primarily using tracking information, within an additional latency of 14 μs. Studies have shown that the required reduction factor can be achieved with the proposed track trigger.

The current readout electronics of the detector systems, which operates at a maximum rate of 100 kHz with a maximum latency of up to 3 μs, is not compatible with the baseline trigger architecture described above. This mandates the replacement of the detector readout electronics to accommodate the new two-step Level-0/Level-1 architecture: the higher trigger accept rates and the extended latencies at both Level-0 and Level-1. The replacement of the calorimeter and muon readout electronics can also be used to provide the full calorimeter granularity at Level-1 and an improved muon p_T resolution at Level-0 or Level-1 using the monitored drift tube (MDT) information. As already mentioned, the present tracker readout does not have sufficient bandwidth to cope with the HL-LHC luminosity and will be replaced as part of the tracker upgrade, including the implementation of a Level-0/Level-1 architecture for a Region-of-Interest based track trigger.

The combination of a Level-1 track trigger and the additional information from the calorimeter and muons will lead to a robust trigger system capable of adapting to the HL-LHC luminosity, ensuring that ATLAS will be able to fully exploit the physics potential of the increased luminosity.

The large particle fluxes may compromise the performance of the hadronic endcap calorimeter (HEC), due to lifetime issues in the cold electronics, and of the forward calorimeter (FCAL), reducing the acceptance. Confirmation of these issues and the course of actions that need to be taken is under investigation.

The computing and software of the experiment must evolve to meet the needs of the upgraded detector systems and to be able to handle the large event samples, and to adapt to new computing hardware architectures such as multi-core processors. The large event samples will require a new approach to simulation; mixing fast simulation, based on detector parameterisations, with full simulation at the event level, to optimise speed and event size. A new software framework will be developed that will be used both offline and for the high level trigger (HLT).

The proposed upgrades to the detector systems represent major changes and require careful implementation that allows for installation of the new tracker and access to the front- end electronics of the calorimeter and muon systems. Plans allowing for the possibility of opening the HEC cryostat to allow replacement of the front-end electronics and to install the new FCAL are under evaluation as they imply a longer shutdown period, and require additional radiation precautions and provision of clean environments, which must be weighed against the possible improvements in physics performance. Radioprotection is a major issue that will have to be faced during the Phase-II installation; work is underway to understand the radiation levels expected during installation and develop safe working practices.

The studies of HL-LHC physics channels are presented in chapter 8, after describing the performance of the upgraded detector. The proposed upgrades to the trigger and data acquisition are described first in chapter 2, since these influence all the other upgrades. The proposed upgrades of the readout electronics for the calorimeter systems are then described in chapters 3 and 4; and those for the muon system and the tracker in chapters 5 and 6. The performance, layout and technical description of the proposed new tracker are also detailed in chapter 6. Studies of the HEC cold electronics possible limitations and studies of the FCAL performance at high luminosity, along with possible upgrades, are discussed in chapter 3. The development of computing and software is discussed in chapter 7. The installation plan and schedule are described in chapter 9, including evaluation of the impact of replacing the HEC electronics and the FCAL. Finally the estimated CORE costs are presented in chapter 10.

2. Upgrade of the Trigger and Data-Flow Systems

The luminosities that will be delivered by the HL-LHC will provide a significant challenge to the ATLAS trigger and DAQ system, both in terms of the rates and increased data volume. This section describes the current plans for the modifications of the trigger and data acquisition systems.

The requirements for the Phase-II trigger are closely tied to the ATLAS physics goals at the HL-LHC. Much of the upgrade of the Phase-II trigger is driven by the desire to maintain efficient triggers for isolated electrons and muons with thresholds around 20 GeV at luminosities of up to $7 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This goal is determined by the need to maintain good efficiency for the leptons produced in the decays of the electroweak bosons. The acceptance for muons from $t\bar{t}$, WH and SUSY processes as a function of the true muon p_T is shown in Figure 2.1, where an increase in threshold from 20 GeV to 30 GeV results in a reduction in acceptance of between $\sim 1.3 - 1.8$. There is also a clear physics need to maintain reasonable thresholds for tau leptons and photons. Finally, the trigger needs to be sufficiently flexible to be able to adapt to any new physics discoveries and changes in background conditions.

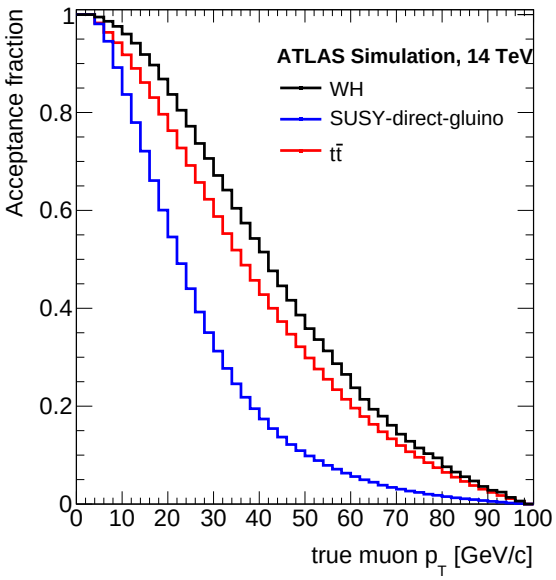


Figure 2.1: Acceptance of muons from $t\bar{t}$, WH and SUSY processes as a function of true muon momentum.

A new trigger architecture is being developed that is compatible with the constraints imposed by the detector and provides a flexible trigger with the potential to deliver the required performance. The main features are summarised below and details described in the following sections. As currently envisaged, the baseline design for the **Phase-II Trigger** is:

- a **split Level-0/Level-1 hardware trigger** with a total level-1 accept rate of **200 kHz** and total latency of **20 μs**.
- the Level-0 trigger would distribute the Level-0 accept at a rate of at least 500 kHz within a latency of 6 μs. The Phase-II Level-0 trigger is functionally the same as the Phase-I Level-1

system and consists of a feature extractor (FEX) based on calorimeter electromagnetic and jet triggers, and the Phase-I Level-1 muon trigger. The Level-0 accept is generated by the central trigger system which incorporates topological triggering capability.

- the Level-1 system will reduce the rate to 200 kHz within an additional latency of 14 μs. This reduction will be accomplished by the introduction of track information within a Region-of-Interest (RoI), full calorimeter granularity within the same RoI and the introduction of a refined muon selection based on the use of the MDT information.

- an increased use of offline-like algorithms in the High-Level Trigger (software trigger) with an anticipated readout rate of 5−10 kHz.

A block diagram of the architecture is shown in Figure 2.2. A more detailed description of the Phase-II trigger and DAQ system is given in the following sections.

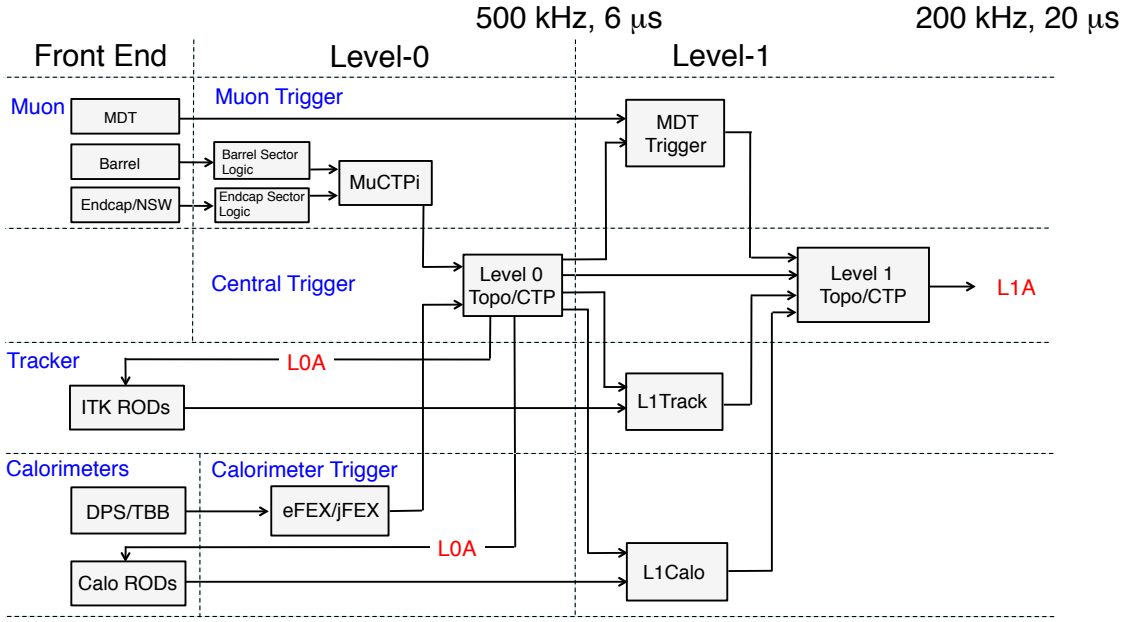


Figure 2.2: A block diagram of the architecture of the split Level-0/Level-1 hardware trigger proposed for the Phase-II upgrade. (The MDT trigger is shown as part of the Level-1 but may be used at Level-0).

2.1 Projected trigger rates for the Phase-I trigger at HL-LHC luminosities

The Phase-I Level-1 trigger [2] is composed of the Level-1 calorimeter trigger (L1Calo) and the Level-1 muon trigger (L1Muon). The resulting triggers and regions of interest are combined in the Topological Processor and Central trigger system where the Level-1 accept is generated within a latency of 2.5 μs. The trigger rate is limited to $\lesssim$ 100 kHz by the detector readout capability.

The performance of the Phase-I system at Phase-II luminosities is taken from a simulation of the functionality of the new electron feature extractor (eFEX) and extrapolations of the performance of the muon trigger based on current data. The performance of the Phase-I trigger at Phase-II luminosities is summarised below.

The rates for isolated electrons/photons and for isolated taus have been obtained from a simulation of the Phase-I L1Calo trigger with $< \mu >= 115$, corresponding to a luminosity of $\sim 4 \times 10^{34}$ cm^{-2} s^{-1}. For Phase-II luminosities, the increased granularity still gives improved discrimination for electrons and taus. The expected rates, plotted as a function of trigger threshold are shown in Figure 2.3. In the case of the tau triggers, two curves are shown, the inclusive tau trigger rate and the exclusive rate which does not include electron triggers. For these plots, it has been assumed that new Phase-I L1Calo-based triggers maintain a 95 % efficiency for electrons. For the tau leptons, the efficiency is assumed to be 95 % relative to the current Level-2 selections. The projected trigger rates for an electron trigger threshold of 25 GeV is 250 kHz without hadronic isolation and 125 kHz with hadronic isolation (`EM_VH`). In order to reduce the electromagnetic trigger rate to approximately 20 kHz, a threshold of 40 GeV would be required. This would have a significant impact on ATLAS physics at the HL-LHC and provides strong motivation for the upgrade of the trigger. The situation with tau triggers is less clear. For a tau trigger threshold at 40 GeV, most of the trigger objects are picked up by the EM trigger with a threshold of 25 GeV, but only if hadronic isolation is not used. If only EM triggers with hadronic isolation are assumed, the exclusive tau trigger rate for a 50 GeV trigger threshold is approximately 50 kHz. Either way, the combined `EM_25` and `TAU_40` trigger rate is around 250 kHz.

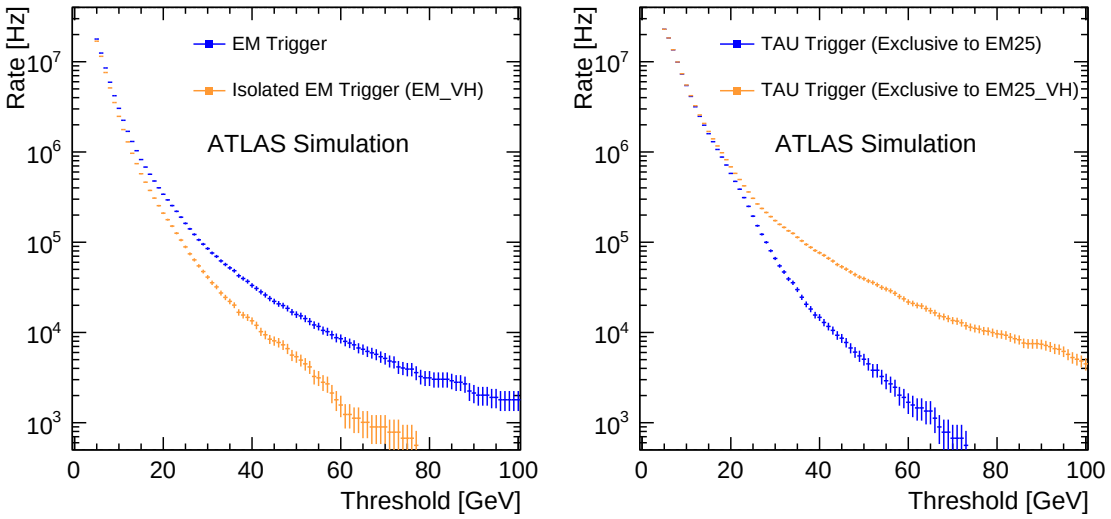


Figure 2.3: The projected trigger rates for $< \mu >= 115$ corresponding to $\sim 4 \times 10^{34}$ cm^{-2} s^{-1} as a function of trigger threshold for isolated electron and tau triggers based on a simulation of the electron eFEX. For the tau triggers, the exclusive rates for triggers which do not generate an EM trigger at the corresponding electron threshold are also shown with respect to the EM trigger at half the tau threshold, with and without hadronic isolation.

The Phase-I upgrade of the Level-1 muon trigger consists of replacing the the muon endcap inner trigger station by the New Small Wheel detectors covering the pseudorapidity range $1.2 < \eta < 2.4$. The New Small Wheel improves the tracking and results in a sharper p_T threshold that significantly rejects backgrounds from muons with low p_T. The performance of the Phase-I muon trigger at Phase-II luminosities has been estimated based on existing data supported by Monte

Carlo studies. The estimated rate for a trigger threshold of 20 GeV is at least 40 kHz, depending on the assumptions for the background rates.

2.2 The Phase-II Trigger Architecture

The requirements of the Phase-II hardware trigger are driven by the ATLAS physics goals at the HL-LHC and the constraints imposed by the readout of the detector subsystems. From the perspective of physics, the trigger must maintain thresholds for electrons and muons at between $20-25$ GeV, and should maintain sufficient flexibility to adapt to emerging physics scenarios. The constraints from the detectors are, in some ways, more difficult to define as they are closely tied to the possibility of replacing/upgrading the readout electronics and data buffering capacity. With this caveat, Table 2.1 summarises the anticipated constraints on the trigger system from the readout of the detector subsystems in terms of both rate and latency. The most stringent requirements come from the MDT. Here approximately 30 % of the electronics of the Barrel Inner (BI) Layer of the muon spectrometer is very difficult to access. The implications for the Level-1 trigger latency of the muon readout are discussed in section 5. The constraints imposed by the MDT detector system are met by a Level-1 accept that satisfies:

$$\boxed{\text{Rate} \leq 200\,\text{kHz} \quad \text{and} \quad \text{latency} \leq 20\,\mu\text{s}}$$

A number of possible Phase-II trigger architectures have been considered. One possibility would be a single stage trigger, along the lines of the current system, formed from the existing calorimeter and muon triggers with a Level-1 trigger accept (L1A) at a rate of less than 200 kHz. However, such a trigger is not compatible with the estimated rates at Phase-II luminosities. For example, Table 2.2 shows the current estimates of the Level-1 trigger rates at 7×10^{34} cm^{-2} s^{-1}, based in part on the studies described in the previous section. The trigger rate for single electrons, muons and taus alone exceeds the 200 kHz constraint imposed by the MDT. In addition to the single lepton triggers, which are necessary for triggering on decays such as $t \to bW \to b\ell\nu_\ell$, it is also necessary to maintain high efficiency for decays to di-lepton and di-photon final states, *e.g.* $H \to \tau\tau$ or $H \to \gamma\gamma$, and di-object trigger thresholds have to be kept as low as possible. For thresholds in the range $10-20$ GeV, the di-lepton/di-photon trigger rate is expected to be of the order of 100 kHz. Even if only 100 kHz is allowed for jet and MET triggers, building in some flexibility and safety factors suggests that a trigger rate of *at least* 500 kHz for the first stage of the ATLAS trigger should be assumed. With the current trigger architecture, the only way of meeting

Detector	Max. Rate	Max. Latency
MDT	~ 200 kHz	$\sim 20\,\mu$s
LAr	any	any
TileCal	> 300 kHz	any
ITK	> 200 kHz	$< 500\,\mu$s

Table 2.1: The anticipated constraints on the L1 accept rate and the total latency from the proposed Phase-II readout of the Muon Drift Tube chambers (MDT), the liquid-Argon calorimeter (LAr), the hadron tile calorimeter (TileCal) and the inner tracking system (ITK).

Trigger	Estimated L1 Rate
EM_20 GeV	200 kHz
MU_20 GeV	$>$ 40 kHz
TAU_50 GeV	50 kHz
di-lepton	100 kHz
JET + MET	$\sim$ 100 kHz
Total	500 kHz

Table 2.2: The current estimates of the expected Level-1 trigger rates at $7 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, based on the Phase-I hardware trigger system. The tau rate is the exclusive rate. The rates for the JET and MET triggers are estimates based on an extrapolation of the current fraction of the trigger budget used for these triggers.

the 200 kHz limit, imposed by the detector readout, would be to raise thresholds significantly above the electroweak scale. The alternative is to bring additional information into the trigger system.

The baseline for the ATLAS Phase-II hardware trigger is to use tracking information to reduce the electron, muon and tau trigger rates to an acceptable level, such that a L1A could be issued at a rate of 200 kHz. The baseline option for implementing a track trigger is a two stage trigger where a Level-0 accept is issued within 6 μs at a rate of 500 kHz. On the Level-0 accept, additional information, e.g. track and full granularity calorimeter information, is brought into the trigger system where it is processed with significantly longer latency generating the L1A within a total latency of 20 μs at a rate of 200 kHz, thus meeting the constraints imposed by the readout of the detector.

The proposed Phase-II trigger architecture is shown in Figure 2.2. The Level-0 trigger is almost identical to the Phase-I system, although now the L1Calo trigger is based on the digital signals provided by both the LAr and TileCal calorimeter systems, and no longer uses the previous analogue signals. The Level-0 accept triggers the full readout of the calorimeter into a new L1Calo system based on regions of interest (RoIs). Similarly, regions of the inner tracking detector are readout into the L1Track trigger. The Level-0 accept also initiates the new L1Muon Trigger using data from the MDT. The split Level-0/Level-1 trigger provides two options for the detector readout: either the data from the full latency of 20 μs could be buffered and readout on the L1A; or two pipelines could be used, where the full detector data are buffered for the latency of the Level-0 trigger (6 μs) and on a L0A the data are passed into a second pipeline. The estimated trigger rates with the baseline Phase-II architecture are listed in Table 2.3. With the split Level-0/Level-1 hardware trigger the lepton thresholds can be maintained at an acceptable level within the 200 kHz budget.

The Phase-II Level-0 hardware trigger is essentially the same as the Phase-I system. The Level-1 trigger system consists of four new elements

- **Level-1 Track Trigger:** the Level-1 Track Trigger will provide track parameters for tracks found within the RoIs found at Level-0.

- **Level-1 Calorimeter Trigger:** The L1Calo trigger would have access to the full calorimeter granularity through the readout from the calorimeter pre-processor modules on a Level-0 ac-

Object(s)	Trigger	Estimated Rate	
		no L1Track	with L1Track
e	EM20	200 kHz	40 kHz
γ	EM40	20 kHz	10 kHz*
μ	MU20	>40 kHz	10 kHz
τ	TAU50	50 kHz	20 kHz
ee	2EM10	40 kHz	<1 kHz
$\gamma\gamma$	2EM10	as above	$\sim$5 kHz*
$e\mu$	EM10_MU6	30 kHz	<1 kHz
$\mu\mu$	2MU10	4 kHz	<1 kHz
$\tau\tau$	2TAU15I	40 kHz	2 kHz
Other	JET + MET	~ 100 kHz	~ 100 kHz
Total		~ 500 kHz	~ 200 kHz

Table 2.3: The expected Level-1 trigger rates at 7×10^{34} cm^{-2} s^{-1} for the baseline split L0/L1 Phase-II trigger. The EM triggers all assume the hadronic energy veto (VH) is used. *For the photon and di-photon triggers it is assumed that the full granularity in the Level-1 calorimeter trigger will bring an additional factor 3 in background rejection power. The $\tau\tau$ trigger rate assumes a factor 2 reduction in the tau fake rate from the eFeX. The *exclusive* rates for $e\tau$ and $\mu\tau$ are not included as these will depend strongly on the exact trigger menu and trigger thresholds used.

cept. It might be sufficient to transfer the data only from a region around the RoIs identified by the Phase-I calorimeter trigger. The additional processing time available within the total latency of 20 μs would allow further refinement of the EM, tau, jet and energy sum triggers. For example the positions of the electrons would be determined more precisely which would improve the matching with track segments. Furthermore, the fine-grained calorimeter information would improve the quality of the standalone EM triggers, which will be essential in order to maintain reasonable thresholds for photons.

- **Level-1 Muon Trigger:** A L1Muon system will introduce the monitored-drift-tubes (MDTs) of the ATLAS muon spectrometer in the Muon trigger, at Level-0 or Level-1. This enables track momentum reconstruction to be performed for muons in the MDT acceptance, providing further background rejection against relatively low momentum muons.

- **Level-1 Central Trigger:** The new Level-1 central trigger would form the final Level-1 accept based on the results of the L1Calo, L1Muon and L1Track trigger RoIs.

2.3 Calorimeter Trigger

In the Phase-II upgrade, the entire calorimeter front-end and back-end electronics will be replaced, as described in sections 3.1 and 4.1. The new front-end electronics will digitise all channels every bunch crossing and transmit the data off the detector on high speed links to new calorimeter backend electronics in USA15. The fibres would be laid in a low-latency route using the holes in the shielding freed by removing the previous analogue trigger cables. The back-end electronics will process these data every bunch crossing to extract the E_T and timing of each pulse.

Figure 2.4 shows a functional block diagram for the Phase-II Level-0 calorimeter trigger. The digital signals from the calorimeters are duplicated optically and organised into overlapping detector regions for processing by the electron and jet feature extractors (FEX s). In both cases, the HCAL data are divided into regions of 0.1×0.1 in η, ϕ. The data from the four layers of the LAr ECAL (pre-sampler and three sampling layers of the ECAL) are delivered on optical links in the 1-4-4-1 arrangement, where the first two layers of the ECAL proper are segmented into 0.025×0.1 in $\eta \times \phi$. For input into the jFEX, these signals are summed in the DPS into a single 0.1×0.1 trigger tower. The total number of 10 Gb/s links into the L0Calo would be 4064, comprising 3200 fibres carrying the fine granularity EM layers and 432 fibres carrying the lower granularity data from both the EM and hadronic layers.

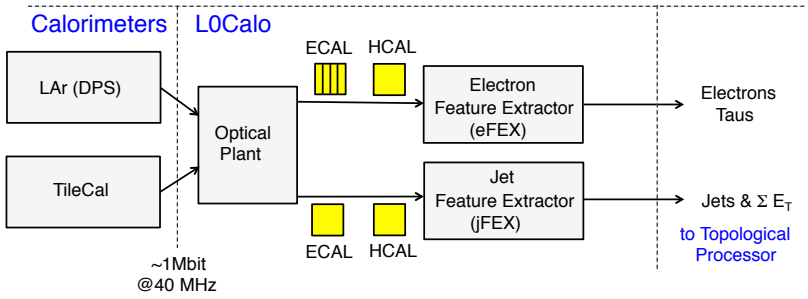


Figure 2.4: Functional block diagram for Phase-II Level-0 calorimeter trigger.

2.3.1 Level-0 Calorimeter Feature Extractor

The L0Calo trigger processing will continue to use the Feature Extractor components developed for Phase-I [2], but with modified firmware. The additional information provides further background rejection. An example that has been shown to give a $3-4$ times reduction in the trigger rate is to calculate the ratio of the energy in the second layer of the LAr ECAL which occurs in 3×2 region of $\eta - \phi$ centred on the local maximum of E_T to that deposited in the corresponding 7×2 region of $\eta - \phi$ as illustrated in Figure 2.5, centre.

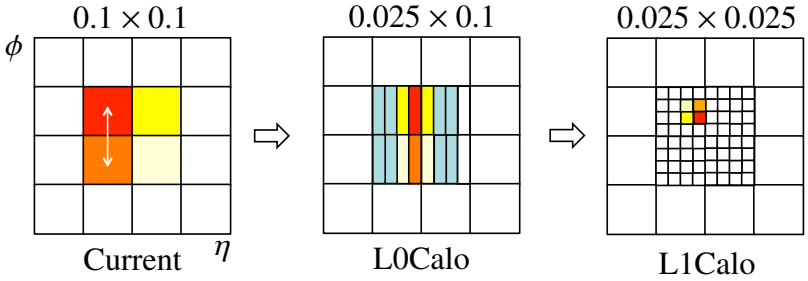


Figure 2.5: The EM granularity available in the current, Phase-II Level-0 and Phase-II Level-1 EM triggers.

2.3.2 Level-1 Calorimeter Trigger

The Level-1 calorimeter trigger will be asynchronous, processing Level-0 accepts as they arrive at an average rate of up to 500 kHz and a peak rate possibly up to 20 MHz. The calorimeter data for

every cell is stored in FIFOs on the calorimeter pre-processor module. On receipt of the L0A for each RoI identified by the Level-0 trigger, the pre-processor modules will send the data for that interaction to the L1Calo system. Each pre-processor module handles a maximum of 2000 cells and will transmit roughly 20 bits per cell, which with an L0Accept rate of 500 kHz, corresponds to a required *average* bandwidth of about 20 Gbit/s per module. (The average bandwidth into the entire L1Calo system would be roughly 2 Tbit/s arriving on up to one thousand 10 Gbit/s links.

The L1Calo system will have access to the full granularity calorimeter data and can therefore deliver improved measurements of the energies and position of the trigger objects. This will result in sharper turn-on curves and more precise matching of clusters to tracks found by the L1Track trigger. The increased granularity available will also provide further background rejection for EM electron and photon triggers. The L1Calo system could also determine more accurate estimates of the E_T of each object and could apply more sophisticated jet-finding algorithms, which are closer to those performed offline, although the possible performance gains have yet to be studied.

Two additional methods can be used to refine the identification of electrons and photons. The granularity of the data in the second sampling layer of the EM calorimeter is 0.025×0.025 in $\eta - \phi$, compared to 0.025×0.1 in the L0Calo system, as indicated in Figure 2.5 right. Secondly, the significant contribution to the fake EM trigger rate due to photons from $\pi^0 \to \gamma\gamma$ decays can be suppressed using the 0.003125×0.1 segmentation of the LAr strip layer, which provides the first sampling in the calorimeter. As shown in Figure 2.6, a single electron/photon will be concentrated in a single region of the strip layer, whereas the photons from π^0 decay may be resolved as two clusters. This provides additional background rejection compared to the layer two information. Furthermore, the strip layer will provide a very fine resolution determination of the η coordinate of any EM RoI. The impact of using the increased granularity on the standalone EM trigger rates is currently being assessed.

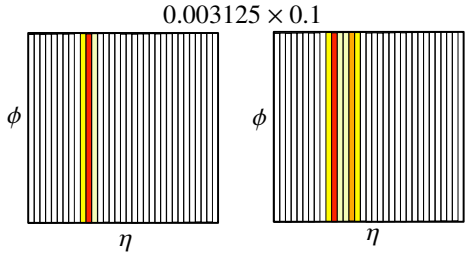


Figure 2.6: The fine granularity information from the front sampling layer of the EM calorimeter, available in the Phase-I L1Calo trigger, for a photon on the left and π^0 on the right.

The results from L1Calo would be similar to those from L0Calo, and would consist of a set of EM, Tau and Jet objects with their E_T and fine position, together with a variety of energy sums. These would be sent to the Level-1 Central Trigger, which also implements triggers based on event topology, on a single fibre per L1Calo module. Assuming the 500 kHz L0A event rate is handled by ten L1Calo modules in a single ATCA crate, the readout bandwidth is modest, and would not be more than 50 Mbit/s per module assuming that the RoI information is less than 1 kbit.

2.4 Muon Trigger

For the required acceptance for physics signatures with muons at the HL-LHC, the p_{T} threshold

of the Level-1 muon trigger must be no higher than about 20 GeV. To achieve this goal within the split Level-0/Level-1 trigger scheme for the ATLAS Phase II upgrade, upgrades of the current Level-1 muon trigger systems in both barrel and endcap regions are required. Furthermore, to improve the momentum resolution of the muon trigger RoIs and to improve the rejection of poorly measured low-momentum muons, hit information from Monitored-Drift-Tubes (MDTs) is used in the Level-0 or Level-1 trigger decision for the first time. Some aspects of the upgrade of the Level-1 muon trigger are introduced below; a more complete discussion is given in Section 5.

The barrel muon trigger system for the Phase II ATLAS upgrade will make use of the current RPC trigger detectors. However, the electronics will need to be replaced in order to adapt to the new Level-0/Level-1 trigger scheme and to cope with the increased rates and increased latency, which exceed the limits of the current trigger electronics. The majority of the Level-0 trigger logic will be performed in off-detector boards. The on-detector trigger electronics will be primarily concerned with the digitisation of the RPC front-end signals and implementing FPGA-based zero-suppression algorithms, in order to reduce the data bandwidth to the off-detector boards. The replacement of the readout electronics will give the opportunity to improve the tracking performance of the present RPC chambers using time-over-threshold information as a measure of the charge deposited on the strip. The centroid of the charge distribution inside a cluster of adjacent strips will yield an improved measurement of the track position, see section 5.2.1.

After the Phase I upgrade, the muon endcap trigger will be based on hits in the TGC big-wheel (BW) and vector information from new small-wheel (NSW) that provides track-segment information with an angular resolution of 1 mrad and spatial resolution of a few millimetres. The angular information means that track segments which are not consistent with originating from the beam interaction point (IP) can be rejected. The matching in space between TGC-BW RoI and the position of NSW track-segment can also be used to provide further background rejection. Thus, the combination of the BW and the NSW removes most of the RoIs originating from charged-particle tracks which do not come from the IP. This Phase I system forms the basis for the Level-0 muon endcap trigger in the Phase II upgrade. In order to adapt to a new ATLAS trigger scheme, all electronics boards will be replaced with the exception of the Amp-Shaper-Discriminator (ASD) boards and associated cables that are mounted on the chambers. The implementation is described in Section 5.

2.4.1 Level-0/Level-1 MDT Trigger

The Level-1 muon trigger rate can be reduced by sharpening up the momentum resolution of muon trigger RoIs, such that the background from lower momentum tracks can be suppressed. In the Phase II upgrade, the monitored-drift-tubes (MDTs) will be used in the Level-0/Level-1 trigger for the first time. The tightening of the threshold curve results in a reduction of the muon rate by a factor of three over much of the detector.

In the front-end boards, the new MDT trigger will implement BCID on a dedicated circuit, shown in the left plot of Figure 2.7, where the leading edges of the hit signals from the ASD's are detected. The BCID hit signals will then be serialised and transmitted to USA15 using high-speed optical links, where they are subsequently deserialized and fed to decoding circuits, as shown on the right of Figure 2.7.

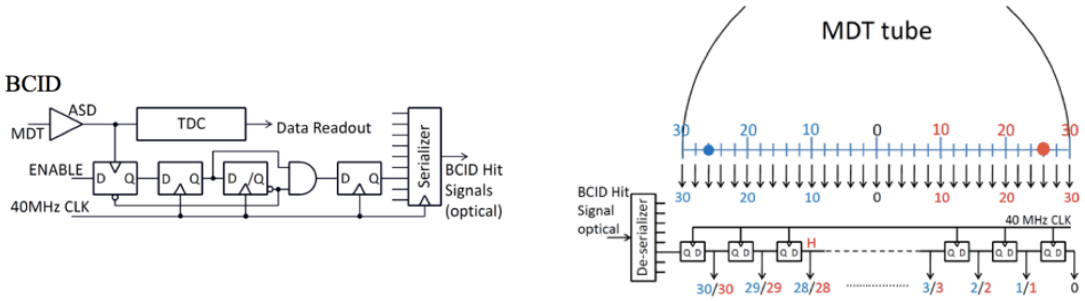


Figure 2.7: The level-0 muon BCID circuits and 31-stage shift register for the Phase II MDT trigger.

The 25 ns bunch-spacing corresponds to a 0.5 mm drift distance in the MDTs. Hence a hit in the 30 mm diameter drift-tubes corresponds to a series of possible incident positions/BCIDs. In the MDTs, this is encoded using a a 31-stage shift-register, where a single BCID hit signal is converted to 61 possible space-time combinations [30, 29,..., 1, 0, 1,..., 29, 30], with the two same numbers corresponding to the two possible drift directions. In the electronics this is achieved by moving a hit signal from one register to the next, from the ends to the centre of the shift register, at the LHC 40 MHz clock speed. The decoded hit signals from each layer can then be aligned in accordance with the expected incidence angle from an infinite p_{T} track. Aligned hit signals from the inner, middle and outer stations can then be fed into the track fitting/finding FPGAs and the transverse momentum of the MDT muons can be estimated with a similar quality to the current level-2 muon stand-alone algorithm (MuFast).

2.4.2 Improvement of the RPC trigger

An alternative way to reduce the Level-1 muon trigger rate, by sharpening up the high-p_T resolution of muon trigger RoIs, would be to substantially increase the present RPC system tracking capability. Indeed, the accuracy of the position measurement in the bending direction of the RPC chambers could be improved by using charge distribution in a cluster of adjacent η-strips. The charge measurement can be achieved by exploiting the property of the present RPC front-end electronics that the duration of the output signals of the Amplifier-Shaper-Discriminator chip is correlated with the input charge. Measuring the duration of the signal with a TDC would thus deliver the charge distribution in a cluster of adjacent strips. The centroid of the charge distribution could be used as an estimator of the particle position. A R&D project has been started to evaluate the potential of this method in detail.

2.5 Level-1 Track Trigger

The proposed upgrades of L1Calo and L1Muon triggers alone are unlikely to be sufficient to meet the ATLAS physics goals at the HL-LHC. In particular, the L1 electron trigger rates would be unsustainable for thresholds of ~ 20 GeV based solely on calorimeter information. Furthermore, the limited resolution of the L1Muon system makes it hard to achieve sizeable rate reductions with modest increase of the p_{T} thresholds.

The only other source of information with the potential to enhance the purity of the events selected at Level-1 is the Inner Detector. Tracking information is already an essential ingredient

in the current Level-2 trigger; where matching tracks found in the Inner Detector with objects in the calorimeters or the muon detectors provides much of the required background rejection. A new hardware trigger, the FTK, that is based on a fast track reconstruction system will provide reconstructed tracks to the High-Level Trigger, is planned as part of the Phase-I Upgrade [2]. This will be used as the basis of a Level-1 track trigger that can be implemented as part of the tracker upgrade.

The impact of a track trigger on the Phase-II Level-1 trigger rates has been estimated both using simulation studies and extrapolations of the current Level-1 trigger rates from data. These studies indicate that the inclusion of tracking information in the Phase-II Level-1 trigger can give a factor of five reduction in the rates for `L1_MU20` and `L1_EM18VH` single lepton triggers, relative to the rates at Level-0.

Simulation studies of the breakdown of the muons for $|\eta| < 1.3$ in terms of their true momentum firing the `L1_MU20` trigger have shown that approximately 80% of the trigger muons have $p_{\mathrm{T}} < 17\,\mathrm{GeV}$. This reflects the approximately exponential transverse momentum distribution of muons produced in proton-proton collisions at the LHC and the limited p_T resolution of the muon L1Muon trigger. Similar performance is expected in the forward region after the installation of the new small wheel in Phase-I. Consequently, if the momentum of the muon RoI can be determined from matching to a L1 track object, then the trigger rate will be reduced by around a factor of five. Figure 2.8 shows the corresponding efficiency curve, showing that using L1Track can sharpen the p_{T} turn-on threshold, even with a p_{T} resolution that is five times worse than the offline track reconstruction. Fast track reconstruction using a similar approach has been studied up to luminosities of $3 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in the context of the ATLAS FTK system [3]. The only additional constraint for the L1Track system is the shorter latency required for processing ($\sim 6\,\mu\mathrm{s}$). This will require a higher level of parallelism in the L1Track system relative to FTK, but this is considered to be achievable.

The track trigger can bring similar improvements to EM triggers by providing additional discrimination against the relatively large background rate from electromagnetic energy deposits due to charged pions and photons from π^0 decays. For example, Figure 2.9 shows the impact of requiring that the RoIs from the `L1_EM18` Level-0 trigger have a matched track which satisfies $0.5 < E_{\mathrm{T}}/p_{\mathrm{T}} < 2.0$; the trigger rate with track-matched RoIs is ten times smaller than without track-matching. The rate reduction factor is even higher when Level-2 quality EM clusters are used for the track matching, indicating that the benefits from L1Track are complementary to any rate reduction that can be achieved by improving the L1Calo system. Finally, the right plot of Figure 2.9 shows the impact of the track association for RoIs passing the current `L1_TAU` trigger. For a constant 20 kHz trigger budget, the inclusion of a track trigger based on requiring 1 or 2 tracks with p_T greater than 2 GeV will reduce the E_T threshold required from 76 GeV, corresponding to an efficiency with respect to offline tau candidates with ET>20 GeV of only ~ 0.08, to 39 GeV providing an efficiency with respect to offline reconstructed taus of ~ 0.3.

2.5.1 RoI-driven L1Track

The baseline design for the L1 track trigger uses an RoI based approach. The RoI-driven L1Track trigger would provide *all* tracks in a number of relatively small regions of the inner tracking detector, where the regions considered are defined by the EM and muon RoIs from the Level-0 trigger,

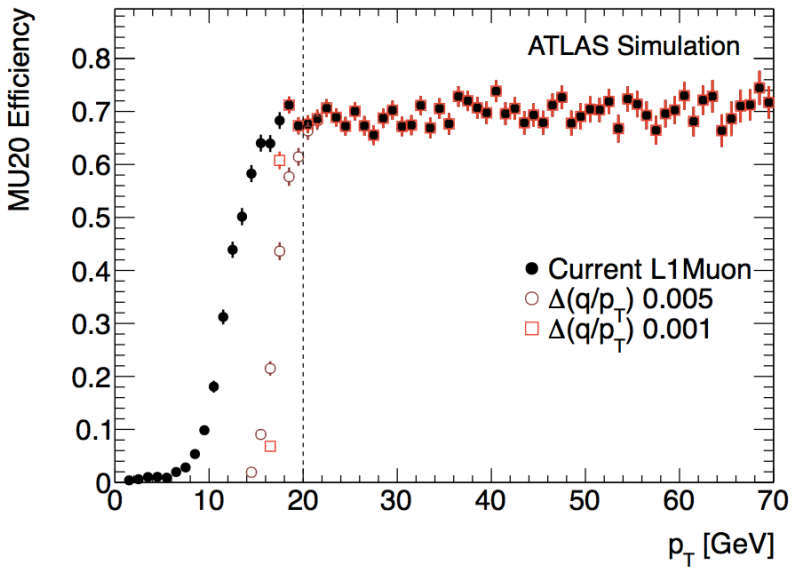


Figure 2.8: The `L1_MU20` trigger efficiencies as a function of true p_{T} after matching with a true muon assuming a track trigger with different p_{T} resolutions, assuming that all tracks are reconstructed.

and can only be implemented in a split Level-0/Level-1 trigger architecture. The alternative self-seeded design that aims to reconstruct high-p_{T} tracks in the *entire* tracker is also being investigated.

The full data from the Inner Detector cannot be read out at the full beam-crossing rate or even

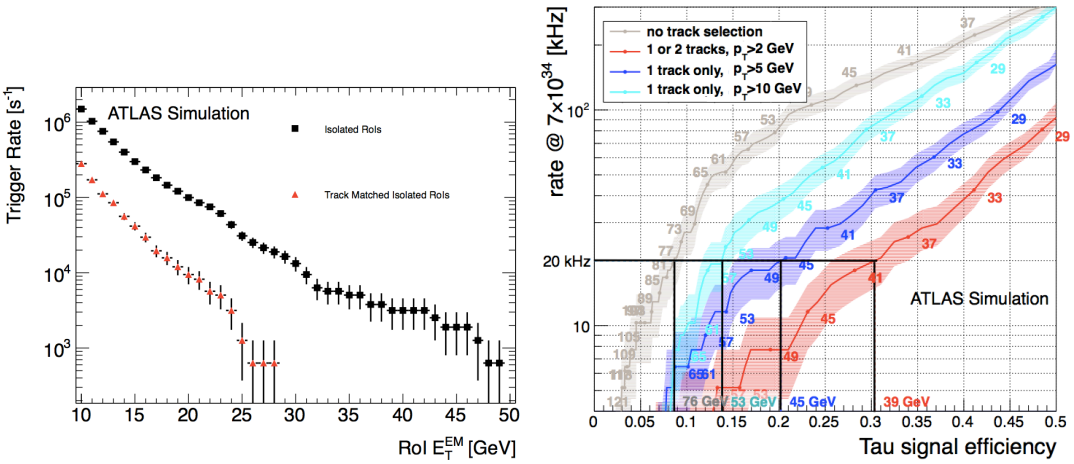


Figure 2.9: Left: the trigger rate vs. L1 EM Cluster E_{T} threshold for simulated minimum bias events with $\langle\mu\rangle = 70$. Right: rate vs. tau finding efficiency curves for taus from the decay of a 120 GeV Higgs boson for the inclusive tau trigger at 7×10^{34} cm^{-2} s^{-1} for different track multiplicity and minimum track p_T requirements. The bands show the rate vs efficiency parametrised for different L1 cluster E_T thresholds, shown as the small numbers next to the corresponding points on each band. The thresholds for each band, such that the integrated rate from the trigger is 20 kHz, are shown at the bottom of the plot. The rates are estimated using simulated minimum bias events at 3×10^{34} cm^{-2} s^{-1} and extrapolated to 7×10^{34} cm^{-2} s^{-1}.

at the Level-0-Accept rate of approximately 500 kHz. The regional readout design of the track trigger addresses this problem by starting from the Level-0 trigger RoIs and then reading only a subset of the tracker data. Conceptually, this is similar to the current design of the ATLAS Level-2 trigger, except that between the Level-0-Accept and the Level-1 decision the data are buffered on-detector in the inner detector readout electronics, as shown in Figure 2.2. After a Level-0 trigger, the data in the front-end chips are copied to a second buffer and labelled with a Level-0 identifier. Some time later, an RoI map device sends a Regional Readout Request (R3) to selected detector front-ends, using information from the Level-0 trigger to identify the modules whose data should be read out. This subset of front-end data is sent to the associated RODs and multiplexed into the main readout data stream. The ROD separates the R3 data from the main readout data and sends them to L1Track. The tracking information from the L1Track trigger would be combined with the output of the L1Calo and L1Muon systems in the central trigger processor. One of the main advantages of the RoI-driven L1Track design is that it has little impact on the inner detector layout which, apart from the depth of the front-end buffers, can be optimised almost independently from the design of the trigger. The double buffering and digital logic for the regional readout has been implemented in the design of the front-end readout ASICs for the strip tracker upgrade and the overall functionality will be tested with prototype chips when they are available in 2013.

The main challenge for the RoI-driven L1Track trigger is to fit within the ATLAS Phase II trigger latency constraints of 20 μs. Preliminary results from latency studies using a discrete event simulation show that more than 95% of the regional data can be read out within 6 μs, with appropriate prioritisation on the hybrid and assuming 160MHz readout lines, which increases to close to 100% if the speed of the readout lines was increased to 320MHz. This would imply that the latency available for pattern recognition is of similar magnitude, which is considered to be feasible in ten years based on the expected evolution of the hardware that will be used for the FTK project [4].

2.5.2 Self-seeded L1Track

The self-seeded L1Track design performs a fast track reconstruction of *all* high momentum tracks ($p_T > 10$ GeV) in the full coverage of the tracker without requiring external seeds. The huge amount of data makes it impossible to process all hit information from the inner detector for every bunch crossing with currently available technology, so only a subset of selected strip layers would be used for trigger processing. The strategy is to reduce the number of trigger layers to a minimum, while retaining sufficient redundancy in the system and minimising the fake track rate.

For the self-seeded track trigger, a data reduction scheme has to be applied in the detector front-end system in order to remove hits originating from low-momentum tracks at the earliest possible stage. For this, two methods can be exploited: using hits with a small cluster size and inferring the transverse momentum of charged particles from the lateral inclination angle between the hits of stacked double strip layers. In simulations, it is found that these two methods can reduce the bandwidth by about two orders of magnitude for $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ [5, 6]. The exact performance depends on the detailed design of the layout of the inner detector layers. The design requirements for self-seeded track trigger have not been implemented in the current design of the upgraded tracker.

The positions of the hit coincidences in the stacked strip layers are sent to track-finder units which perform a global hit matching in the L1Track processor boards. The expected data rate at

the HL-LHC is about 4 Gbps per stave. This is believed to be sufficiently low that high-momentum tracks can be obtained with a small fake rate. There are a number of open technical questions with the self-seeded L1Track concept, including the implementation of the front-end filtering algorithms, the front-end read out of the trigger information and the technology choice for the fast pattern look-up in the L1Track processor boards.

2.5.3 L1Track pattern recognition and FTK upgrade for Phase-II

The associative memory (AM) technology used in the FTK for the pattern recognition would also be a prime candidate for the L1Track pattern recognition strategy. The very high level of pile-up and the larger number of silicon detector layers, compared to the current ATLAS, would inevitably require a much larger number of patterns to be stored in the AM chips in order to achieve high efficiency for real tracks and low level of fakes. Therefore, the effort to develop new generations of the AM chip, using novel ASIC design technologies (e.g 3-D) is essential for the ATLAS Phase-II programme, regardless of whether such a chip will be used for an upgraded FTK or for the L1Track pattern recognition. Studies will also be required to optimise the pattern recognition hardware, including the data formatting and data preparation, and the potential use of hit filtering, profiting from the few millimetres spacing between the double layers of silicon wafers in the strip detectors, to reject early hits from low momentum tracks.

2.6 Central Trigger System

The main advantage of the proposed L0/L1 architecture is the availability of additional inputs to the trigger decision, which would otherwise not be made available at the bandwidth and latency required for the first trigger stage. The split L0/L1 design is shown in Figure 2.2, which shows the elements of the central trigger system at L0 and L1: the central trigger logic, the Trigger, Timing and Control (TTC). While the block diagram clearly shows the different roles the L0/L1 Central Trigger (CT) play in the system, the physical implementation could well be an integrated design, for example electronic boards housed in the same crate. If possible, the trigger functionality will be carried out using generic boards, which share the same interfaces and design, and which are dedicated by means of specialised firmware to the functionality of the L0 or L1 trigger respectively. The following two sections outline the proposed design of these two stages, while another section is concerned with the Trigger, Timing, and Control (TTC) system, which is used to distribute the Trigger and Timing signals within ATLAS, and between ATLAS and the accelerator. The TTC system is also expected to be redesigned for the needs of the Phase-II upgraded detectors and machine. A block diagram of the split L0/L1 design is shown in Figure 2.2.

As mentioned below the subsystems will have a choice whether or not to participate in L0, hence a whole range of functionalities will be made available for both the L0 and L1 CTP. Among these are the flexible generation of trigger items from the available inputs, pre-scaling of trigger items, bunch group masking, dead time generation and trigger veto handling, as well as the online monitoring of the trigger rates, dead-time, and busy fractions. In addition the trigger on both levels will be partitionable, with one main partition feeding the data acquisition for the physics running, and up to three further partitions for testing and commissioning purposes. that derive their trigger signals from a subset of the trigger menu used for the main partition, but have separate veto and dead time handling.

The L0 Central Trigger (L0CT) combines the functionality of the L0CTP and the Topological Processor. It resembles the Phase-I system in that it receives its main inputs from the L1Calo and L1Muon Trigger processors plus additional trigger inputs, for example from forward and luminosity detectors, and a trigger derived from the signals recorded in the MDT chambers if it is implemented at L0. However, a new system will be required for the Phase-II upgrade to ensure that the Phase-I system does not impose constraints on the Phase-II system. The L0CT has to provide trigger accept signals at rates of at least 500 kHz, while fitting into the overall L0 latency of 6 μs. The L0 muon and calorimeter trigger systems will identify regions of interest, that are used in a topological trigger system in order to perform trigger algorithms based on correlations between different types of trigger objects, such as muons and jets. The regions of interest identified at L0 may be used to seed the inner detector and muon trigger at L1. The current candidate electronics standard for the implementation of the L0CT is ATCA, using high-speed optical links for both incoming trigger signals and outgoing trigger and timing signals. A limited number of electrical inputs can be made available for specific use cases and system tests. Trigger signals will be delivered to all sub-systems by the TTC as discussed below, however subsystems can decide if they want to readout on receiving a L0 accept or not.

The L1CT in many ways resembles the L0CT in that it forms a logical combination of both the results of the L0CT and additional detector inputs, namely the L1 Track Trigger, the L1Calo Trigger, and the L1Muon Trigger derived from the signals recorded in the MDT chambers if it is implemented at L1. Again the trigger objects will be made available to a topological trigger processor, which in turn will provide the output of the topological trigger algorithms to the CTP of L1CT. For the L1 trigger the required accept signal frequency is at least 200 kHz, while the overall latency of this trigger stage is 20 μs. Given the desire to integrate the L0CT and L1CT systems if possible, it is clear that the candidate technologies are the same, namely ATCA with high speed optical links for the signal input and output interfaces. The trigger signals from the L1CT will be distributed to all subsystems. Here all sub-systems need to participate in order to trigger the readout of detector data for the following stages of trigger and data acquisition.

For the Phase-II upgrade it is foreseen to introduce upgraded systems for both the data transfer between the detector front-ends and back-ends, and for the distribution of trigger and timing signals. One of the requirements in introducing these new systems is to maintain backward compatibility with the legacy TTC system, as subsystems will still be using the TTCrx receiver components after the upgrade.

For the communication between the detector front-ends and back-ends a new bidirectional standard is foreseen which will carry trigger and timing signals, control signals, as well as readout data. This link system consists of on-detector and off-detector GBT elements connected by a Versatile Link.

For the upgrade of the off-detector trigger, timing and control system, the basic requirements are bi-directionality and increased bandwidth with respect to the current system, while maintaining backward compatibility. In addition flexible partitioning should be possible, and the new system will accommodate the requirements of different clients (accelerator and experiments) in order to allow the use of a common system. The current optical network implements a point to multipoint architecture with relatively high bandwidth downstream and lower bandwidth upstream. This is the typical topology of a Passive Optical Network (PON), with extra requirements on clock recovery

and fixed latency signals added. The current demonstrator model for the new PON system using 1G-EPON components has a downstream bandwidth of 1.6 Gb/s or ten times the current TTC system for about twice the transmission latency. The upstream bandwidth is 800 Mb/s, where the actual payload and latency depends on the splitting ratio of the system. The current plan is to use 10G components for the final system, which should dramatically improve the values for the demonstrator given above.

2.7 Data Acquistion

As described in this document, changes are required to the ATLAS detector to meet the challenges of operating at a peak luminosity of $7 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The increased Level-1 accept rate and expected event size (> 4 Mbyte) alone indicate that at least a factor of four increase in bandwidth will be required compared to the data acquisition system that will be in operation up to the Phase-II shutdown. Changes to the detector readout architecture to cope with these requirements are described in Section 2.7.1. Other functions of the data acquisition will also have to evolve to meet the challenges imposed by increased: level-1 accept rate, the rate and quantity of data that will have to be transported from the detector readout to the HLT, event building and throughput to the mass storage. The rate and bandwidth requirements will pose particular challenges to the data acquisition networks and may necessitate a change of architecture and network technology to meet the extended performance requirements. As described in Section 2.8, a $\langle\mu\rangle$ value of 200 implies a significant increase in the average event processing times, even taking in to account the increase in compute power per node that technology evolution can expect to provide. Subsequently an increase in HLT computing power, and its associated infrastructure, will be required beyond that which will be in operation up to the Phase-II shutdown.

Experience has indicated that evolutions in hardware technologies necessitate equivalent upgrade of software technologies used in a data acquisition so as to fully exploit the full potential of the underlying hardware technology. For these reasons following a period of targeted technical evaluations, particularly in the areas of network technologies, online databases, information sharing and expert systems, new open-software technologies will be selected and where necessary the data acquisition software re-designed and implemented.

The upgrade of the data acquisition system will leverage the changes that will have been implemented for Phase-I operations and where it is necessary to meet the additional Phase-II performance requirements, leverage technology advances expected in the domain of commodity computing and networking. However, given the long lead time to Phase-II it is premature to present details of a possible data acquisition architecture that will be so dependent on technology evolution.

2.7.1 Readout System

The ATLAS dataflow architecture is schematically depicted in Figure 2.10. The readout system is the "core" of the dataflow and interfaces the detector-specific front-end to the DAQ/HLT units which build and filter the events before the data are transferred to mass storage. In general terms, the following functional blocks can be identified:

- Detector front-end (FE) electronics: conditioning and digitization of the analog input signals, implementation of the Level-1 pipelines and derandomizing buffers, transmission of the FE

data stream on a Level-1 Accept (L1A) trigger.

- Readout Drivers (RODs): multiplexing of several FE data streams, optional processing (e.g. calibration and signal feature extraction), formatting of the FE streams into output ROD fragments.
- Readout Links (ROLs): transmission of the ROD fragments at the L1A event rate. Currently based on the S-LINK protocol, they provide XON/XOFF flow control and error detection.
- Readout Buffers (ROBs): temporary storage of the ROD fragments during the LVL2 decision making and event building, buffer management, error detection and recovery, diagnostics and on-line monitoring.
- Readout Systems (ROSs): housing and control of the hardware modules (ROBIns) that implement the ROBs, buffering and multiplexing of several ROB's into a single event builder input, and servicing Region of Interest (RoI) requests to LVL2.

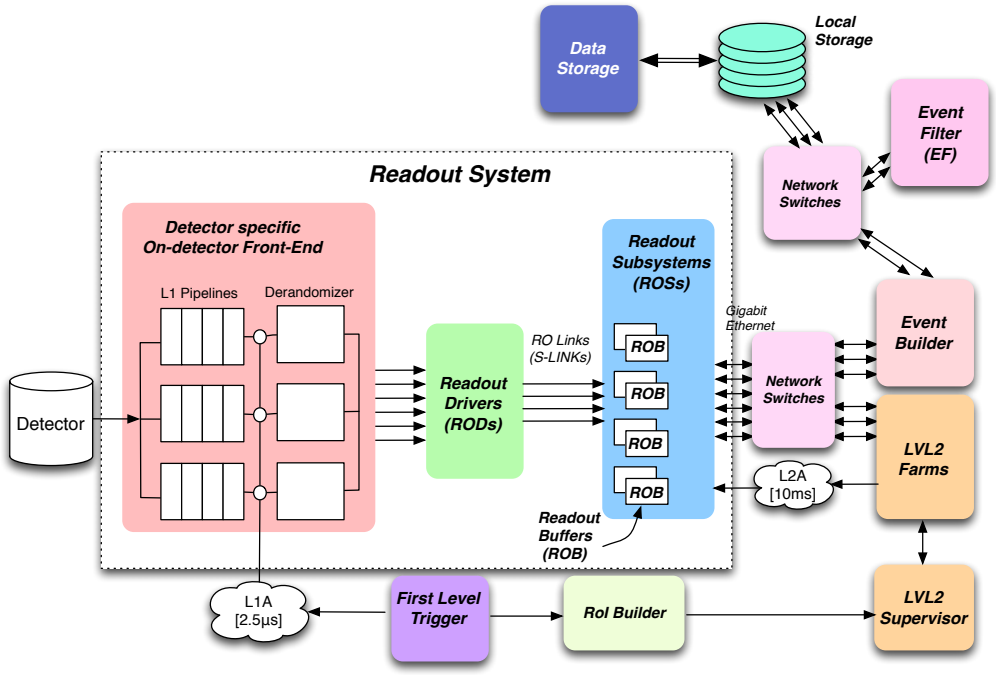


Figure 2.10: System overview of the ATLAS dataflow and readout system

The flexible design and implementation of today's readout system has enabled the readout system to be consistently operated beyond its design specification, reflecting the evolving needs of ATLAS during the first years of operations. However, the performance of this implementation will become incompatible with the conditions at the HL-LHC because of increased Level-1 trigger rates, new latency specifications, and significantly more challenging data bandwidth requirements: e.g. the maximum ROL bandwidth today is 160 MB/s is based on a ROD fragment size of 1.6 kB for a L1A trigger rate of 100 kHz, while at HL-LHC the event-size will increase significantly because of pile-up and an L1A trigger rate of at least to 200 kHz. Furthermore, component ageing will also become

a concern: most of the electronics will be at least 15 years old, raising issues of maintenance and reliability on the timescale of the Phase-II operations.

For the upgrades, total production costs and ease of maintenance during the HL-LHC operation should be one of the prioritized design criteria. The experience accumulated over several years of operations, emerging technologies and the evolution of commercially available components should allow more commonality across the detector ROD systems in terms of functionalities and implementation.

A readout scheme that addresses the need for increased performance, e.g. throughput, at the HL-LHC and places emphasis on commonality is being evaluated by all the ATLAS detector systems, see Fig. 2.11. In this scheme the detector specific FE electronics interface to a common GBT aggregator that implements the multiplexing of FE data streams, data fan-out (e.g. duplication of selected data streams) and translation to standard commercial network protocols. Subsequently the aggregated FE data streams via a commercial high-speed network are routed to the RODs. This allows a common interface for data input to the RODs based on commercial high-speed links.

This scheme, which introduces commercially available high-speed switching networks and links at the earliest stages of the readout and possibly the implementation of a common ROD, introduces several advantages: the capability of re-organizing the readout connectivity without physical re-cabling, scalability and staging capabilities, large scale production of a single ROD design thus lower production and maintenance costs, easier long-term maintainability by avoiding multiple implementations of common functionality (e.g. ROD data formatting), faster and more effective integration and commissioning, and more effective usage of manpower and resources.

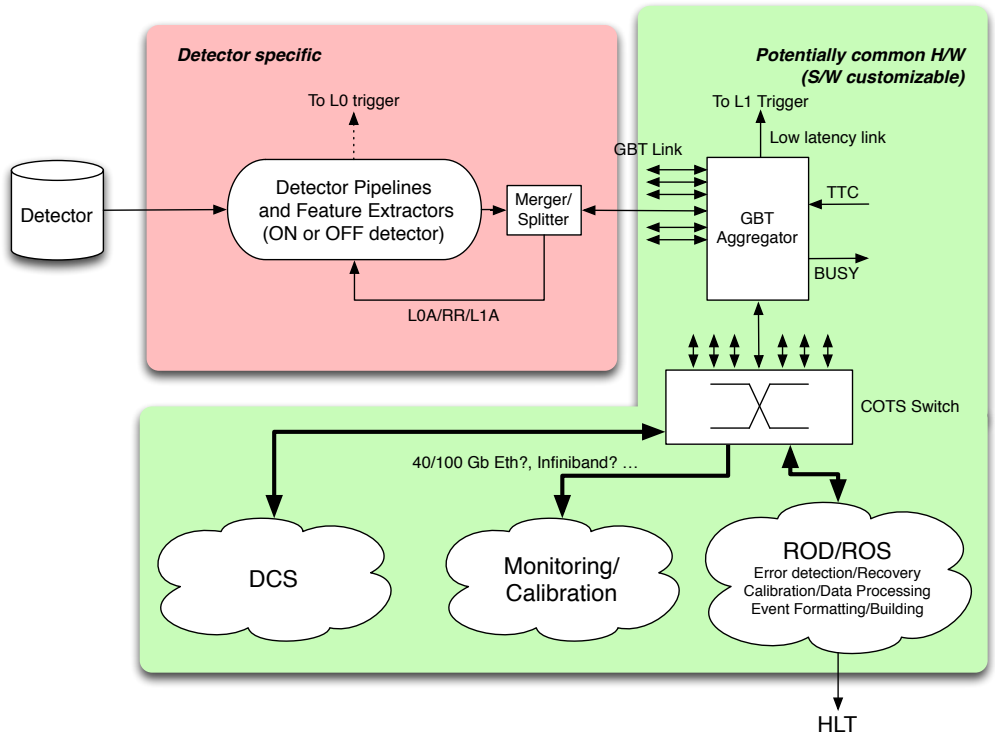


Figure 2.11: Readout architecture overview for the Phase-II upgrades

From a technical perspective this scheme also offers the possibility of increased integration between the ROD and ROB/ROS components. The ROD/ROS physical integration would relieve the technical complications of implementing a reliable flow control mechanism (XON/XOFF) across a switched network. It also introduces more flexibility to cope with different trigger strategies, e.g. a more significant component of the high-level trigger selection strategy no longer seeded by RoI.

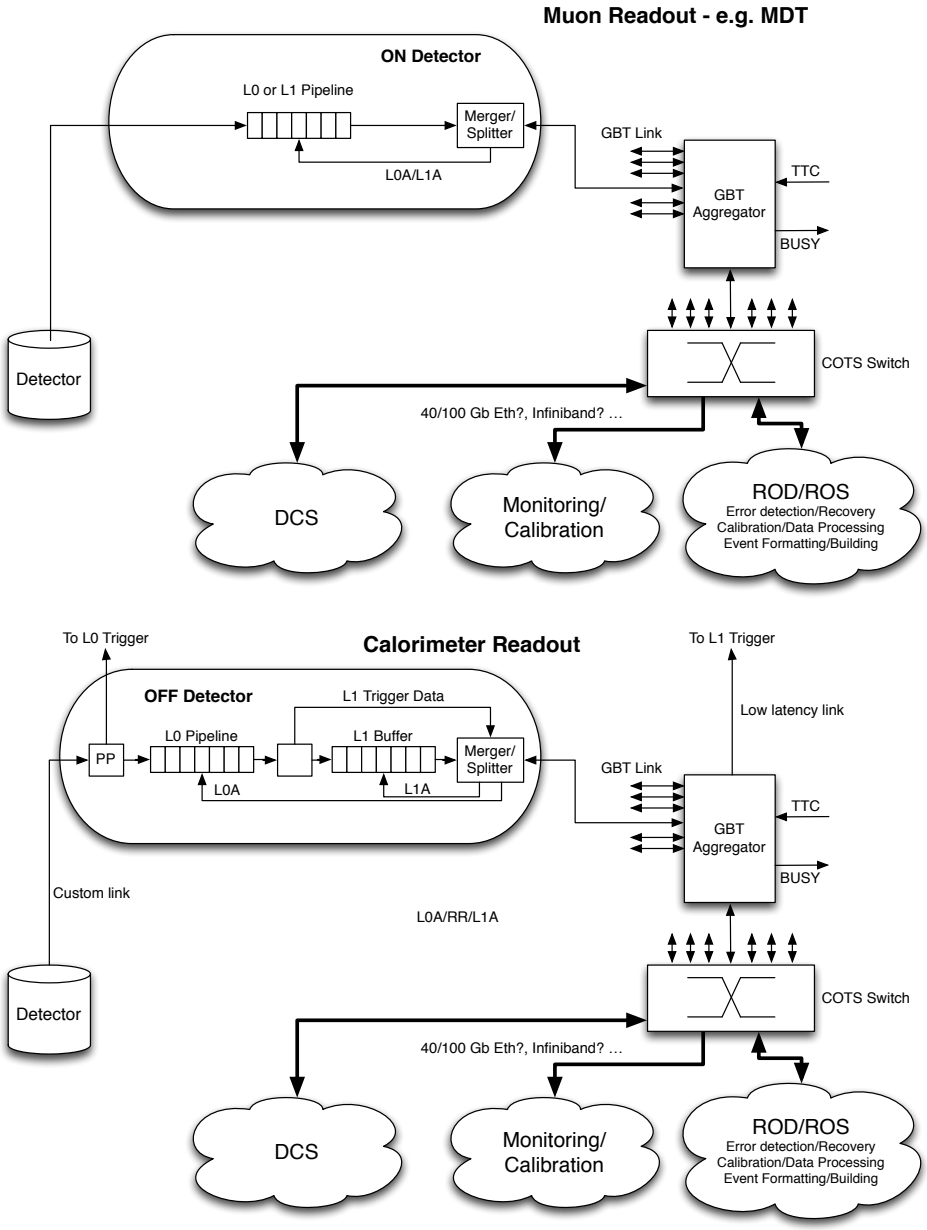


Figure 2.12: Possible implementation of the Phase-II readout architecture for the Muon spectrometer (top) and for the calorimeters (bottom).

Fig. 2.12 illustrates how the readout scheme may be implemented for a Muon detector (MDT)

and the calorimeters. For the former the front-end electronics (on detector) are interfaced directly to the GBT aggregator modules. The calorimeters stream off detector the digitized samples to a Pre-Processor (off detector) that implements the pipelines, processing of the raw data samples and preparing of the L0/L1 trigger primitives. A full analysis of the functional and technical requirements, as well as implementation details detector by detector is ongoing.

2.8 High-Level Trigger

The HLT selection software must be upgraded to match the detector upgrades, to maintain HLT rejection following the Level-1 upgrade and to fully exploit the evolution in computer hardware. Following the Phase-II Level-1 upgrade, many of the selection tools that are currently employed at L2 will be moved to L1. This includes tracking, fine granularity calorimeter information and topological triggers. An HLT rejection factor of 20-40 is required to reduce the Level-1 accept rate of 200 kHz to about 5-10 kHz for recording. To do this the HLT will employ offline-type selections which are likely to emphasise multi-object signatures, for example combinations of leptons, jets, jets with b-tagging and E_T^{miss}. In order to maintain the required HLT rejection, new selection software will be developed to fully exploit the increased computing power available through advances in computer hardware and upgrades to the HLT farm.

Selections will be based on trigger objects reconstructed with near-offline quality. In addition to the refinement of Level-1 objects within the RoIs, whole event reconstruction can be performed by the HLT. Calorimeter clustering will benefit from improved calibration compared to Level-1 and can exploit offline-type tools for suppression of calorimeter noise and pile-up. HLT tracking will be adapted to the new tracker. HLT inner detector and muon tracking will benefit from being able to use the full conditions information and magnetic field-maps, and offline-quality track reconstruction to refine Level-1 information and reconstruct tracks for which there are no Level-1 seeds.

In addition to improvements in cluster and track reconstruction, the software that identifies trigger objects such as leptons and jets will be upgraded. Lepton reconstruction will be upgraded to benefit from offline tools for electron Bremsstrahlung correction, reconstruction of converted photons and secondary vertex reconstruction for hadronic tau decays. HLT code will be upgraded to use more offline-type selections, for example multivariate selections. Event by-event primary vertex reconstruction would allow impact parameter requirements to be applied to suppress pile-up. The HLT jet reconstruction will be upgraded to benefit from advances in offline jet reconstruction algorithms and will be based on full-event information. The addition of track information could be used to reconstruct track-based jets and suppress pile-up by calculating the fraction of transverse energy associated to a given primary vertex.

Significant changes to the HLT software are planned in preparation for Phase II. These changes are needed to exploit the evolution in computing hardware. While it is difficult to predict the precise nature of the changes over the next 10 years, it is likely that the current trend towards increased numbers of cores will continue. The HLT software will need to adapt to fully exploit many-core architectures through the parallelisation of code. This will involve significant changes both to the HLT framework, to support execution of algorithms in parallel threads, and to the algorithms themselves to allow deeper parallelisation e.g. at the level of a track. Substantial effort

will be needed to design and implement these upgrades. Commonality with offline software will be exploited where possible.

3. Upgrade of the Liquid-Argon Calorimeter

The HL-LHC allows for substantial increases in instantaneous and integrated luminosities at the cost of corresponding increases in radiation damage and operational challenges. During HL-LHC operation, particle fluxes and the average energy deposited in the calorimeters are expected to be typically five to ten times higher than specified in the LHC design values. Furthermore, at the start of HL-LHC operations the current electronics will be 15–20 years old.

Under these conditions it is certain that the front-end electronics of the LAr Calorimeters will have to be replaced. This is due to both radiation damage and the need for ATLAS to upgrade the trigger system, with the latter requiring real-time performance capabilities that the current electronics cannot satisfy. Replacing the front-end electronics implies replacing the back-end electronics as well. The current design for the upgraded electronics is described in Section 3.1.

The conditions for the optional upgrade of the cold electronics of the Hadronic Endcap Calorimeter (HEC) and of the forward LAr calorimetry are detailed in Section 3.2.

3.1 Liquid-Argon Readout Electronics

3.1.1 Introduction

To minimize analog signal distortions, the front-end electronics of the liquid argon calorimeters are located on (or, in case of the HEC, partially in) the calorimeter cryostats. They are therefore subject to significant radiation exposure. The current system was designed to maintain performance throughout ten years of LHC operation at design luminosity, with substantial safety factors. The LAr front-end ASIC components were required to be able to sustain a total ionizing dose of 5.8×10^2 Gy, non-ionizing energy loss of $1.7 \times 10^{13} n_{\mathrm{eq}}/\mathrm{cm}^2$ (n_{eq} = 1-MeV neutron equivalent in Si) and resist single-event effects from hadrons of energy > 20 MeV for $3.2 \times 10^{12} h/\mathrm{cm}^2$ [7]. Table 3.4 summarizes the increased radiation tolerance criteria for the LAr front-end electronics at HL-LHC for a total luminosity of 3000 fb^{-1}, including safety factors as in Ref. [7], but without uncertainties on the expected integrated luminosity.

Table 3.4: Radiation tolerance criteria of the LAr electronics for operation at HL-LHC for a total luminosity of 3000 fb^{-1}, including safety factors given in brackets. For ASICs, the pre-Phase-II and Phase-II conditions are compared. For all other electronics components (HEC pre-amplifier and summing (PAS) chips, COTS, and low-voltage power supply (LVPS) electronics) the Phase-II requirements are listed.

	TID (kGy)		NIEL ($n_{\mathrm{eq}}/\mathrm{cm}^2$)		SEE (h/cm^2)	
ASICs (pre-Phase-II, 1000 fb^{-1})	0.58	(5.25)	1.7×10^{13}	(5)	3.2×10^{12}	(5)
ASICs (Phase-II)	1.74	(5.25)	5.0×10^{13}	(5)	9.6×10^{12}	(5)
ASICs PAS HEC	5.0	(5)	4.1×10^{14}	(5)	5.1×10^{13}	(5)
COTS (multiple lots)	23.4	(70)	2.0×10^{14}	(20)	3.9×10^{13}	(20)
COTS (single-lot)	5.7	(17.5)	5.0×10^{13}	(5)	9.6×10^{12}	(5)
LVPS (EM barrel and endcap)	1.35	(70)	2.3×10^{13}	(20)	6.0×10^{12}	(20)
LVPS (HEC)	0.1	(5)	6.9×10^{12}	(5)	6.9×10^{11}	(5)

As seen in Chapter 2, the expected trigger rates at HL-LHC will be exceeding the detector front-end capabilities even after the foreseen Phase-I upgrade.

The current front-end electronics use many ASICs fabricated in obsolete technologies making impossible to replace individual components and imposing a full replacement. The proposed upgrade exploits technological progress to implement a more robust architecture, in which the trigger pipelines are moved off-detector. This simplifies the front-end electronics, but requires high rate digitization and the transmission and handling of large amounts of data (approximately 140 Tbps). This also implies upgrading the back-end electronics.

3.1.2 Front-End Electronics

The LAr front-end electronics will be upgraded in steps, with the addition of new trigger boards in Phase-I, and the replacement of the primary read-out and calibration boards in Phase-II. Before Phase-II, the front-end electronics crates (FEC) will contain four types of boards, implementing four different functionalities: the front-end boards (FEB) reading out signals from 128 calorimeters cells, the LAr trigger digitizer boards (LTDB and LTDDB), the calibration boards and the control boards for the clock and configuration signals handling [8]

The main architectural difference between the existing readout system and the planned upgrade is the switch from an (analog) on-detector Level-1 pipeline to a "free-running" design in which signals from all calorimeter cells are digitized at 40 MHz and sent off-detector. This approach effectively removes all constraints imposed by the LAr readout on the trigger system, since full granularity data will be available and the latency and Level-1 bandwidth become essentially unlimited. Furthermore, due to its minimal coupling to the trigger system, the new architecture provides significant flexibility for changes/evolutions in the ATLAS trigger system or overall detector readout. It can accommodate a trigger system with a low-latency, low-granularity Level-0 trigger stage just as well as a system with a high-latency, full-detector Level-1 trigger, or both. Figure 3.1 shows the front-end architecture in Phase-II. The calibration board, which is not used during collisions, is not shown. The control board functionality may or may not be integrated directly into the new generation front-end boards.

The upgraded front-end board (FEB2)

The new front-end board architecture is designed to remove the bottlenecks inherent in the current boards by exploiting progress in technology, while maintaining the analog performance: 16-bit dynamic range (currently achieved with three gains at 12 bits each) and coherent noise below 5% of the incoherent noise level. The main functional blocks are: analog preamplification and shaping, production of (summed) analog signals for the LTDB, analog-to-digital conversion of all signals at a rate of 40 MHz, multiplexing and serialization of digital data, and transmission over high-speed optical links. Since there will be no Level-1 pipeline, the control logic will be limited to clock distribution and slow controls for configuration and calibration.

The analog preamplification and shaping stages will be integrated in a single ASIC. The Liquid Argon Preamp and Shaper (LAPAS) test-chip, fabricated in 2009 in IBM's 8WL SiGe process, validated the design approach of implementing a wide dynamic range single ended preamp followed by low power differential shaping stages with multiple gains to achieve the required 16 bit resolution. An integral non-linearity of less than 0.6% was demonstrated in both the $1\times$ and $10\times$

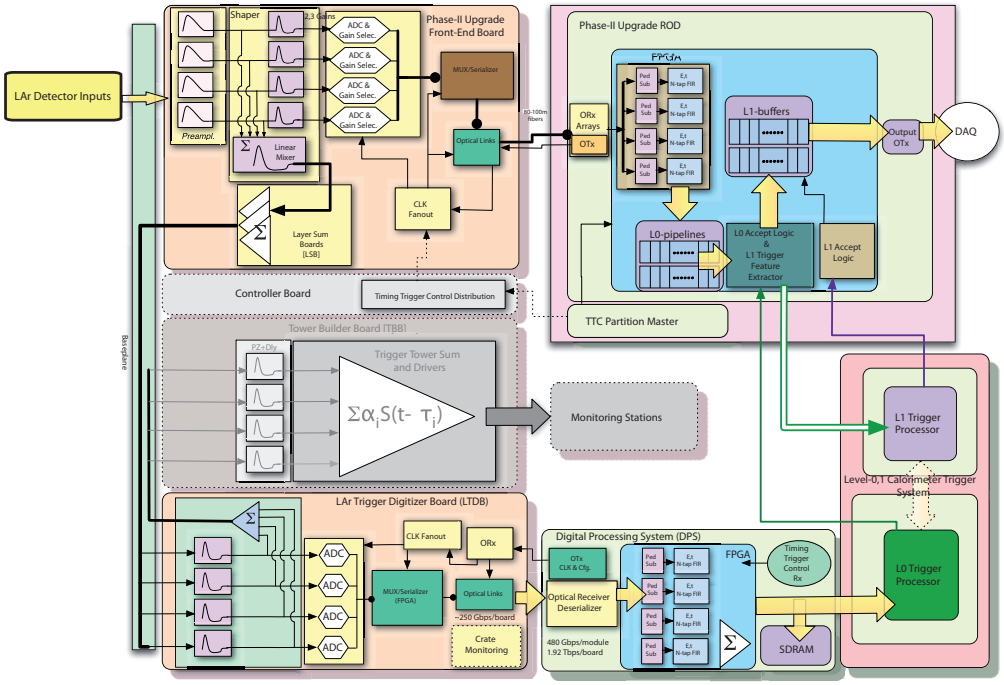


Figure 3.1: Phase-II front-end readout architecture for the liquid argon calorimeters. The calibration board, which is not used during regular data-taking, is not shown. The LTDB will have been installed in Phase-I, superseding the original Tower Builder Board.

gain shaping stages and measurements of ionizing radiation indicated robust performance for TID in excess of 1 MRad [9]. Less expensive SiGe process alternatives are currently being explored: IHP SG25H3P and IBM's 7WL.

The requirements on the ADC are a large dynamic range (at least 12 bits), low power (less than 100 mW/channel), high speed (40 MSPS), small footprint (128 channels on a 50 cm-wide board) imposing serialized outputs, and substantial radiation tolerance (see Table 3.4). Many commercial devices meet most of these criteria but are very sensitive to single event upsets (SEUs). Irradiation tests are being performed to verify if recent devices are suitable.

In addition to radiation tolerance, ASIC solutions have lower latency and power consumption. Two ASIC designs are currently being pursued: one uses four 1.5-bit pipeline stages followed by an 8-bit SAR, and the other a 12-bit SAR. Both are implemented in IBM 130 nm CMOS (8RF) technology. A test-chip for the first architecture has allowed to demonstrate analog performance at least equivalent to that of a commercial 12-bit ADC together with excellent radiation tolerance. For both architectures, multi-channel pre-prototypes with most of the required features are expected to be submitted for fabrication in 2012.

After digitization, the raw data produced on a FEB2 represents 128 (channels) $\times$ 14 bits (incl. two gain scale bits) $\times$ approximately 40 MHz, or about 72 Gbps. Conservatively assuming 25% overhead for control words and encoding, each board needs to multiplex, serialize and transmit 90 Gbps of data, leading to a total data transmission rate of 140 Tbps. A radiation-tolerant serializer in 250 nm silicon-on-saphire technology able to run at 5 Gbps has already been devel-

oped. An 8 Gbps test-chip will be submitted in 2012, and with a newer process, speeds of 10 Gbps should be achievable. It should therefore be possible to use a single 12-fiber ribbon to transmit the data produced by each FEB2.

The upgraded calibration board

Given irradiation and aging considerations, as well as changes needed in the power distribution and control systems, replacement with a new board is likely to be more cost-effective than attempting to adapt the existing boards. The main challenges for the calibration board are a 16-bit DAC with stable performance after irradiation, and a board layout that allows simultaneous pulsing of calorimeter channels. Since this board is only used during "beam off" time, SEU sensitivity is of significantly reduced concern. Compared to the current calibration board, the capability to delay individual signals will probably be dropped as this is not used in current operations.

Other boards

Since the FEB2 latency need not be different from the LTDB's, the full detector granularity and precision can be made available to an early trigger stage if needed. However, the LTDB, introduced in Phase-I, can be kept to provide signals to the Level-0 and/or Level-1 trigger stages.

At this time, no control board is foreseen for Phase-II. Both clock and control directives are expected to be sent over a dedicated optical link to each individual FEB2 as well as the calibration board. This approach, somewhat more complex off-detector, removes the need for a separate slow bus and dedicated ASIC.

Power and infrastructure

In the present design, each Front End Crate is powered by an AC/DC converter located in USA15 which provides a 280 V output. This is fed to a water-cooled 3 kW Low Voltage Power Supply (LVPS) DC/DC converter located in the Tile Calorimeter finger gaps which in turn provides 7 independent low voltages over a power bus located on the FEC. Linear regulators on the boards are used for final voltage regulation.

For the Phase-II upgrade, fewer voltages will be used in the front-end system. Given the cooling constraints, the total power is expected not to exceed the current values. The powering concept is based on distributed power architecture with main converters and point-of-load converters (POL) close to the front-end or an intermediate bus architecture with an additional set of bus voltages. Modular main converters with 1.5 kW per module are under development with conversion efficiencies above 80%. Operation in external fields up to 300 Gauss, as present at the location on the LAr detector, has been successful. POL converters need to operate in even stronger magnetic fields, which prevents the usage of magnetic materials for inductance cores. Several developments on air-core based POLs are currently ongoing and different technologies, based on Si CMOS and GaN, are under study.

3.1.3 Processing of Level-0 Trigger Signals and Compatibility With Phase-I Upgrade

In the baseline design, the low-latency Level-0 trigger signals will come from the high-granularity trigger output installed in the Phase-I upgrade. In the LAr Electromagnetic Barrel and Endcap Calorimeters, this produces "super-cells" of granularity 0.1×0.1 in $\Delta\eta \times \Delta\phi$ in the pre-sampler and third layers, and 0.025×0.1 in the first and second layers [2]. This provides a reasonable data

volume (24 Tbps) to be processed by the Level-0 trigger to achieve manageable rates. The full LAr calorimeter granularity will be available to the Level-1 trigger to further refine the trigger decision, possibly based on refined reconstruction in a region-of-interest determined by Level-0.

The system is very flexible, since the front-end electronics impose no limits on latencies or rates. It can accommodate different choices of Level-0/Level-1 architecture, including a potential future evolution in which the full calorimeter granularity is used at Level-0, or even a very long latency single-level trigger architecture. Trigger upgrades beyond Phase-II would therefore not require intervention at the front-end level.

3.1.4 Back-End Electronics

The main tasks of the back-end electronics of the LAr Calorimeters are the reception of digitized data from the FEB2 boards, data filtering and processing, as well as data transmission to the trigger and DAQ system. The expected raw data rate from each 128-channel FEB is 90 Gbps. The total input data rate to the LAr back-end amounts therefore to about 140 Tbps.

The back-end electronic boards are housed in the underground area USA15, in a non-radiation environment. In order to comply with the ATLAS requirement for a common architecture for the trigger information buffering and the data buffering and formatting, the back-end electronics will have well defined functions as described in Section 3.8.

The Pre-Processor boards will be an evolution of the FPGA-based LAr Digital Processing System, which will have been developed for the Phase-I upgrade. The data are received using serial optical links on multi-fiber ribbons. The conversion to electronic signal will be performed by commercial components and deserialization will be handled by fast FPGA transceivers.

The Pre-Processor boards will apply digital filtering using FPGAs to calculate calibrated energy deposits together with the signal time for each LAr cell as well as signal quality criteria. The digital filtering will take electronic and pile-up noise into account and will be adjustable to the expected high-luminosity conditions by parameterized FPGA algorithms. The processed data will be buffered in digital memory blocks. These can be arranged according to the number of hardware trigger levels and their corresponding trigger latencies.

The data from the individual LAr cells are further processed by the Pre-Processor boards in order to provide input to the Level-1 hardware trigger. The signal processing may go beyond the simple sums of cell energy or transverse energies, E_T: more complex algorithms like the extraction of features of electromagnetic shower shapes or fast tagging of $\pi^0 \to \gamma\gamma$ signals using the finely segmented first calorimeter layer may be possible given the availability of the full granularity of the LAr calorimeters. The detailed algorithmic layout is the subject of on-going R&D, which will take into account how many channels can be concentrated in one processing FPGA and the latency required by the trigger system.

On the Pre-Processor boards further functionalities for receiving trigger, timing and control (TTC) signals need to be integrated, in order to control the data flow and signal buffering. In addition, data monitoring functions will be implemented.

In a scenario of 14 FEB2s connected to one Pre-processor board, 32 towers of size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ are covered by one Pre-Processor board. The incoming data rates per board are estimated to be 14×90 Gbps of digitized input data. At the output, low-latency data transmission of about 250 Gbps and $5-10$ Gbps is expected to the Level-0 and Level-1 trigger systems, respectively. In

addition, data consistency and proper functioning of the data processing will be verified, exploiting the availability of pre-Level-0 data. Depending on the monitoring concept and the amount of raw data to be checked, the data rate may also well exceed 10 Gbps per board.

For the proposed mapping of 14 FEB2 to one Pre-Processor board, about 110 Pre-Processor boards are needed, and a larger number if fewer data links or lower processing capability can be concentrated in one Pre-Processor board. As with the digital processing system installed to be in Phase-I the Pre-Processor boards are foreseen to be in ATCA format with corresponding powering, cooling, and board monitoring capabilities. The Advanced Mezzanine Card concept of the ATCA standard allows for 4 Processing Units (sPU) per 9U-format ATCA board.

With 10 Pre-Processor boards per ATCA shelf, a total of 11 ATCA shelves will be needed, to be installed in 6 USA-15 racks, to be compared with the 17 VME crates currently installed.

Upon receipt of a Level-1 accept signal, the calibrated data will be presented to GBT aggregators and redirected to ATLAS standard ROD/ROB/ROS systems through a switched network as proposed in Section 3.8. Calibration data analysis, histogramming, monitoring and Level-2 trigger algorithms can be implemented at that stage.

3.2 Liquid-Argon Calorimeter Detectors and HEC Cold Electronics

3.2.1 Detector Upgrade Motivation and Options

The LAr detector system consists of a barrel region and two end-cap and forward regions. The radiation levels increase with $|\eta|$. From the barrel to the end-cap and from the end-cap to the forward calorimeters the flux and average energy of the particles from minimum bias events increases with the consequent growth of multiplicity and density of particle showers. This results in a power density, and hence radiation flux, deposited in the calorimeters reaching levels not seen in previous collider detectors. The ATLAS calorimeters are designed to cope with a peak luminosity of 10^{34} cm^{-2} s^{-1} and an integrated luminosity of 1000 fb^{-1} as foreseen at the LHC.

Under HL-LHC conditions both the maximum instantaneous luminosity of 7×10^{34} cm^{-2} s^{-1} and the maximum integrated luminosity of 3000 fb^{-1} collected over an anticipated HL-LHC lifetime of 10 years correspond to an increase over LHC parameters of typically an order of magnitude. It is safe to assume that the LAr barrel electromagnetic calorimeter will continue to function at highest peak luminosities and over the expected integrated luminosity. Depending on the running conditions and the actual level of radiation, the performance of the forward calorimeters may be degraded due to the increased peak luminosity. On the other hand, it is unclear how much the end-cap calorimeters (EMEC and HEC) will be affected. To understand the effects of luminosity on the end-cap and forward regions, the so-called HiLum R&D project has been launched.

One element which might be affected by the integrated luminosity are the analog front-end electronics of the HEC which are installed at the perimeter of the HEC in relatively high radiation fields, inside the endcap cryostats. Measurements in radiation tests [10] show that the amplifiers start to degrade when the neutron fluence exceeds $3\times10^{14}n_{\rm eq}/{\rm cm}^2$. Compared to ATLAS requirements of $0.28\times10^{14}n_{\rm eq}/{\rm cm}^2$ this amounts to a safety margin of ~ 10 for a total integrated luminosity of 1000 fb^{-1} during LHC operation. Assuming an additional integrated luminosity of 2000 fb^{-1} at the HL-LHC the safety margin is strongly diminished, i.e. the present HEC cold electronics would be operated at their limit. R&D has therefore started towards a new ASIC that would

be more radiation hard to safely withstand the total luminosity integrated during HL-LHC operation. Such a new ASIC should be available to replace the present GaAs chips at the HL-LHC. The plans are outlined in the HEC electronics Sec. 3.2.2.

In the forward region a possible option is to design a new FCal detector with a reduced gap size, a modified HV distribution circuit, and the addition of cooling loops. This upgrade version, named the sFCal, will be described in Sec. 3.2.3. An alternate upgrade option to deal with the limitations of the FCal, is to install a small calorimeter just in front of the current FCal. This detector, called the "Mini-FCal", will be discussed in Sec. 3.2.4.

The data collected in Phase-0 as well as physics simulation with degraded HEC cold electronics and with degraded FCal detectors will show whether the existing detectors may be used for HL-LHC (option 0) or whether ATLAS will have to apply one of the following three options (1-3):

- Option 0: No change neither of the HEC cold electronics nor of the FCal detectors.
- Option 1: If the HEC cold electronics have to be replaced (Sec. 3.2.2), the large cold cryostat cover would have to be opened and the irradiated FCal would have to be removed. A newly built cold FCal (sFCal) (Sec. 3.2.3) would then be inserted before closing the cryostat.
- Option 2: If the HEC cold electronics do not have to be replaced, the cold FCal would be replaced by a new one of the sFCal type (Sec. 3.2.3). It is anticipated that only the small cover of the cold vessel, the FCal bulkhead, would have to be removed.
- Option 3: If the HEC cold electronics do not have to be replaced, the cold FCal would stay in place and a new small calorimeter (Mini-FCal) (Sec. 3.2.4) would be placed in front of it. In this case only the cryostat warm vessel would have to be opened.

Each of the three upgrade options has specific requirements, which are discussed in the relevant detector sections.

3.2.2 Cold Electronics of the Hadronic End-Cap Calorimeter

In the vicinity of the HEC cold electronics a maximum neutron fluence of $0.82 \times 10^{14} n_{\mathrm{eq}}/\mathrm{cm}^2$, γ dose of 1.2 kGy and hadron fluence of $1.0 \times 10^{13}\,\mathrm{h}/\mathrm{cm}^2$ are expected after a total luminosity of 3000 fb^{-1}. For the radiation tolerance criteria, these numbers are multiplied by a safety factor of five to account for uncertainties in the simulation. The radiation hardness against all three types of radiation has been studied with both pre-production and production versions of the current HEC pre-amplifier and summing (PAS) chips both at room temperature and submerged in liquid nitrogen [10]. It was found that neutron irradiation has by far the largest impact on the functioning of the electronic devices. Measurements in these tests showed that the current HEC PAS amplifiers start to degrade rapidly when the neutron fluence exceeds $3 \times 10^{14} n_{\mathrm{eq}}/\mathrm{cm}^2$ (Si-NIEL equivalent for ATLAS).

3.2.2.1 Present HEC cold electronics The signal processing of the HEC employs the notion of 'active pads' which keeps the detector capacities at the input of the amplifiers small and thereby achieves a fast rise time of the signal [11, 12]. Short coaxial cables are used to send the signals from the read-out pads to preamplifier and summing boards (PSB) located at the perimeter of the

wheels inside the liquid argon. The detector capacitance varies from 40 to 400 pF giving a rise time variation from 5 to 25 ns.

These PSB's carry highly integrated amplifier and summing chips in Gallium-Arsenide (GaAs) technology. The signals from a set of preamplifiers from longitudinally aligned pads (2, 4, 8, or 16 for different regions of the calorimeter) are then actively summed forming one output signal, which is transmitted to the cryostat feed-through. The GaAs TriQuint QED-A 1 μm technology has been selected for the front-end ASIC because it offers excellent high frequency performance, stable operation at cryogenic temperatures and radiation hardness [13]. The front-end chip consists of 8 identical preamplifiers and two drivers.

Neutron and proton irradiation tests were performed at Řež and PSI respectively. Results for the neutron tests in warm and cold conditions are shown in Figure 3.2 for the voltage gain and pre-amplifier non-linearity. It can be seen that both parameters degrade much earlier in cold than in warm conditions. The non-linearity is evaluated in power-law fits of the form out $= a \times \text{in}^x$ for each neutron fluence, where out and in are the output and input pulse heights, respectively. The non-linearities of the pre-amplifiers are especially important here as four pre-amplifiers are usually summed together for one read-out channel and thus any non-linearity on the pre-amplifier level can not be corrected. At a fluence of $3.8 \times 10^{14} n_{\text{eq}}/\text{cm}^2$ (Řež Si NIEL [1]), the gain degradation is about 30% and the power-law exponent quickly rises. At exponents of 0.97 and 1.03 the maximal deviation from perfect linear behavior normalized to the maximum amplitude is 1%, as indicated by the horizontal dashed lines.

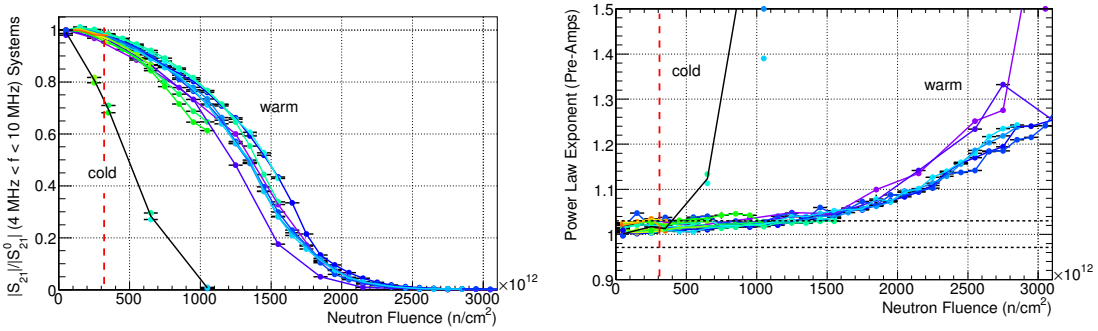


Figure 3.2: Left: Relative gain after neutron irradiation for the current GaAs HEC electronics in warm and cold conditions in the frequency range relevant for the shaper. Shown is the ratio to the signal amplitude before the irradiation as function of the Si-NIEL fluence as measured in Řež. Right: Measured non-linearity of the HEC pre-amplifier response as a function of Si-NIEL fluence as measured in Řež, expressed in terms of a power-law exponent, which has a value of one in the linear case before any irradiation. The red dashed vertical lines indicate the equivalent Řež fluence for $4.1 \times 10^{14}\,\text{n}/\text{cm}^2$ in ATLAS, which corresponds to the integrated neutron flux at the HEC electronics location after 3000fb^{-1}, including a safety factor of 5.

[1] To convert the Si NIEL fluence for GaAs devices in Řež to Si NIEL equivalent in ATLAS one has first to convert to GaAs NIEL and then back to Si NIEL. For Řež the NIEL ratio GaAs/Si is found to be 1.82 and for ATLAS at the location of the HEC cold electronics it is estimated to be 1.35. Thus a Řež Si NIEL fluence of $3.8 \times 10^{14} n_{\text{eq}}/\text{cm}^2$ corresponds to a ATLAS Si NIEL fluence of $5.1 \times 10^{14} n_{\text{eq}}/\text{cm}^2$ for GaAs devices in the HEC PSB region

The measurements in the cold show that at the expected HL-LHC fluences the HEC PAS will possibly operate at their limit and would fail entirely at ATLAS Si NIEL equivalent fluences of $1.2 \times 10^{15} n_{\mathrm{eq}}/\mathrm{cm}^2$. More details about the irradiation tests and consequences for HL-LHC physics in case the cold electronics are not replaced can be found in Ref. [14].

3.2.2.2 Development of new ASICs To address the possibility that the performance of the HEC PAS chips will lead to an unacceptable degradation in physics performance at the HL-LHC, new ASICs are being developed with improved radiation tolerance. To install these, the HEC wheels will need to be extracted from the cryostat, but not disassembled further.

The new ASICs will need to fulfill the same performance requirements as the existing devices but with an improvement in radiation tolerance by a factor of ten. The radiation hardness against neutron irradiation has been studied for transistors in SiGe, Si and GaAs technologies. IHP 250nm SiGe CMOS and SiGe bipolar Technology were selected and further tests were conducted. The observed gain loss after $1.7 \times 10^{15} n_{\mathrm{eq}}/\mathrm{cm}^2$ and $1.8 \times 10^{14} p/\mathrm{cm}^2$ are 0% and 11% for the SiGe transistors and -5% and -7% for the bipolar transistors, respectively. Based on these studies both options, bipolar SiGe as well as CMOS FET technologies, are considered sufficiently stable under neutron irradiation and preamplifiers in both technologies are being developed. Studies of the dynamic range and noise performance are ongoing.

3.2.2.3 Replacement of the HEC cold electronics For the replacement of the HEC cold electronics the cryostat will need to be fully opened, i.e. the warm covers removed on both faces, the inner tube including the LAr FCal lifted to the surface and the cold covers removed. While the EMEC will have to be uncabled from the feedthroughs, the EMEC wheels will remain in place inside the cryostat, well protected against the extraction of the HEC wheels and from the work on the HEC electronics at the outside of the cryostat.

After warming up the cryostat and positioning it for optimal access, one of the main technical issues will be opening it. The cryostat is made of various parts: outer warm vessel, cold vessel, cold cover with inner bore for the FCal, cold bulkhead, inner cold tube, etc. The cold covers, on both faces, are welded. Therefore the welds will need to be cut to allow the opening of the cryostat.

Once the cold inner tube (enclosing the FCal) will have been released, the FCal (see Sec. 3.2.3) and cold cover will be removed. The extraction of the two HEC wheels will then be done with the "Super-T6 [15]" tooling.

Figure 3.3 presents a schematic view of the HEC wheels extracted from the cryostat outside the EC cyrostat.

After replacement of the HEC electronics and successful functional testing, the procedure will be reversed.

The estimated time for the complete operation, once the cryostat has been warmed up and emptied is estimated to 15 months for both sides.

3.2.3 An Upgraded FCal, the sFCal

The ATLAS FCal [16], consisting of three modules (FCal1, FCal2 and FCal3 with FCal1 closest to the interaction point), is designed to operate in the harsh radiation environment produced by the showers from particles in minimum bias events at $\mathscr{L} = 10^{34}$ cm^{-2} s^{-1}, with a comfortable margin [17, 18].

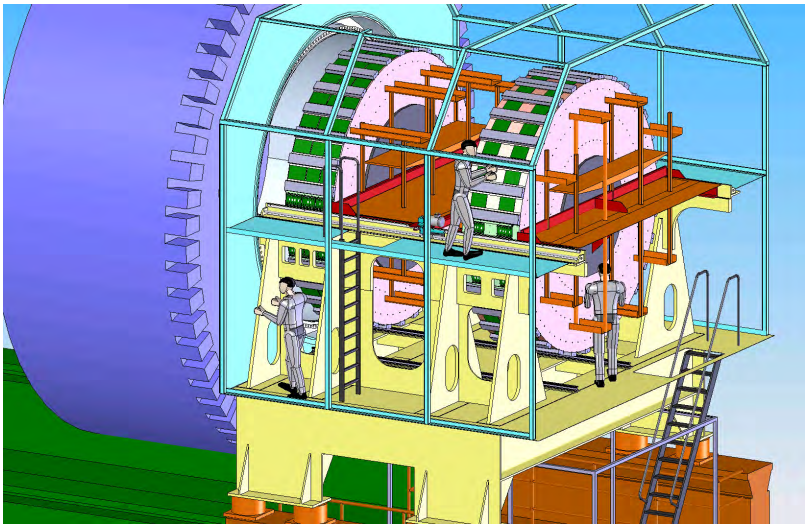
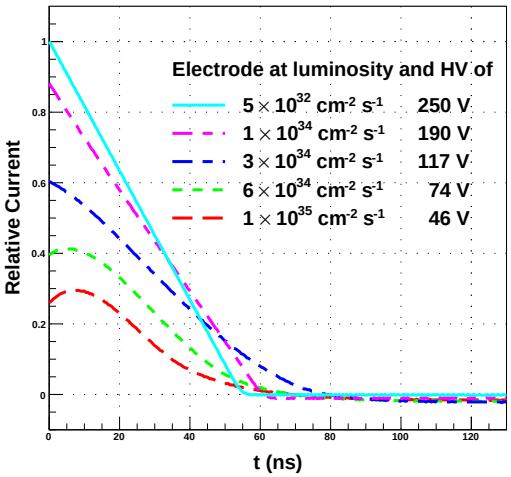


Figure 3.3: The two HEC wheels extracted from the cryostat and supported by the Super-T6 tooling. The protective tent's frame is also shown

Three main effects will alter the performance of the FCal calorimeter when the luminosity rises above the design value: space charge, HV drop and possibly heating up of the liquid argon. The impact of the larger current drawn through the protection resistors, which will drive the gap potential down the HV plateau curve, ultimately leading to a stronger degradation and an enhanced gain instability, will also have to be taken into account. These effects are all included in Fig. 3.4, where the relative degradation of the current pulse is calculated at different luminosities, but averaged over instabilities.

Figure 3.4: Simulated signal pulses at $\eta = 4.7$ for different luminosities, assuming a constant rate from minimum bias background. For each luminosity the same signal energy is deposited in the FCal1 module on top of the minimum bias background and the signal pulse is extracted. For luminosities below 10^{34} cm^{-2}s^{-1} the signal on the electrodes is the familiar triangular pulse. At higher luminosities both the magnitude and shape of the pulse are degraded.



Two on-going basic R&D programs address some of the poorly known parameters of liquid argon in high ionization rate environments. Better estimates will allow an improved prediction of the onset of the space-charge limitations in the ATLAS calorimeters.

The HiLum R&D project is described in [19]. The results, expected sometime in 2013, will

play a key role in deciding which of the options to follow. The goal of the HiLum R&D project is to study the limitations in operation, signal reconstruction and radiation hardness of the endcap and forward calorimeters at very high intensities. Small modules of the electromagnetic end-cap (EMEC) [20], hadronic end-cap (HEC) [12] and forward (FCal) [16] calorimeters have been built, placed in separate cryostats and exposed to hadronic showers from high intensity proton beams at IHEP (Protvino).

The second R&D program uses two specialized FCal1 electrodes with a coaxial cavity machined in a short section of the copper rod. Inside each cavity is a beta source. For the one with the hotter beta source, the result is an approximately 0.5 cm^3 section of the FCal1 module in an ionizing environment similar to that seen near the electromagnetic shower maximum at $|\eta| = 4.7$ at $\mathscr{L} = 10^{34}$ cm^{-2} s^{-1}. The nearly uniform ionization will allow an accurate estimate of the onset of space charge degradation [21]. For the other the ionization rate is about 21 times smaller. Within the standard picture of ionization where electrons and monatomic positive argon ions are produced, the ion mobility and the effect of oxygen contamination will be determined. This second R&D project will also provide valuable input into future decisions.

3.2.3.1 The sFCal For options 1 and 2, the two FCal assemblies at each end of ATLAS would be replaced with upgraded versions (sFCal). The summing boards would also be replaced.

The sFCal modules will have electrodes with smaller liquid argon gaps in order to fend off space-charge limitations. The upgraded sFCal1 module will have gap widths in the range of 100 μm, rather than the 260 μm gaps of the original FCal1 (see Figure 3.5).

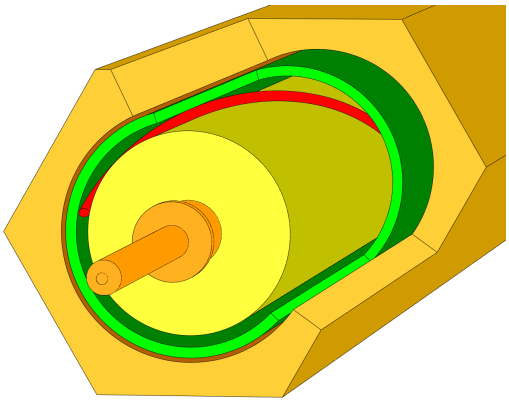


Figure 3.5: Structure of a LAr unit cell in the FCal1 modules. The cell is made by a copper tube (5.250 mm inner diameter) and a concentric copper rod (4.712 mm). An insulating PEEK fibre is helically wound around the rod to keep it centered within the tube. The LAr gap between the rod and the tube is 0.269 mm. The two FCal1 modules are made of 12,260 cells each.

The electrode rods will increase in diameter and the electrode tubes will decrease in diameter so that the mid-line of the gap remains at the same radius for optimal sampling uniformity. The module sizes and weights will remain roughly the same as for the existing FCal.

The summing boards will be replaced with ones having protection resistors with smaller resistances. On the present summing boards the resistors serving the high $|\eta|$ half of the channels are 1.0 MΩ. These will be reduced by about a factor 10, to about 100 kΩ. In the HiLum test beam run the HV distribution system uses protection resistors of this value. Schemes with added protection, such as inductors, to augment the traditional resistors are being considered.

By construction, keeping the sampling frequency the same, the initial current in the sFCal will be 1.5 μA/GeV as in the current FCal1. The electron drift time will go from about 61 ns to about

23 ns, looking approximately like a delta function current pulse to the upgraded LAr electronics.

In option 2, four new feedthroughs would be implemented in the sFCal cold bulkhead in order to carry signal and HV to the sFCal, and as a main consequence, some additional space would be required between the sFCal bulkhead and the cryostat warm cover. The current gap of 28 mm would need to be increased up to 83 mm, affecting the calorimeter envelope inside the ATLAS layout. A modification of about 20 mm of the position along the beam axis of the JD shielding would be necessary. The current plug (behind the current FCal) will be removed. Consequences on the muon background have to be evaluated.

The entire operation of FCal extraction, sFCal insertion and feedthrough modification is estimated to last approximately 5 months per end-cap, starting with the end-cap calorimeter at warm temperature and at the right location. This estimate does not include the cool-down time of the cryostats. In order to save time, activities on the second end-cap are foreseen to start in the shadow of the first one, so that the operation on both sides should stay within 9 months [22].

3.2.4 Mini-FCal Options

3.2.4.1 Introduction This section describes the option for a small calorimeter, named the Mini-FCal, located in front of the current FCal, thus reducing the particle flux in the FCal and allowing it to operate at HL-LHC luminosities. The Mini-FCal is described in detail in Ref. [23] and its location is shown in Fig. 3.6. The Mini-FCal uses copper absorber, has a depth of 300 mm and inner and outer radii of 130 and 350 mm respectively.

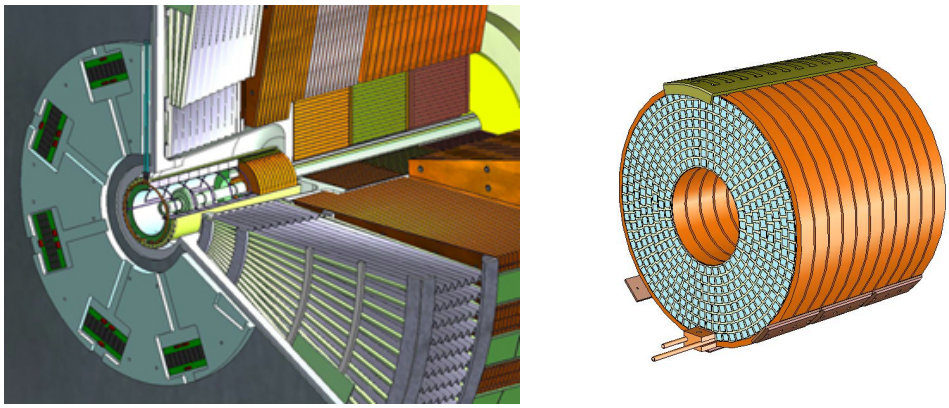


Figure 3.6: The left hand illustration is an overview of the Mini-FCal showing the surrounding detectors and cryostat with part of the beam pipe still in place. The right hand diagram shows the diamond detector option for the Mini-FCal with the first absorber removed so that the diamond detector layer can be seen. The cooling pipes are visible at the bottom.

Three different technologies have been considered for this detector; diamond sensors (DD), High Pressure Xenon gas (HPXe) and Liquid Argon (LAr). Both the DD and HPXe versions of the Mini-FCal are standard parallel plate calorimeters with 12 copper absorber discs and 11 sensor gaps and are designed to fit in and be supported by the existing inner warm tube of the end-cap cryostats [24]. Finite Element studies have been completed that demonstrate that the current warm tube has sufficient strength to support the Mini-FCals which weigh approximately 200kg. The LAr

option has an electrode geometry that is identical to that proposed for the new sFCal1. This option is contained in a small single wall cryostat that is itself located in the end-cap cryostat vacuum between the inner warm tube and the FCal. This option would require a new inner warm tube to be designed and built.

Although extensive R&D has been carried out on options for the Mini-FCal, further work is required to select the best option for ATLAS. Recently the major supplier of diamond detectors went out of business; new companies are developing this technology and these are likely to be available in the near future. Basic R&D work is required for the HPXe option as there is relatively little information on gas properties (drift and recombination rates) at the required pressures. On the other hand the detector technology for the LAr option is well understood, but engineering is required for developing the cryogenics that will in turn allow an estimation of the amount of material required for this option. These material amounts would then be used as input for simulating its response.

The readout electronics for the Mini-FCal would be placed in the empty slots of the LAr end-cap FE crate containing the FCal electronics. Use of the LAr front-end crate will require the installation of special transition boards to route the signals to the baseplanes and the FEB2s. A monitoring board will also be required to process auxiliary data such as temperature and pressure. There are two possible locations for the Mini-FCal preamplifiers, on the FEB2s or on the front face of the LAr end-cap cryostat as shown in Fig. 3.6. Locating the preamplifiers on the face of the cryostat would minimize the cable lengths to the detectors, however it would also require radiation tolerant electronics that would work at this location.

An extensive set of simulations has been carried out to investigate the use of the Mini-FCal in ATLAS including studies of particle flux levels in the end-cap detectors with and without the Mini-FCal, determination of the single particle response of the end-cap system with the Mini-FCal and studies of energy resolution as a function of sensor ganging and electronics noise for the DD option. These are described in detail in Ref. [23]. The studies of particle fluxes demonstrated that the Mini-FCal, with 18.8 radiation lengths, would reduce the particle flux in the high η region of the FCal by a factor of 2-3 which is sufficient to prevent boiling of the liquid argon and keep the ionization rates in the FCals at a level where the signals would not be degraded. Single particle studies using full simulation, including the effects of radiation damage to the diamond sensors after an integrated luminosity of 3000 fb^{-1}, showed that the installation of a Mini-FCal maintained the detector response at high η, but did introduce a region of reduced response at the transition region at the outer edge of the detector. This is illustrated in Figure 3.7.

For the Mini-FCal installation, the impact on the cryostat is limited. Indeed, only the warm inner tube may be affected, depending on which technology is selected. The estimated time for the installation is four months.

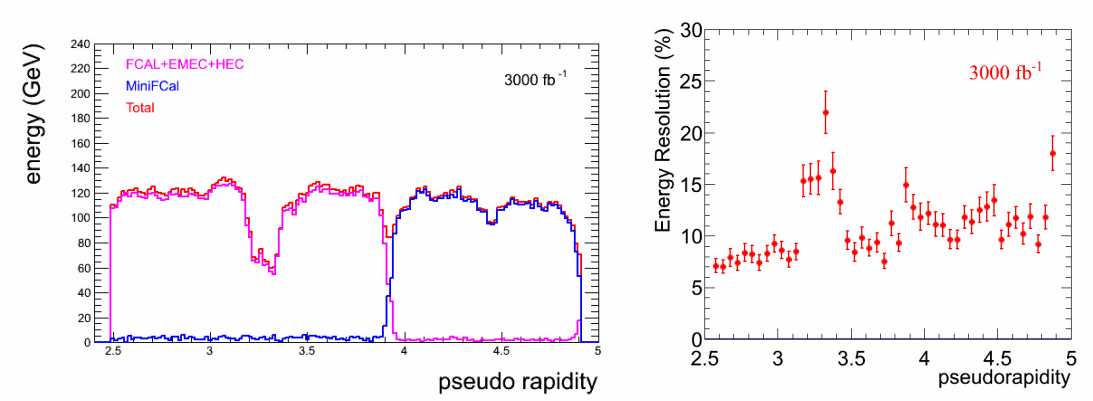


Figure 3.7: Plots of the single 120 GeV electron response (left) and resolution (right) for the calorimetry system with a Mini-FCal as a function of $|\eta|$. The energy response plot shows how much energy is deposited in each of the Mini-FCal, the FCal, the HEC and the EMEC and the energy sum.

4. Upgrade of the Tile Calorimeter

A full replacement of the readout electronics of the hadronic Tile Calorimeter is foreseen for the Phase-II upgrade to be compatible with new L0/L1 trigger architecture, to meet the increased radiation tolerance requirements of the front-end readout electronics and to provide higher granularity information to the trigger processors. The upgraded on-detector electronics will be organized in independent modules (e.g. separate power, cooling and monitoring services), which cover half the size of the existing drawers. To improve serviceability and for ease of maintenance, the mechanical design of the drawers is being reviewed.

4.1 Upgrade of the TileCal readout

The readout architecture is based on continuous digitization and data transfer off-detector at each bunch crossing of all the readout channels. The design includes a fully redundant optical transmission of the data: no gain selection logic will be implemented in the front-end and the digitized signals of both gain ranges will be transmitted on two different transmitters. Pipelines and derandomizing buffers are implemented in Pre-Processor (PrP) units in USA-15. The PrP modules will apply energy scale calibrations as in the existing RODs and prepare trigger primitives to be transmitted to the Level-0 trigger feature extractors and the Level-1 processors.

4.1.1 On-detector electronics

The front-end readout design relies on the use of high reliability components, with a strong emphasis on radiation tolerance. Since the TileCal electronics are shielded from radiation by the calorimeters, the radiation dose requirements to be met is at most 100 kRad over 10 years at the HL-LHC conditions. ASIC designs based on modern technology processes with feature sizes of 130nm or less are known to meet the requirements in terms of total ionizing doses and neutron displacement damages. A few commercial available components have been already qualified for the doses and fluences expected in the location of TileCal drawers. However, it may be challenging to find suitable fast switches, FPGAs, and optical transceivers because of the sensitivity of such devices to single-event upsets, whose cross-section in fact increases as the technology feature size becomes smaller. Soft errors in the configuration memory and the logic in FPGAs are a particular concern, for which error correction and triple redundant designs must be studied. The design needs also to adapt to new design strategies and new infrastructure, such as the use of the GBT [25], new timing and trigger protocols (TTC) and the use of high bandwidth optical transceivers.

At the system level, the front-end readout is being redesigned to optimize the reliability of the interconnects and of the sub-assemblies in the drawers. The experience gained during the detector operations provided valuable information on how to simplify the stacking of the different boards in the drawers and to improve the reliability of the interconnections and avoid single-point failures, in particular for the power distribution, electromagnetic compatibility and interference.

The dynamic range required to resolve the highest jet energy densities is 16-bit above the noise level. Two 12-bit gain ranges allows to cover the full dynamic range and still provide sufficient precision for low transverse energies and muons. It will also improve significantly the performance of the trigger primitives sent to the first level of the trigger, which currently uses 8-bit digitizers and covers an equivalent 1 GeV - 250 GeV range in transverse energy.

In the current planning, the system would be configured to have three types of PCBs in the low voltage drawer electronics (Fig. 4.1a), namely front-end boards (FEB) for pulse conditioning and calibration, a "MainBoard" for transferring data (and digitizing if this is not done in the FEB) to the processing DaughterBoards, which contain all other functions. The MainBoard would act as an interface to the multiple front-end cards, and the DaughterBoard as the host of the connections and communication with the off-detector electronics. Three strategies are being pursued for the redesign of the front-end cards. One of the options considered is an optimized redesign of the current 3-in-1 front-end board that uses state-of-the-art integrated circuits. The signal is filtered in a passive shaper, amplified in two separate ranges, and sent to 12-bit ADCs on the MainBoard. The new 3-in-1 board has shown better linearity than the present version and was found to be sufficiently radiation tolerant.

A second approach is to integrate much of the functionality of the current 3-in-1 card into a custom ASIC: preamplifier, shaper and ADC are integrated in a single customized IC. The ASIC design is based on a 130nm CMOS process already qualified in ATLAS for applications in other subsystems with much higher radiation doses and fluences. Given the limited maximum voltage swing of 130nm feature size devices, three gain ranges would be necessary to cover the full dynamic range.

A third approach is based on the QIE ASIC developed at FERMILAB. The QIE consists of a current splitter that splits the current into 4 different paths corresponding to 4 different signal ranges, integrates each of them in gated integrators and digitizes them. The entire operation takes 4 clock cycles and the performance is well within specifications. A full prototype is expected in the spring of 2013.

A technology decision will be made around 2015 when the relevant R&D have been completed, and when the different designs will be evaluated in a test beam environment. The MainBoard design must be adapted to the chosen FE-board design but the DaughterBoard that serializes the data and transmits it to the Pre-Processor boards would be the same. In all schemes, the Charge Injection System and Cs integrator are similar to the existing system and are located on the FE-board.

To protect against single point of failure and to improve reliability of the system a full redundant readout is implemented as follows: each Tile calorimeter cell is viewed by two PMT. Signal condition, digitization, serialization and data transmission of those two PMT is kept separate from a functional point of view. Thus, each MainBoard/DaughterBoard combination essentially consists of two decoupled halves that each process 6 phototubes. Furthermore, each MotherBoard/DaughterBoard module is fed by its own power source and uses its own optical transceivers.

Data will be transmitted via the GBT protocol, which uses limited forward error correction (FEC). Utilizing the inherent redundancy in TileCal, data from the two halves will be transmitted in separate data streams containing both high and low gain data, and as additional protection we intend to duplicate these streams. The present GBT system can accommodate this with 8 (4.8Gbps) GBT transceivers. A solution based on radhard Modulator/WDM combinations allowing transfer speeds of 10 Gbps and more) is also under study.

The low-voltage power distribution system in the drawers will also be upgraded to address: space constraints, the presence of a magnetic field, and concerns about efficiency and heat. The power distribution system continues to rely on low-noise DC-DC converters. The design includes

a bulk +10V DC-DC converter, bussed to all the different loads, with a proximity "point-of-load" (POL) regulation. Radiation tolerant POLs are presently under development at CERN. Their use will help eliminate stringent regulation requirements at the load over long distances, as well as improving the efficiency and load-balancing. In addition, a redundant scheme with two +10V sources connected to the point of load via diodes is being engineered in order to eliminate single-point failures.

Higher luminosity will increase the current flow inside the PMT divider, decreasing the inter-dynode gains of the last stages and causing a nonlinear behavior for very high signal levels. R&D is currently ongoing to identify a new divider design in which the 2 or 3 last stages use radiation tolerant transistors to balance charge effects.

4.1.2 Off-detector electronics

Pre-Processor boards that can process data for one or two entire ϕ-slices ($\Delta\phi = 0.1$, $|\eta| < 1.7$) in a single board can be designed based on current technology, minimizing the board-to-board communications in the back-end crates. All high speed communication would then occur inside the boards, leaving the backplane to configuration and monitoring. The algorithms that extract the signal features (e.g. energy, timing and quality factors) can be implemented in FPGAs, combining computational power and flexibility. The main functional blocks of the back-end electronics are similar across the calorimeter systems (see Figure 2.12 in section 2.7). A possible implementation to process the data from one TileCal ϕ-slice is shown in Figure 4.1b. The trigger Pre-Processing and the link with Level-0/1 calorimeter trigger system are accomplished in separate FPGAs, whereas data reception, pipeline and data processing would be implemented in 8 FPGAs processing data from 4 miniPOD optical fibre connector/receivers each (POD, Parallel Optical Device). The connection with front-end modules can be done via front panel MT fibre connectors. With front panel MT fibre connectors linked to miniPOD receivers on the circuit board and higher performance FPGAs, it should be possible to double the density to 2 ϕ-slices per Pre-Processor board.

Data transmission to the next stages of the TDAQ system, as well as control and configuration functions are performed through the backplane, based on the ATCA standard. The configuration and control interface needs to provide communication with the Pre-Processor, for example to download weights and calibration constants used for the energy reconstruction and the extraction of the other signal features.

The total bandwidth required by the Tile calorimeter is around 9.5 Tbps. It takes into account a total of 9856 PMT signals, two gain settings each digitized (12-bits) at 40 MSPS. If the readout is logically partitioned in ϕ-slices (2 Extended and 2 Long Barrel modules in the same ϕ coordinate) the input bandwidth needed for the 154 channels in a single PreProcessor board would be 150 Gbps and 300 Gbps with redundant duplication. This could be taken care of by 84 4.8 Gbps GBT fibres, but different levels of modularity (6PMT blocks, miniPODs) suggests 200 or 400 fibres for 2ϕ-slices (Figure 4.1b).

The data should be continuously processed to deliver energy estimations for detected pulses with a secure time relation to the correct bunch crossing. This requires extending the algorithms used in the present Level-1 calorimeter trigger preprocessor. Whether to perform these operations of a tower or cell level has to be determined. A middle way is to operate pulse recognition on the tower level and energy extraction on cell level. The pre-processing block should provide merging of

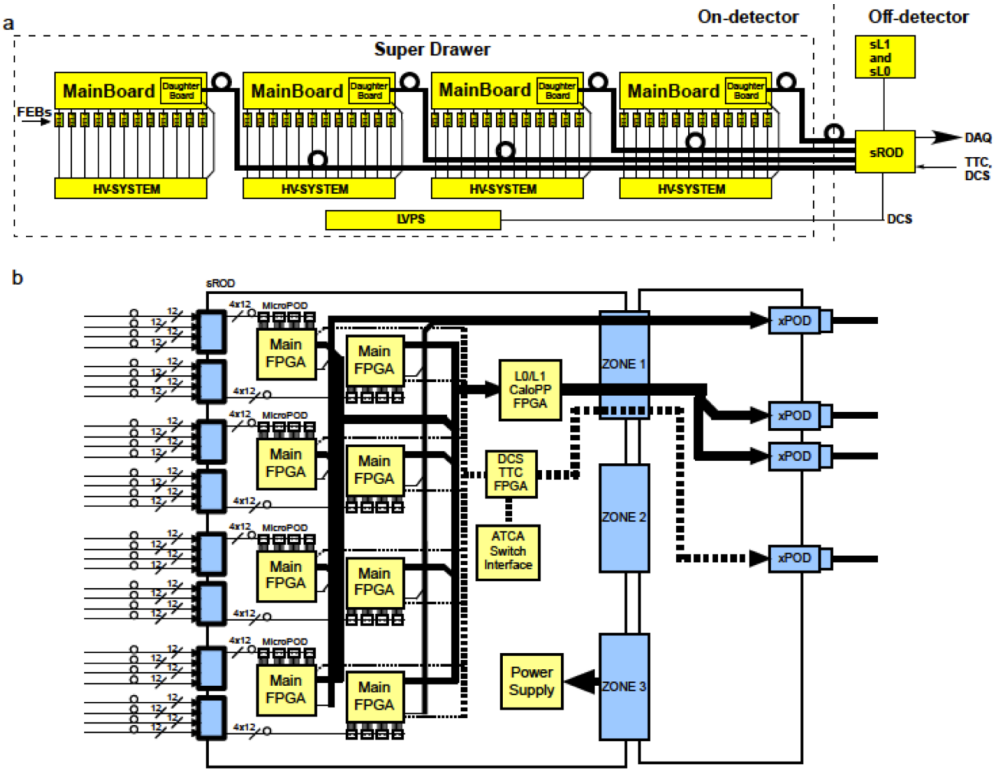


Figure 4.1: (a) Block diagram showing the proposed TileCal front-end electronics upgrade. (b) Architecture proposal for a Pre-Processor module (not including the down links).

cell sums, pulse identification and feature extraction for the L1Calo system. The feature extraction and merging require extensive simulations studies before any complete algorithms can be coded.

The Pre-Processor boards will reconstruct energy, timing and quality factor at the rate of the Level-0 trigger system. DSP-like slices in modern FPGAs allow to implement optimal filtering or more sophisticated algorithms under development. Upon receival of a Level-1 accept signal, the calibrated data will be presented to ATLAS standard ROD/ROB/ROS systems through a switched network as proposed in Sec. 2.7. Calibration data analysis, histogramming, monitoring and Level-2 trigger algorithms can be implemented at that stage.

4.1.3 Precision calibration

The basic function of the three independent calibration systems used by TileCal, i.e. Cesium, laser and charge injection, is the same as in the present system. However, most on-detector electrical components need to be replaced and upgraded due to aging, to adapt to the new readout system and to cope with the increased radiation levels. It is also desirable to improve the reliability wherever possible.

4.2 The TileCal Demonstrator

To validate the readout architecture, its interface to the trigger system, and the hardware implementation of the novel trigger algorithms, installation of a hybrid “demonstrator” front-end readout in

ATLAS is planned. Here a probable phase 2 drawer implementation is combined with circuitry to produce analog trigger signals so that the drawer can adapt seamlessly to the present system. The first new Tile drawer will be installed during the Phase-0 shutdown (LS1). If successful, i.e. if does not compromise normal data taking, three more drawers will follow during the winter shutdowns of the following years. The "demonstrator" will read out a $\Delta\eta \times \Delta\phi = 0.7 \times 0.4$ slice of the barrel calorimeters, matching the size of the current Level-1 electron trigger algorithm clusters. The demonstrators will be based on modified 3-in-1 front-end option as it can be easily modified to produce the analog trigger data.

5. Upgrade of the Muon System

The HL-LHC luminosity upgrade requires a substantial performance increase in elements of the muon spectrometer system. This is true for the muon precision tracking as well as for the triggering system. The Monitored Drift Tube (MDT) precision tracking chambers will have to maintain excellent spatial resolution and tracking efficiency in a high hit rate environment caused by γ conversions and neutron reactions. The MDT readout system must cope with the higher occupancies and an increased L1 trigger rate, which will well beyond the present limit of 100 kHz.

The trigger system, on the other hand, will have to maintain a high level of efficiency for finding high-p_T tracks, improve the rejection of fake triggers which is presently already responsible for about 90% of the L1 triggers. The goal is to keep the total L1 trigger rate due to the muon system inside acceptable limits. Figure 5.1 shows the composition of the overall L1 muon trigger in the 2011 and 2012 runs as a function of η as classified after the full event analysis. The magenta diagram shows the distribution of valid L1 triggers with $p_T > 20$ GeV. Tracks which do not pass through the Small Wheel are in white, in red are those that do not come from the interaction point and in green are tracks with radial coordinates in the EI and the EM station that do not match. The β cut, finally, selects tracks with a small deflection angle in the magnetic field, consistent with a track above 20 GeV. With the Phase-I upgrade due in LS2, where new Small Wheel (nSW) will be constructed and installed, the first three sources of fake triggers in the end-cap will be eliminated. In addition, the muon angle θ will be measured with sufficient accuracy (< 1 mrad) to discard tracks which do not come from the IP.

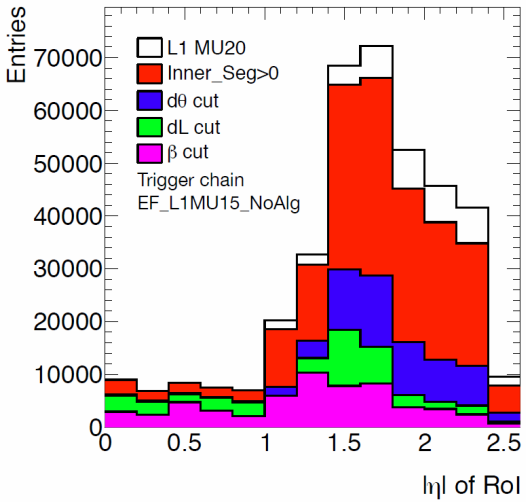


Figure 5.1: Composition of the L1 muon trigger. The magenta diagram shows the distribution of valid high-p_T triggers, coming from the IP (see text).

The new ATLAS-wide L1 trigger system, foreseen for Phase-II, will impose new requirements to the trigger readout systems as well as on the precision chambers. In the new scheme the present L1 trigger, with its 2.6 μs latency, is split into two separate steps with latencies of ~ 6 μs and $\sim 14-19$ μs respectively. The first step results in an accept when a track candidate is detected

by one of the primary trigger sources, i.e. muon spectrometer or calorimeter (e.g. a muon track with $p_T > 20$ GeV). This "L0 trigger" is broadcast to the detector front-ends to flag data of the corresponding beam crossing for intermediate storage. The L1 trigger is generated after an additional delay of $\sim 14 - 19\mu s$, filtering the L0 trigger sample using additional criteria applied to event topologies ("topo trigger"), for example an isolated high-p_T muon as defined in the trigger menu.

The maximum value of the L1 latency $\sim$25 μs is dictated by a technical limitation of the present MDT readout electronics, which, at the expected tube hit rates, has a limited storage capacity. To overcome this, the readout electronics cards should be substituted with new ones. This is possible for most of the chambers, but access is difficult for chambers located mostly inside the toroid coils. Installation of the replacement electronics would require a major intervention to move the toroids to allow access to the chambers. However, upgrading the readout electronics will allow the precision coordinates of the MDT to be used at the trigger level to improve the turn-on of the high-p_T trigger threshold, significantly reducing the number of low-p_T muons passing the muon trigger.

The muon spectrometer upgrade performance at HL-LHC luminosity is based on extrapolations of trigger rates and hit frequencies from simulation and on the rates observed in the present running experiment. The performance of the existing MDT chambers together with the new EI stations, upgraded in LS2, appears to be sufficient for operation at the HL-LHC. However, the MDT readout electronics will have to be adapted to the new trigger scheme and to the increased hit rates.

The present readout system of the trigger chambers in the barrel (RPC) and in the end-cap (TGC) will not be able to cope with the new L0/L1 trigger scheme, as the present electronics is designed for maximum latencies of 3.2 μs (TGC) and 6.4 μs (RPC), and for trigger rates up to 100 kHz. Consequently, the whole readout electronics chain will have to be replaced, offering the opportunity to rebuild the readout chain with modern technologies. This will allow improvements to important performance aspects such as the readout bandwidth and the flexibility of trigger algorithms, as discussed in sections 5.2 and 5.3.

Another important aspect for the upgrade of the trigger electronics is the necessity to improve the selectivity for high-p_T tracks, which calls for an improved spatial resolution in the bending direction η. The spatial resolution is presently determined by the construction of the chambers, such as the width of pick-up strips in the RPC and the modularity of wire-ganging in the TGC, which cannot be changed unless the current trigger chambers were replaced. The present readout electronics of both trigger chamber systems does not supply information on the signal pulseheight and therefore charge interpolation can not be used to possibly refine the position measurement. For the upgrade of the readout electronics with respect to charge measurement in the RPC chambers see section 5.2.1.

The determination of the p_T of the candidate trigger track could be improved using the precision spatial information from the MDTs. This requires the development of new MDT readout electronics to allow the MDT coordinates to be made available to the trigger logic within the L0 or L1 latency. According to simulation, sharpening of the high-p_T threshold would reduce the fake muon trigger rate by up to a factor 3, which makes this concept an attractive option for the trigger system upgrade in Phase-II.

5.1 Upgrade of the MDT tracking system

At local hit rates above 500 Hz/cm^2 the MDT will start to lose spatial resolution and efficiency due to the build-up of space charge in the gas volume, leading to drift field fluctuations and reduced gas gain, see [24], [26] and [27]. Hit rates at HL-LHC luminosity, as extrapolated from recent measurements, indicate that the existing MDT chambers with their 30 mm diameter drift tubes can safely operate in most of the muon spectrometer, as the hit rates remain significantly below the 500 Hz/cm^2 limit. Higher rates only occur at the inner part of the EI station ("Small Wheel"), which will be in be addressed by the installation of the new Small Wheel in LS2. Column 4 of Table 5.5 shows the expected tube hit rates at 7×10^{34}cm^{-2}s^{-1} in the barrel and end-cap (excluding chambers of the EI station). The distribution is relatively uniform with values up to ∼60 kHz/tube in the Outer Wheels (OW), ∼100 kHz/tube in the barrel and ∼150 kHz/tube in the Big Wheels (BW). This indicates that efficiency, spatial resolution, link saturation and relative buffer occupancy will remain at acceptable levels in barrel and OW.

MDT tube location	Tube length	Expected hit rate	Hit rate / tube	Occup-ancy	Tube eff.	Link sat. @200kHz	Buf. occ. @25μs	
							absol.	relative
		Hz/cm^2	*kHz*	*%*	*%*	*%*		*%*
BW, inn	150	273	123	9	91	75	147	58
BW, mid	240	196	141	11	90	82	169	66
BW, out	430	91	117	9	92	73	141	55
OW, inn	202	70	42	3	97	45	51	20
OW, mid	298	56	50	4	96	48	60	23
OW, out	500	42	63	5	95	53	76	30
BI	260	105	82	6	94	60	98	38
BM	353	98	104	8	93	68	125	49
BO	494	63	93	7	93	65	112	44

Table 5.5: Typical hit rates and performance of the MDT chambers and their readout in Phase-II at 7×10^{34}cm^{-2}s^{-1}. Occupancies and efficiencies in cols. 5 and 6 are calculated on the basis of a dead time of 750 ns, corresponding to the maximum drift time in the MDT. Column 7 gives the saturation of the link from mezzanine card to CSM (80 Mbit/s) at a L1 trigger rate of 200 kHz. Cols. 8 and 9 give the average filling of the 256 words deep L1 buffer in the front-end TDC at a latency of 25 μs. (BW = EM wheel, OW = EO wheel, BI, BM, BO: inner, middle and outer MDT layers in the barrel.)

In the present trigger system the L1 trigger arrives at the MDT front-end with a latency of about 2.6 μs relative to the corresponding beam crossing (BX), and the drift time data need to be stored in the buffers of the TDC during this period. After LS3, the L1 trigger processor (CTP) will send out two distinct triggers, L0 and L1, after a latency of ~ 6 and $\sim 14-19\mu$s, respectively. The longer latencies imply an increased residence time of the data in the buffers and, thus higher buffer occupancies. If the average occupancies grow too high, the derandomizing function of the buffers will become ineffective and depending on local hit rates, may result in data loss.

The MDT readout electronics will also have to comply with the increased trigger rates. This means higher data rates to be read out from the front-end cards ("mezzanines") to the Chamber

Service Module ("CSM") and from there to the off-detector electronics in USA15. Rates up to 500 kHz and 200 kHz are presently under discussion for the L0 and L1 trigger respectively.

Columns 7 and 9 give the link saturation and L1 buffer occupancy. In the Big Wheel (BW) the figures show a link saturation beyond 70 %, a figure which is considered the limit for operation without data loss due to rate fluctuations. New readout electronics in modern technology will have to be installed to overcome this limitation.

The replacement of the front-end readout electronics will be a major issue during LS3. It is not obvious that it is possible to exchange the front-end electronics everywhere. Possible intervention scenarios for the BI chambers are under investigation, but currently it is not clear that such an access is possible without a major dismount of the chambers. If no solution can be found then the usage of the present readout electronics will have to continue for the HL-LHC running period. Under the conditions assumed for the figures in Table 5.5, the L1 buffer occupancy of 38% in the BI chambers (at 25 μs latency) would still be acceptable, while the link saturation of 60% would provide little safety margin for a possible unforeseen increase of the tube hit rates. In the case of an increase of link saturation beyond the limit of data loss, the readout of the trailing edge would have to be suppressed, reducing the link saturation by nearly a factor of 2.

To fully exploit the inherent rate capabilities of the MDT drift tubes, even at latencies beyond 20 μs, a number of modifications to the readout electronics will be necessary. A detailed description of the present MDT readout electronics is given in [28]. The most important performance improvement would come from the replacement of the presently used TDC (see [29]) by the more performant HPTDC [30], see figure 5.2, which provides longer L1 buffers and thus allows for a corresponding increase in the L1 latency. The replacement of the TDC would also lead to a reduction of the data rate to be transferred via the 80 Mbit/s link towards the CSM. The data of the leading and trailing edges, currently represented by two separate 32-bit words, could be merged into one 32-bit word, reducing the link saturation by nearly a factor of two. In the standard readout mode the duration between trailing and leading edge does not correspond to the Time-over-Threshold but is used to encode the ADC reading of the ASD preamplifier, which provides information for time-slewing corrections to the TDC measurements and monitoring the gas gain [28].

In addition to these modifications at the front-end, the CSMs and RODs will also have to be adapted to the new readout architecture, the GOL link being replaced by the GBT [31] and the presently used TTC system by a more complex and flexible trigger distribution system.

5.2 Upgrade of the RPC Muon Barrel trigger

The present RPC Level-1 Muon Barrel Trigger system will not work correctly at the proposed 500 kHz L0 trigger rate. The present limitation is a maximum of 100 kHz readout rate, although the maximal acceptable latency of 6.4 μs satisfies the future L0 latency requirement. The trigger rate limitation mandates significant changes to the existing infrastructure, including a redesign of large parts of the trigger system.

The system redesign will take advantage of the evolution of technology since the mid-nineties. Most of the circuits performing trigger algorithms will be moved from on-detector to USA15 using modern optical links capable of providing large bandwidths. Another important change will be the use of FPGAs to execute the trigger algorithm, rather than the use of dedicated ASICs. Moving

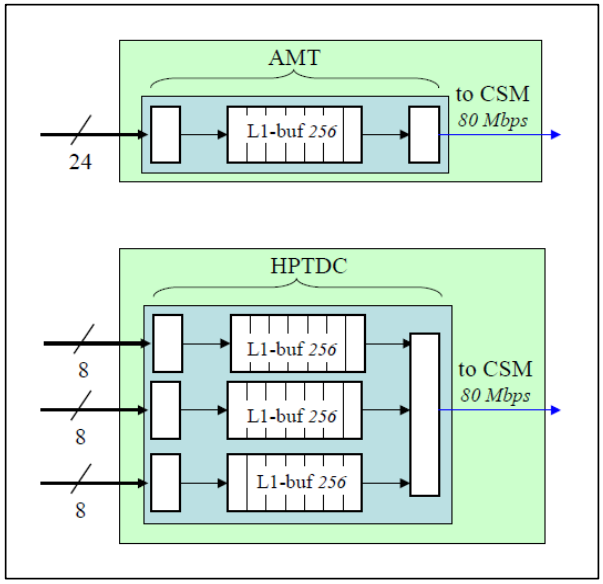


Figure 5.2: The readout of the MDT front-end (schematic). The top diagram shows operating parameters of the existing readout electronics in the present operation mode. Below: the use of a new TDC type (HPTDC) provides more storage space in the L1 buffers and supports a more efficient readout mode (”pair mode”).

the trigger algorithm circuits off the detector grants much more flexibility and simplicity in the commissioning, operation and maintenance of the system.

A new Data Collector and Transmitter (DCT) trigger box will replace the present pad box, see figure 5.3 incorporating the function of the present splitter box. The main task of the DCT, located on top of the detector, is to receive data from the RPC front-end and split and send it to the pad box in a format required by the trigger algorithm. The box can be either remain unchanged or be removed. This solution will avoid a complex and difficult reworking of the on-chamber cabling.

The DCT box will collect trigger data from a single tower, synchronize them with the LHC clock and apply zero suppression before transmitting them to USA15 together with their timing information. Zero suppression is necessary to reduce the 40 MHz of trigger data to be sent to USA15 to a manageable size.

In addition, one foresees the separation of the on-detector trigger data for high- and low-p_T, which will be transmitted separately to USA15 where the trigger algorithm will be performed. The existing connection between low-p_T and high-p_T part of the barrel will be removed, while the system will maintain the actual logical organization in trigger towers and ROIs (four per tower).

The exact link speed needed to transmit the trigger data from on-detector to USA15 will be known once the reduction algorithm is defined. A possible link candidate is the GBT system in the foreseen new 10 Gbit/s version. Other systems could be viable candidates too. The GBT has the advantage of a certain number of accessory facilities that could be very useful in the design of the system. A total number of about 800 DCT boxes and optical links will be used. The number of required links will double, since low-p_T and high-p_T will have separate connections.

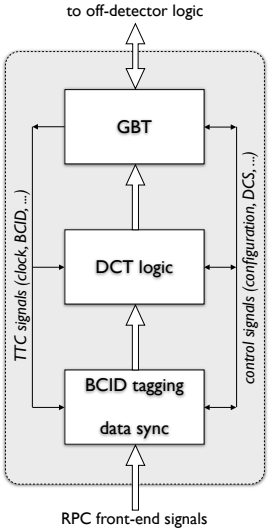


Figure 5.3: Schematics of the new RPC readout electronics

The off detector electronic structure will be fully redesigned, maintaining the present subdivision in 64 sectors. Each Sector Logic will receive data from 12 or 14 DCT boxes (depending on the sector type). After executing the trigger algorithm, the Sector Logic will send out the trigger results to the MuCTPi via 64 optical links, which for compatibility reasons will be the same as the one which will be implemented in the phase-I upgrade.

The Sector Logic will be probably housed on an Advanced Computer Telecommunication Architecture board (ACTA). One board per sector will be used. The trigger algorithm will use FPGAs, which are powerful enough for this task and do not require the design of any new ASIC.

The RPC data readout from on-detector to USA15 will use the same link as for the trigger data, while the ROD can be expected to be the same as for the readout of the MDT. The present readout architecture will be maintained with one ROD for two sectors for a total number of 32 RODs. An approach of one ROD per sector might also be considered.

For the timing and control distribution the plan is to use the TTC channel implemented in the GBT system, following CERN developments. For the Slow Control, the plan is to use the CERN GBT-SCA Slow Control Adapter ASIC, which works in conjunction with the GBT.

5.2.1 Improvement of the RPC tracking accuracy

The replacement of the RPC readout electronics allows new functionality to be added to the front-end readout system with the aim to exploit up-to-now unused information supplied by the RPC chambers. Measurements have shown that the signal duration from the Amplifier-Shaper-Discriminator chips (ASD) is a measure of the charge deposited on the corresponding strip. For Phase-II the signal duration could be measured by TDCs with appropriate time resolution. The charge distribution in a cluster of strips could subsequently be used to determine the position of the track with higher accuracy than what is available in the present system. Indeed, the present RPC readout is tuned to get a small cluster size (1.5 in average), and no pulse-height information is read

out. The successful use of the charge centroid as position estimator in RPC detectors was already demonstrated in beam tests [32, 33], with RPC types of different strip pitches.

The accuracy to be obtained with the strip pitch of the presently used RPC depends on a number of performance parameters, and a R&D project has been started for a detailed evaluation.

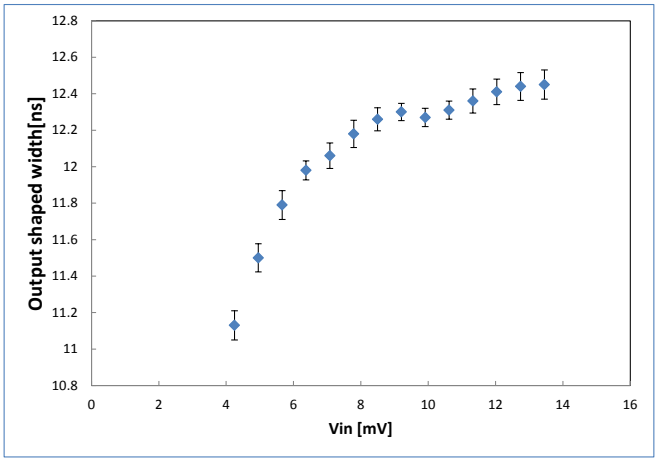


Figure 5.4: Output duration of the RPC signal versus input amplitude on a RPC strip.

A preliminary test of the charge measurement has been carried out with a standard RPC front-end board. Figure 5.4 shows the plot of output duration versus input amplitude. The reported amplitude range covers the small amplitudes expected for the strips adjacent to the one crossed by the muon. The test shows that a factor of 3 in the amplitude of the analog input signal produces a change of 2 ns in the duration of the shaped output signal. A study is ongoing to further improve the response dynamics for the large input charges. The result of the test is that an adequate measurement of the charge for the centroid calculation would require a time resolution of the order of 0.1 ns.

5.3 Upgrade of the Muon End-cap trigger (TGC chambers)

The muon Level-1 trigger in the end-cap is generated by Thin Gap Chambers (TGC), which determine the deflection of tracks in the endcap toroidal magnet by measuring the track angle behind the magnet in three TGC layers in the Big Wheel [24]. The deviation of this angle from the direction of a straight line through the nominal position of the interaction point (IP) is used as a measure of the muon momentum.

The current data taking has shown that one of the major sources of fake triggers is due to tracks which do not come from the primary vertex. To veto those triggers, a precise measurement of the track angle *before* the magnet is mandatory. A new type of TGC with high spatial resolution will be installed in the EI station already for the Phase-I upgrade as part of the new Small Whell (nSW) project. These new trigger chambers will measure the track angle in η and ϕ with a 1 mrad accuracy. Tracks with a significant angular deviation from a straight line through the primary vertex will be discarted.

There are two principal sources of fake muons in the muon trigger: tracks arising from backgrounds and that do not point to the primary vertex and low-p_T muons due insufficient sharpness of the high-p_T threshold. The first set of fakes will be addressed by the installation of the new Small Wheel during LS2. The second set of fakes are due to insufficient spatial resolution in the η coordinate of the TGCs in the Big Wheels. A possible solution would be to replace the Big Wheels with high resolution TGCs that are being built for the NSW, in particular for the high-η region where trigger and background rates are particularly high. If this option is pursued, the inner ring of the TGCs of the Big Wheel will be replaced by new TGCs.

In the outer part, however, such a possibility does not seem realistic given the large number of chambers necessary. Therefore, the only way to sharpen the trigger threshold in the outer part would be the use of MDT coordinates for a better determination of the deflection angle, see section 5.4.

The new L0/L1 trigger scheme will require a complete replacement of the TGC readout electronics chain, with the possible exception of the ASD preamplifiers. This will allow the implementation of a new architecture for the readout, where most of the logic functions will be moved to the radiation-free zone in USA15. This allows FPGAs to be used for L0 trigger formation as well as for data processing and transmission, and provide a higher level of flexibility.

5.4 Use of MDT precision coordinates to sharpen the high-p$_T$ threshold

In the present architecture, trigger and precision chambers operate independently. The information of both chamber systems is read out separately, the precise tracking information of the MDT is currently only used at Level-2, where the majority of trigger candidates is then discarded. Many of them are caused by tracks with p$_T$ below threshold or by tracks not pointing to the vertex.

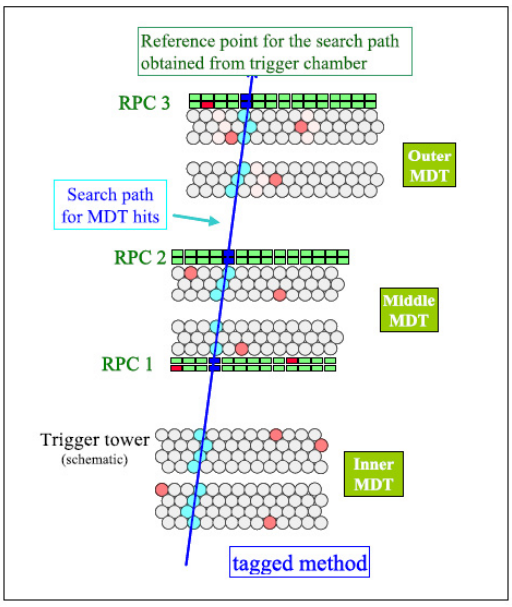


Figure 5.5: The "RoI seeded" method. The RoI information for a high-p$_T$ track from an adjacent trigger chamber (here the outer chamber RPC3) is used as a search road for MDT hits of the candidate track. Track segments of several MDT chambers can be combined to give a precise p$_T$ estimate for the candidate track.

To apply the same effective method of trigger rate reduction already during the formation of L0 or L1, the precise MDT track coordinates must be fed directly to the trigger logic. This can be done fast enough to be used for a refinement of the p_T measurement inside the L0 latency of 6μs. It can implemented by introducing a second fast readout path, parallel to the existing one into the MDT readout architecture. This "fast readout" would have to supply MDT tracking information in a synchronous way with the beam crossing (BX), making the accurate MDT coordinates available to the trigger logic at a fixed time interval after the corresponding BX and assuring the correct assignment of MDT tube signals to a certain BX.

To simplify the drift time measurement for the fast readout, the clock frequency of 1.28 GHz, used internally by the TDC, corresponding to a least significant bit (lsb) value of 0.78 ns, can be relaxed to 40 MHz (or 80 MHz) for a fast, approximate drift time measurement. With the average drift velocity of 25 μm/ns in the MDT tubes, the corresponding lsb of the position measurement is 0.5 mm (or 0.25 mm). These numbers for the η binning of the MDT are more than an order of magnitude better than what is available from the trigger chambers, namely ~ 3 cm strip width in the barrel (RPC) and 3–6 cm η-segmentation in the end-cap (TGC). Consequently, a significant improvement of the p_T measuring accuracy can be achieved in all regions of the spectrometer.

Two different approaches can be pursued, depending on whether or not the early information from the trigger chambers, identifying a region of interest (RoI) for the high-p_T track candidate, is used or not.

The usage of the trigger chamber information as an RoI (ROI seeded method) discards the large majority of MDT hits outside the RoI from the track search, reducing fake tracks and minimizing processing time and bandwidth for the data transfer to the off-detector processors. In this case the local processors of the MDT are activated by the request from the trigger chambers in the same trigger tower to start a search for hits in the RoI, as supplied by the trigger chambers, measuring the track segment in this chamber layer, see figure 5.5. In a second step the track segments of three MDT layers are used to compute a sagitta, taking into account the alignment among the 3 layers. The resulting estimate for p_T is communicated to the Sector logic (SL), where the final trigger decision is taken. As the frequency of L0-triggers in any given trigger tower is very low (< 100 Hz), little output bandwidth is required for data exchange between trigger logic and MDT readout. The communication between trigger and MDT chambers leads to some additional latency. Timing estimates show that a latency of about 4 μs is required to produce a L1 in the present trigger system, which is well below the foreseen limit of 6 μs (detailed discussion in [34]).

If the RoI from the trigger chambers is not used as a guideline ("self-seeded method"), all hits from all MDT layers will have to be searched for each beam crossing, looking for a sequence of hits in adjacent MDT tubes, pointing to the vertex. A logical diagram for track identification in the end-cap is shown in Figure 5.6. In the EM and EO stations ("Big and Outer Wheel") the direction of high-p_T tracks, having been little deflected by the end-cap toroidal field, shows little variation around the direction to the primary vertex. The search for track segments can be limited to a narrow range around this direction. The necessity of transferring *all* incoming MDT hits to the processors in USA15 leads to bandwidth requirements of about 5–10 Gbit/s per group of 24 tubes. This means that each mezzanine card has to be equipped with an optical link, resulting in ~ 5000 links for the Big and Outer Wheels. Timing estimates for this method lead to a short latency of $< 2.55\,\mu s$, compatible with the present system. A detailed description of the method is given in [35].

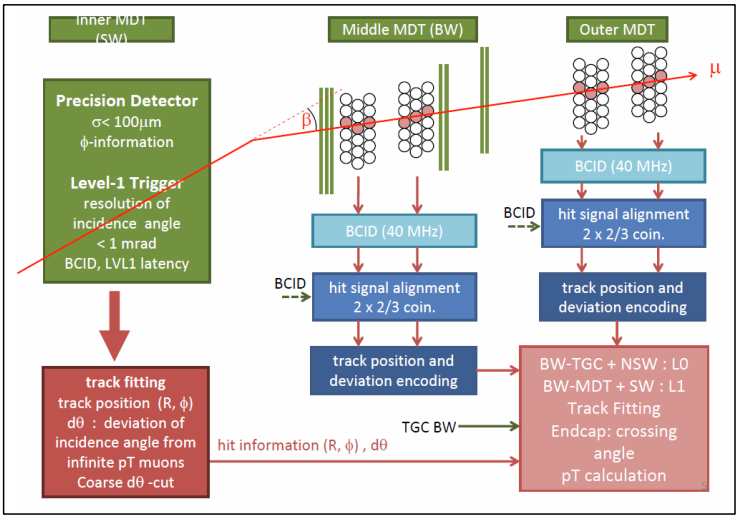


Figure 5.6: Use of MDT tracking information for the determination of p_T, looking for track segments with small deflection angle β, which is the signature for high-p_T tracks. Including the Outer MDT Wheel (OW) in the determination of β results in excellent β-resolution due to the lever arm between BW and OW of about 6 m.- No previous RoI information from the trigger chambers is used in this method.

The schematics of both methods are shown in figure 5.7. In the RoI-seeded method (top) only those regions of the MDT are considered where the trigger chambers have found a high-p_T candidate. In the self-seeded method *all* MDT hits are transferred to the Sector Logic (SL) and processed for track candidates. Track candidates without a match in the trigger chambers are subsequently discarded, removing fake triggers.

The rate of fake tracks found with the self-seeded method strongly depends on the background hit rate. Background hits tend to mask valid hits, while random hit patterns may generate a considerable rate of fake trigger candidates. This comes in addition to other sources of inefficiencies of the MDT, caused e.g. by δ-rays and by particle passage through dead material.

Whether a self-seeded "stand-alone" MDT track trigger (i.e. without any reference to trigger chambers) could help to increase the muon trigger efficiency in detector regions not covered by trigger chambers, would have to be clarified by detailed Monte Carlo simulation, taking inefficiencies and background hit levels into account.

A stand-alone MDT track trigger in the barrel, using only the MDT of the middle and outer layers (BM and BO) would face the additional problem that the lever arm for the angular measurement of the candidate track, i.e. the distance between the two MDT multilayers, is a factor of 2–3 smaller then the one between the corresponding RPC layers (see [24], p. 174 and 193) as well as the fact that the spread of the primary vertices along z is larger in Phase-II, such that the angle θ of the incoming track can not be defined with high accuracy. The inclusion of the Inner MDT layer (BI) would decisively improve the angular measurement of the incoming track, however, as mentioned above, this would require replacement of the BI readout electronics.

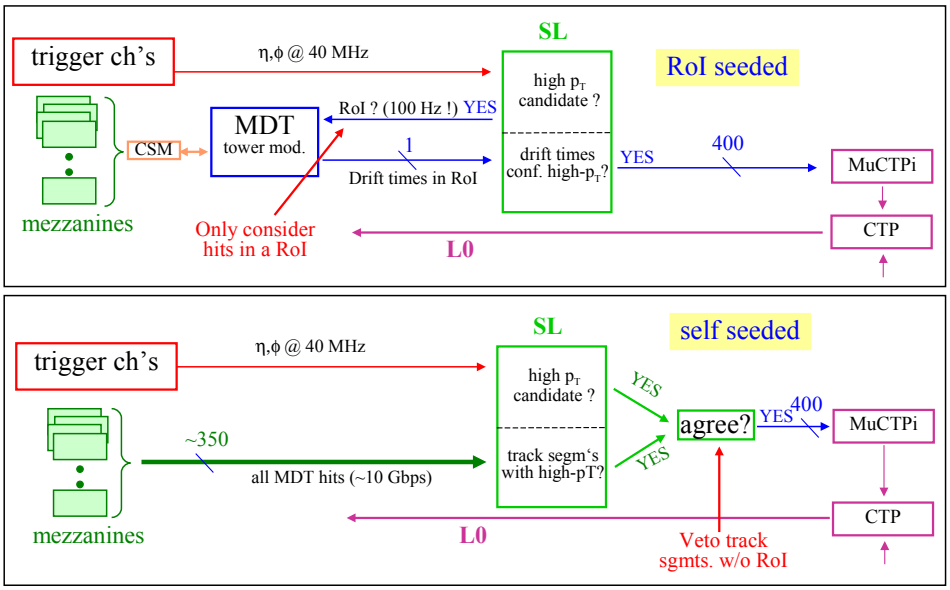


Figure 5.7: Logic flow diagram of the MDT Based L0-trigger for the RoI-seeded (top) and self-seeded approach (bottom). The numbers of fibers before the Sector Logic (SL) are given per trigger tower. There are about 400 trigger towers in the muon spectrometer. Details of the L0/L1 generation are discussed in the text.

6. Upgrade of the Inner Tracking System

The present ATLAS inner detector (ID) has performed outstandingly well in measuring charged particles emerging from proton-proton and heavy ion collisions. As anticipated, it plays a fundamental role in: the identification and reconstruction of electrons, photons, muons, tau leptons; in tagging b-jets and in fully reconstructing certain hadronic decays. With the excellent performance of the LHC, the number of proton-proton interactions in the same bunch crossing (pile-up) is already beyond the design value. The tracker is essential to reconstruct primary vertices and identify the one associated with the hard scattering event of interest. This information is increasingly being used to improve jet energy measurements, to identify isolated particles, and in reconstructing missing transverse energy. A highly performant tracker underpins the entire ATLAS physics program.

To withstand the much harsher radiation and occupancy conditions of the HL-LHC necessitates a complete replacement of the present ID. The new tracker concept presented here is an all-silicon design, based on technologies that are already being prototyped, or are a realistic improvement on existing solutions. The sensors are of finer granularity than the existing tracker, to meet the challenges of very high pile-up and to be able to reconstruct tracks in the core of multi-TeV jets. In addition, the replacement tracker has to be much more radiation hard and the readout links need to provide much greater bandwidth.

6.1 Motivation and requirements

The new tracker design is driven by the performance requirements that cannot be met by the present ID [24]. The current detector consists of 3 layers of pixels, 4 layers of silicon microstrips (SCT) and a straw tube tracker equipped with radiators to generate transition radiation (TRT). An additional innermost layer (IBL) of pixel detectors will be added to the ID during the LS1 [36].

The ID was designed to operate for 10 years at a peak luminosity of 10^{34} cm^{-2}s^{-1}, with an assumed 23 pile-up events per 25 ns bunch crossing, and a level-1 trigger rate of 100 kHz. It cannot survive the planned high luminosity operation nor meet the performance requirements.

Radiation damage The current pixel detector was designed using radiation hard sensor and electronics technologies to withstand 10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$ (1 MeV neutron equivalent per square centimetre). This is estimated to correspond to 400 fb^{-1}, based on measurements of the leakage current increase observed during the current running [37] [38] [39] [40]. The IBL is designed to survive for 850 fb^{-1}. The SCT detector can operate up to a fluence of 2×10^{14} $n_{\mathrm{eq}}/\mathrm{cm}^2$, which is significantly lower than the $\sim 10^{15}$ $n_{\mathrm{eq}}/\mathrm{cm}^2$ required at the HL-LHC.

Bandwidth saturation The frontend electronics of both the pixel and SCT detectors employ zero-suppression and were designed to accommodate occupancies of up to 50 pile-up events, about twice the LHC design value. Limitations in the buffering and the links between the on-module electronics and the read-out driver card (ROD) will lead to inefficiencies in the pixel detector when $0.2 \sim 0.4$ pixel hits per 25 ns bunch crossing are exceeded, which is expected for luminosities of $\sim 3 \times 10^{34}$ cm^{-2}sec^{-1}. A similar saturation effect is expected for the SCT, where limitations in the link between the ABCD front-end chip and the RODs would also imply major data losses above 3×10^{34} cm^{-2}sec^{-1}.

Occupancy With the expected pile-up at the HL-LHC, and in particular in the core of high-p_T jets, the SCT would be unable to resolve close-by particles, and the TRT straws will approach 100% occupancy. Some degradation in the TRT performance has already been observed in the most central heavy-ion collisions.

Even given the design requirements of much greater robustness and redundancy, the extreme conditions at HL-LHC also dictate a more modular concept. For the upgraded tracker, the first two pixel layers must be designed to be replaceable in a standard opening of the experiment without removing the beam-pipe. The outer pixel layers and pixel discs should also be removable in a long opening without disturbing the strips. Furthermore, the end-cap strips should be retractable, to give access to their services and those at the ends of the barrel.

Alongside these mechanical constraints and those dictated by the HL-LHC environment, the guiding principle in designing the Phase-II inner tracker has been the anticipated physics program. Experience with the current detector suggests ways to maintain and improve on the expected ID performance, including the IBL, even in the presence of up to 200 pile-up events. Based on ATLAS measurements from the current LHC running [41], a multiplicity of more than a 1000 tracks per unit of rapidity is expected in the tracker acceptance for this number of pile-up events. In the core of high transverse momentum jets, which are a key part of the HL-LHC program, there will be an even higher local track density. Even if the longitudinal extent of the beam spot has a standard deviation of $\sigma = 75$ mm (longer than the present LHC beam spot but possible with crab cavities or other techniques), it will be a major challenge to associate tracks with the correct primary or secondary vertices at high efficiency. Specifically, the main requirements of the new inner tracker should be to:

- measure the transverse momentum and direction of isolated particles, in particular electrons and muons. The combination of tracker and muon spectrometer information gives the optimum evaluation of the muon transverse momentum;
- reconstruct the vertices of pile-up events and identify the vertex associated with the hard interaction. This information can then be used to improve the evaluation of particle isolation, and the measurements of jet energy and total missing energy;
- identify secondary vertices in b-jets with high efficiency and purity, even in highly boosted jets;
- measure tracks in the core of high energy jets, with good double-track resolution to allow high efficiency and sufficient redundancy to reduce fakes;
- identify the decays of tau leptons, including impact parameter information and in combination with calorimeter measurements;
- reconstruct the tracks associated with converted photons.

6.1.1 Baseline layout

Detailed tracker specifications to meet the above requirements, taking account of engineering issues, have been drawn up [42]. The baseline layout based on these requirements is presented in

Figure 6.1. An all-silicon-detector tracker is proposed, with pixel sensors at the inner radii surrounded by microstrip sensors [43]. In the central region, sensors are arranged in cylinders, with 4 pixel layers followed by 3 short-strip layers then 2 long-strip layers. From current knowledge of the LHC conditions the outer radius of the beam pipe is assumed to be at 33 mm. Given the required modularity discussed above, an inner support tube (IST) will be implemented at a radius of 110 mm, and a pixel support tube (PST) at 345 mm, taking account of the required clearances for service routing. The forward regions will be covered by 6 pixel disks and 7 strip disks. Strip layers are double-sided with axial strip orientation on one side and sensors rotated by 40 mrad on the other side, giving the second coordinate measurement. The tracker is surrounded by a polyethylene moderator to reduce the energies of neutrons, which decreases the 1MeV neutron equivalent silicon damage fluence arising from the flux of neutrons entering from the calorimeters [44] (which for the current ID are partially moderated by the material of the TRT).

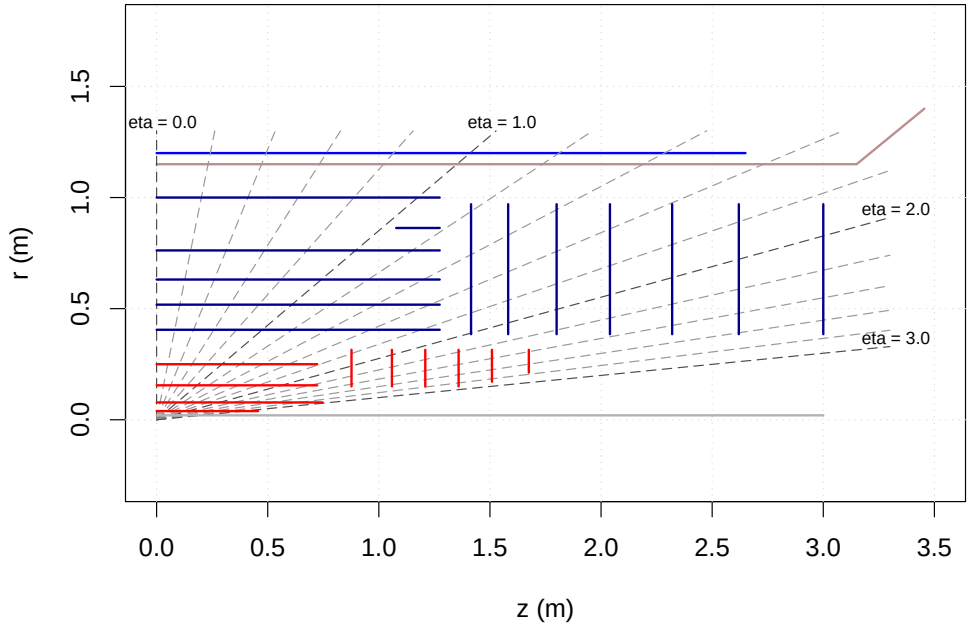


Figure 6.1: The baseline layout of the replacement tracker showing the active areas of silicon detectors arranged on cylinders and disks.

In the optimisation process, gaps have been preserved between subdetector parts to allow for supports, services, and insertion clearances. The resulting sensor areas and channel counts are shown in table 6.6.

The biggest changes to the current inner tracker are replacement of the TRT with 47.8 mm long silicon strips; the pixel system extends out to larger radii; more pixel hits in the forward direction to improve the tracking in this dense region; and smaller pixels and 23.8 mm long inner strips to increase the granularity. The outer active radius is slightly larger, improving momentum resolution. Services have been routed out of the active area as soon as possible, minimising the effects of non-sensitive materials. The layers of silicon are more evenly spaced, especially in the forward region,

Detector:	Silicon area [m^2]	Channels [10^6]
Pixel barrel	5.1	445
Pixel end-cap	3.1	193
Pixel total	**8.2**	**638**
Strip barrel	122	47
Strip end-cap	71	27
Strip total	**193**	**74**

Table 6.6: Inner tracker active area and channel count.

which is beneficial for track reconstruction. The region of fully hermetic coverage from primary-vertex spread in z has been increased, allowing the HL-LHC beam-spot to be increased in length which will increase the average separation between primary vertices.

6.1.2 Expected radiation fluences

The Radiation background simulations for this layout have been performed using FLUKA. A detailed geometry and material description of the proposed new tracker has been implemented. The event generator used in these studies is PYTHIA8 to take advantage of the latest 7-8 TeV ATLAS minimum bias tunings. The new PYTHIA8 also has much improved models for describing the diffractive component of the inelastic interactions. All previous simulations were performed using PHOJET, and the new PYTHIA8 simulations predict fluences up to 15% greater than PHOJET [45], with the largest differences seen in the forward regions. These differences can be taken as a measure of the sensitivity of the radiation background predictions to the choice of inelastic proton-proton event generator.

Two of the most important quantities of interest for understanding radiation damage in the silicon detectors and electronics are the 1 MeV neutron equivalent fluence and the ionising dose. Shown in Figure 6.2 are the 1 MeV neutron-equivalent fluences, normalised to 3000fb^{-1} . Close to the interaction point the backgrounds are dominated by particles coming directly from the proton-proton collisions, but at larger radii, neutrons originating from high energy hadron cascades in the calorimeter material dominate the radiation backgrounds in the inner detector.

The predictions for the maximum 1MeV-neq fluence and ionising dose for 3000fb^{-1} in the pixel system is $1.4 \times 10^{16} cm^{-2}$ and 7.7 MGy at the centre of the innermost barrel layer. For the outer pixel barrels, the maximum fluence and dose are $1.7 \times 10^{15} cm^{-2}$ and 0.9 MGy at the end of layer 3. For the pixel end cap, the maxima are $1.8 \times 10^{15} cm^{-2}$ and 0.9 MGy in the inner region of disk 1. The predictions for the maximum fluence and dose in the strip barrel detectors are $5.3 \times 10^{14} cm^{-2}$ and 216 kGy for the short strips in layer 1, and $2.9 \times 10^{14} cm^{-2}$ and 63 kGy for the long strips in layer 4. For the strip end cap, the maximum fluence and dose are predicted to be $8.1 \times 10^{14} cm^{-2}$ and 288 kGy is seen in the inner regions of disk 7. No safety factors have been applied to these numbers. Further details can be found in [45].

Up to a factor of two difference in 1 MeV new fluences is observed with and without the new polymoderator system [44]. Uncertainties in the simulations have been evaluated by comparing the

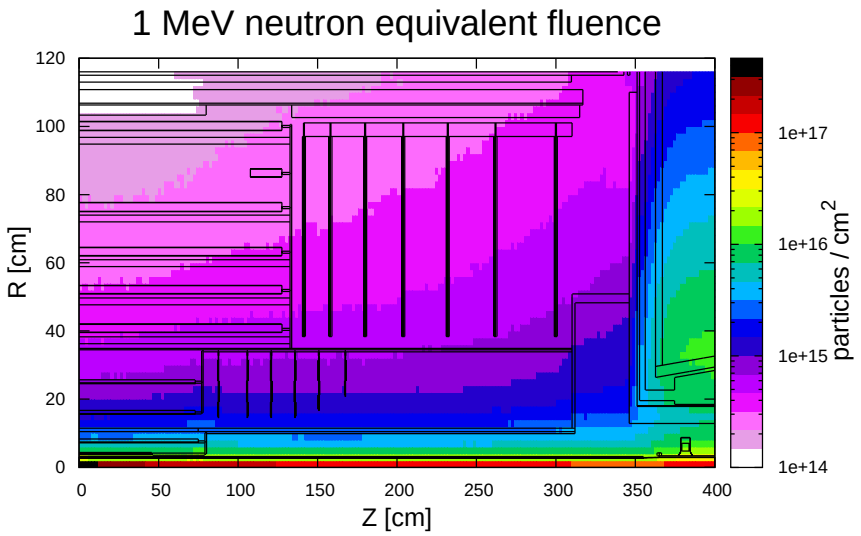


Figure 6.2: *RZ*-map of the 1 MeV neutron equivalent fluence in the Inner Tracker region, normalised to 3000 fb^{-1} of 14 TeV minimum bias events generated using PYTHIA8.

predicted fluences for 7 TeV running with the fluences and doses obtained from Pixel, SCT and RadMon measurements using 2011 data [39]. In general, differences between simulation and data throughout the inner detector volume are found to be well within 50% for both 1 MeV new fluences and ionising doses. The largest discrepancies, a factor of two, are found for the innermost modules of the SCT end-caps.

6.1.3 Trigger information from the tracker

Several possibilities are being considered to bring the tracker information into the trigger chain as early as possible. Details are discussed in 2.5. The baseline solution is a region of interest (ROI) based approach that has no implications on the layout but requires the readout of the pixel and strip detectors to provide sufficient bandwidth to have a Level-0 (Level-1) accept rate of 500 (200) kHz with a latency of 6 (20) μs. However, in contrast, a self-seeded trigger based on directional information in double layers would strongly affect the layout of the strip detector. Although the required technologies for this are being studied, a tracker layout incorporating such a self-seeded trigger architecture is not considered further in this Letter of Intent.

6.2 Phase-II tracker layout and performance

The proposed tracker layout is mainly based on requirements of excellent performance to fully exploit the physics at HL-LHC and in particular to be able to cover the multi-TeV range. In addition to meeting these requirements, the layout must respect constraints from integration, modularity and cost. Figure 6.1 provides details on the proposed tracker including the overall dimensions.

Some characteristics of the performance of this layout compared with the present inner detector are summarised in table 6.7.

Track parameter $\lvert\eta\rvert < 0.5$	Existing ID with IBL no pile-up $\sigma_x(\infty)$	Phase-II tracker 200 events pile-up $\sigma_x(\infty)$
Inverse transverse momentum (q/p_T) [/TeV]	0.3	0.2
Transverse impact parameter (d_0) [μm]	8	8
Longitudinal impact parameter (z_0) [μm]	65	50

Table 6.7: Performance of the existing ID with IBL, and of the Phase-II tracker for transverse momentum and impact parameter resolution. $\sigma_x(\infty)$ refers to σ_x for $p_T \to \infty$, to remove the contribution due to material.

Some of the main considerations leading to this layout are listed below:

Number of hits along a track: Experience with the current detector indicates that to minimise the number of fake tracks (tracks found by the software with hits that do not belong to a single particle) at high pile-up requires tracks have at least 11 hits. Figure 6.3 (a) demonstrates an increase in the ratio of reconstructed to generated tracks with increasing pile-up, which signifies an increase in the number of fake tracks, for reconstructed tracks requiring at least nine hits. This is avoided in the case of requiring at least eleven hits per track, as shown in Figure 6.3 (b).

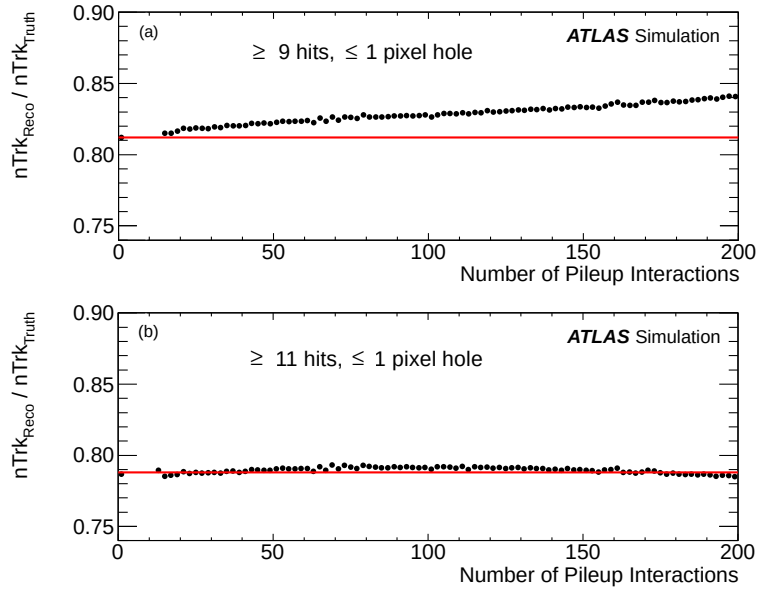


Figure 6.3: Ratio of reconstructed to generated tracks from $t\bar{t}$ at various levels of pile-up for two different track selections: requiring track reconstruction with at least 9 hits per track (a), and with at least 11 hits per track (b).

The 11 hit requirement leads to a close correlation of the radii and lengths of the barrel layers with the z-positions of the disks. Given that each strip module provides two hits, to ensure robust tracking even in the presence of dead modules, the layout aims to have at least 14 silicon hits. Figure 6.4 shows this is effectively achieved to $|\eta| = 2.5$ for primary vertices

within $\pm$150 mm, although at high $|\eta|$ more reliance is placed on the pixel system (up to 7 hits) which provides true space points from single-sided modules. With at least four pixel hits everywhere, robust track seeding in the pixel system is possible as a start for track finding. They constrain the tracks especially in the core of high energy jets. To allow combined muon tracking over the full muon coverage, the pixels extend to $|\eta| = 2.7$. A short ("stub") strip barrel has also been introduced to avoid a gap in hit-coverage between the barrel and end-cap. This also improves momentum resolution in this region, and allows the second end-cap disk to be placed at larger z.

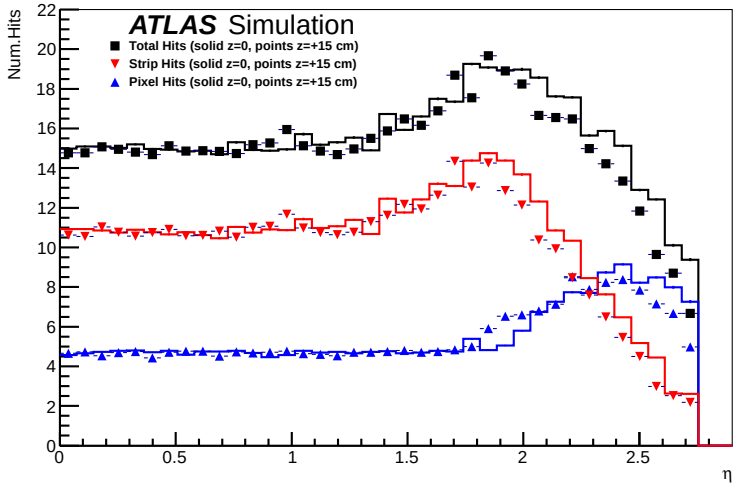


Figure 6.4: Number of hits on muon tracks with $p_T > 5$ GeV as a function of pseudorapidity, η. The line shows the result for tracks originating at the centre of the detector, and the symbols for the tracks originating from $z = 150$ mm.

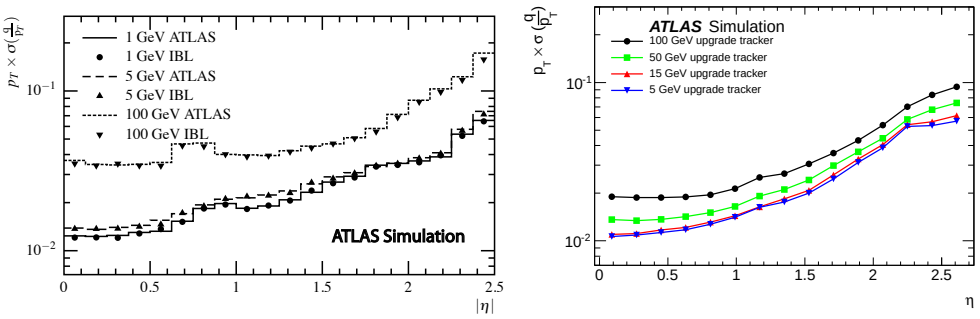


Figure 6.5: p_{T} resolution as a function of pseudorapidity, η, for different momenta from simulation of the current ATLAS ID (with and without IBL) (left) taken from the IBL TDR [36], and from simulation of the upgrade tracker layout using the track selection defined in [46] (right).

Momentum resolution: The layout has been arranged to maximize the length of the trajectory inside the solenoid of the particles and measurements are made at points that minimize the

variations in resolution as a function of the pseudo-rapidity of the track. The outer cylinder of the strips has been placed at the largest radius consistent with the restrictions of the polymoderator, outer support cylinder, detector services, and insertion clearance. Similarly, the last strip disk is at the maximum z coordinate allowed by engineering constraints. Figure 6.5 shows the expected resolution for various values of p_T. The comparison with the current ATLAS Inner Detector (with and without the IBL) shows a significant improvement due to the longer lever arm and more precise hit position. The resolutions have also been found to be largely independent of the level of pile-up.

Two-particle separation and occupancy: The high pile-up and the dense tracks in high p_T jets need a significantly finer granularity than the current detector. This is achieved by using a significantly smaller pixel size in the innermost two layers of 25×150 μm^2 and thinner detectors. This reduces the fraction of clusters with more than one track to a few percent even in the core of a high p_T jet, as can be seen from Figure 6.6. Also the two outer pixel layers have a smaller pitch in z, i.e. 50×250 μm^2, with an option to reduce this further if required for physics performance. In addition the lengths of the strips in the innermost layers and disk sectors are significantly reduced compared to the current detector. The hit occupancies anticipated in this layout for 200 pile-up are shown in Figure 6.7. To obtain a good particle separation for highly dense jets but to minimise gaps in the η coverage, the outermost pixel radius is chosen to be 250 mm. For the three innermost layers the strips are only 23.8 mm long and for the outermost layers they are 47.8 mm.

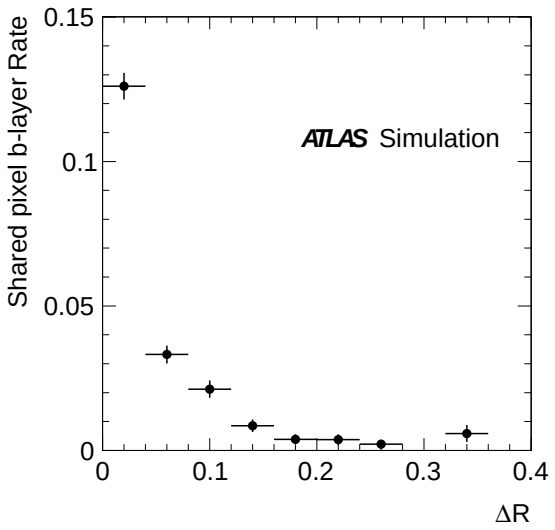


Figure 6.6: The fraction of shared hits on the innermost pixel layer from tracks in high-p_T jets as a function of ΔR from the jet axis, defined as $\Delta R = \sqrt{(\eta_{\text{jet}} - \eta_{\text{track}})^2 + (\phi_{\text{jet}} - \phi_{\text{track}})^2}$.

Inefficiencies due to interactions in the material: It is important to minimise overall tracker material to limit the tracking inefficiency for hadrons by losses due to hadronic interactions and effects due to Bremsstrahlung for electrons. In addition, reduced multiple scattering leads

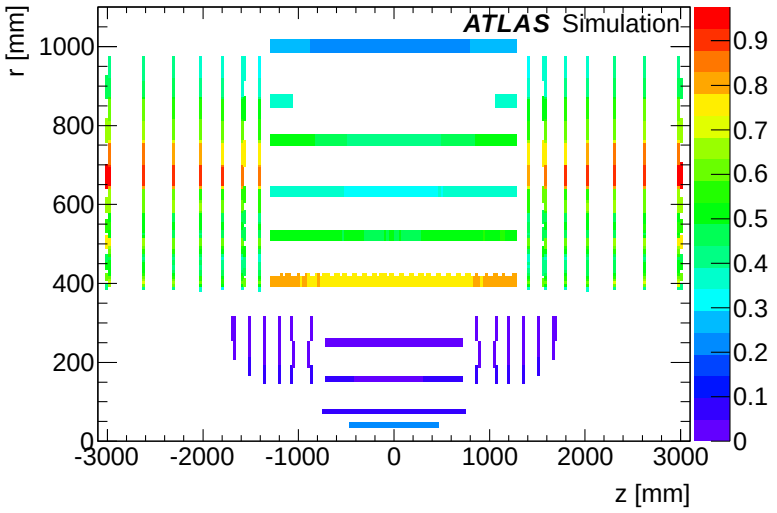


Figure 6.7: Channel occupancies (in percent) with 200 pile-up events.

to a better precision, less photon conversions reduce tracking confusion and provide a better photon identification and the reduced number of secondaries from interactions leads to increased detector occupancy. The expected material distribution shown in Figure 6.8 represents a major improvement with respect to the current ID; the current ID (including the IBL) contributes $> 1.2\ X_0$ for all regions $|\eta| > 1$ [36], while the new tracker remains below $0.7\ X_0$ up to $|\eta| = 2.7$, excepting a few small regions. To improve tracking performance in general and particularly to improve the electron efficiency in the forward region, the amount of services material inside the tracking volume has been minimised. This is achieved through services from the pixel barrels going inwards to small radius, then out along z, which removes them as soon as possible from the tracking region. It also allows the pixel detector to be removed for repair without disturbing the strips.

Consideration has also been given to avoiding large distances between successive hits on a track. This improves the pattern recognition, reducing fakes due to multiple-scattering induced effects. Those fake tracks which are produced, constituting less than 10^{-3} of all reconstructed tracks, are located mainly in the region $1.5 < |\eta| < 2.5$, as discussed in [46].

Tracking efficiency: Efficiencies for tracking muons, pions and electrons in the presence of 140 pile-up are shown in Figure 6.9. The longer barrel with respect to the current ID moves the difficult transition region to higher $|\eta|$. The very high electron efficiency is in part due to the use of bremsstrahlung refitting in the reconstruction.

Primary and secondary vertexing: The tracker must be able to reconstruct the primary vertex for the hard scatter of interest in the event. In Figure 6.10 (right), the number of reconstructed primary vertex candidates is displayed as a function of the number of pile-up interactions. Whereas a significant fraction have too few tracks for a vertex to be reconstructed, the figure

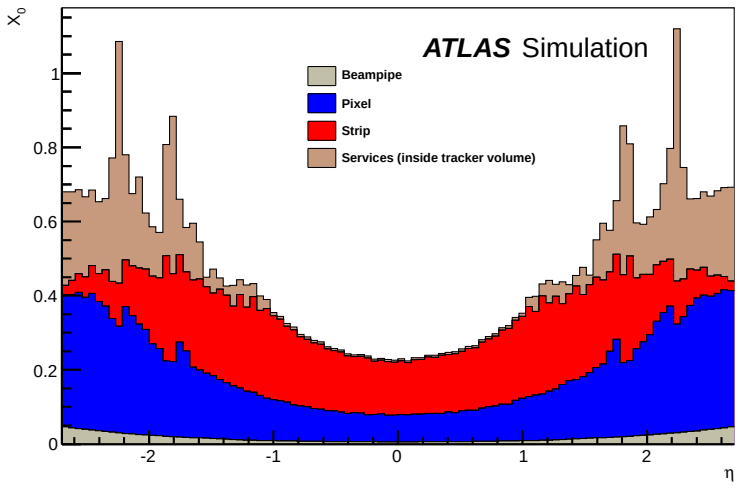


Figure 6.8: The material in X_0 as a function of η for the Phase-II tracker layout.

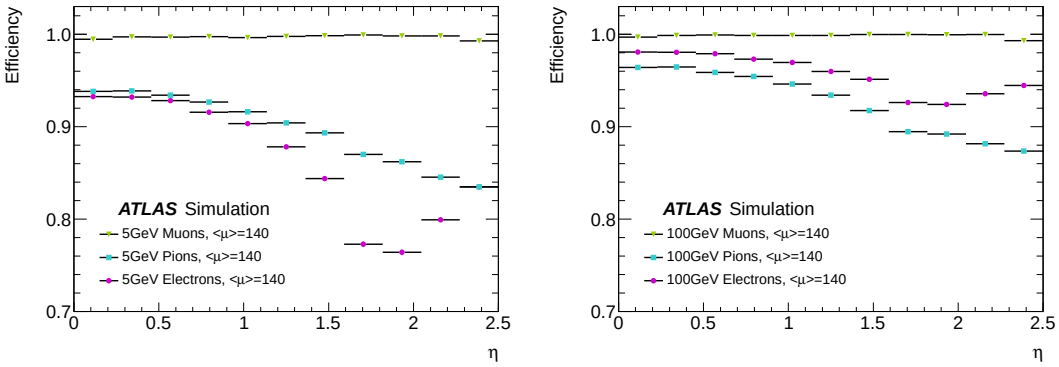


Figure 6.9: Efficiencies as a function of pseudorapidity η for $p_T = 5$ GeV (left) and 100 GeV (right) muons, pions and electrons at pile-up of 140.

shows that the fraction of vertices that are reconstructed is almost independent of the number of pile - up events.

The performance of b-tagging, in particular for high energy jets, depends critically on the rate of fake tracks, and on the two-track resolution. The innermost layer is therefore as close as possible to the beam-pipe at a radius of 39 mm from the beam. Also the smaller pixel sizes help in providing a better resolution and granularity for achieving good fake rejection at high pile-up and in the dense environment of multi-TeV jets. Using current algorithms, the light jet rejection of the new tracker at high pile-up is found (Figure 6.10, left) to be as good as that for the current ID (including the IBL) at zero pile-up. Tau lepton identification also relies on excellent impact parameter resolution, and on good two-track resolution in the case of 3-prong decays or higher multiplicity decays.

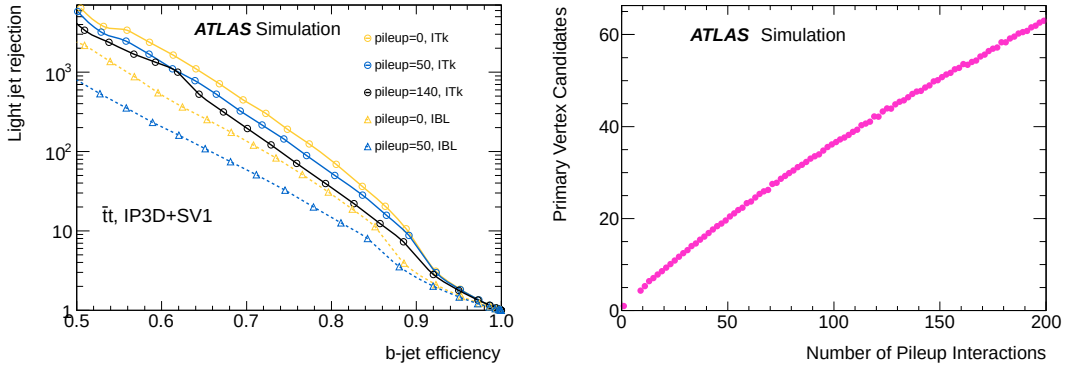


Figure 6.10: Performance of b-tagging in $t\bar{t}$ events, for a range of pile-up levels for the proposed Phase-II Tracker layout in comparison with ID+IBL (left). On the right, the number of reconstructed primary vertex candidates as a function of the number of pile-up interactions.

6.2.1 Alternative layouts

A few alternative layouts are considered, which modify the pixel part of the inner tracker. These are the conical layout, which smoothes the transition between barrel and end-cap with a cone shaped structure at the end of the barrel; the 5-layer pixel layout, which simply adds an extra pixel barrel; and the alpine layout, with a novel arrangement of sensors on structures. These layouts are considered as options requiring more detailed studies and development work.

Conical layout

The conical layout [47] is based on a bent integrated stave, with a flat middle section and bent ends. This concept reduces the material in the forward region, because the end of stave cards are moved to higher η. The crossing angle for particles incident on the stave is closer to perpendicular in the conical region, further reducing the material traversed by a track.

One possible layout is shown in Figure 6.11. The outer barrel layers are at the same radius as the default layout, but they are shortened, taking advantage of the fact that the modified conical section covers the gap. As a result, the barrel pixel silicon area is reduced from the default 5.1 m^2 to 4.6 m^2. The resulting material budgets are depicted in Figure 6.12 and shows some reduction at $|\eta| > 1.5$.

Other layouts with conical structures use a larger radius of the outermost pixel layer without introducing a significant gap to the disks, and the end cap outer radius can be reduced to match the barrel, which has additional advantages for mechanical assembly and integration.

Five pixel layers

The impact of modifying the layout by adding a fifth pixel layer. This allows, e.g., a more robust pattern recognition seeded in the pixel detector alone, and a better two-particle separation in high p_T - jets.

A possible layout is shown in Figure 6.13. It assumes shorter outer pixel barrel layers relative to the baseline layout [48] to reduce costs. The strip stub barrel is removed, since the extra pixel

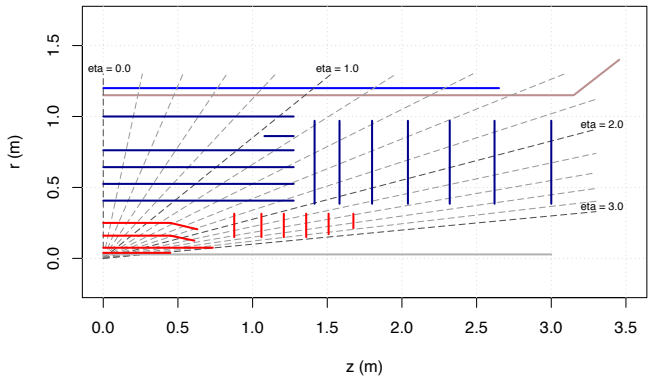


Figure 6.11: An example layout using the concial geometry for the pixel barrel, and reducing the length of the outer barrel layers.

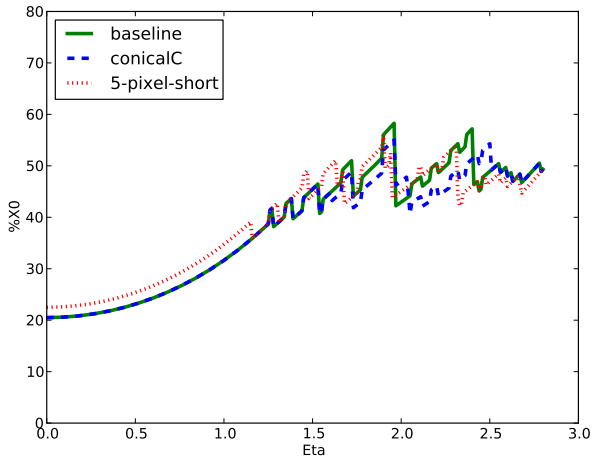


Figure 6.12: The estimated material as a function of η for two conical layouts compared to the baseline layout.

layer brings one extra hit to compensate the loss of two strip hits in the relevant η range. Relative to the baseline layout, the active pixel area in the barrel is increased to 6.7 m^2

Alpine layout

The idea of the alpine stave is to place the modules at high η with an increasing inclination angle [49], so as to approximate an ideal geometry in which all the silicon sensors are perpendicular to the incident particles ('alpine' structure). As shown in Figure 6.14, the innermost layer is not changed, so as to remain as close to the beam line as possible. Relative to the baseline layout, the axial part of the other three layers becomes shorter, the alpine structures starts at about $\mid \eta \mid \sim 2.25$. The modified geometry removes the need for end cap disks, and reduces substantially the area of

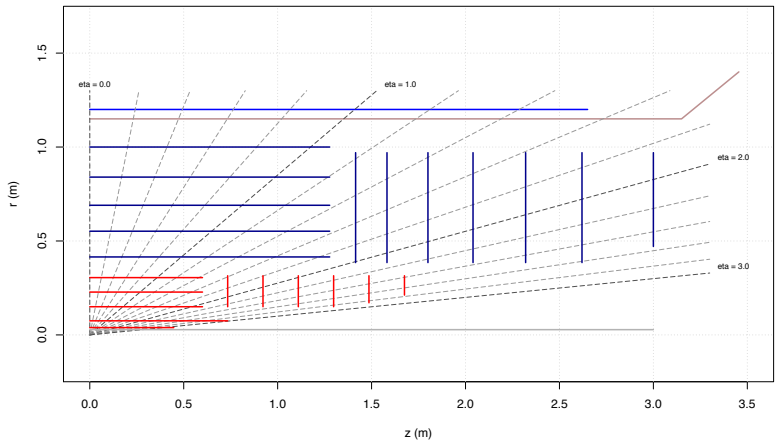


Figure 6.13: The alternative layout with five pixel barrel layers.

silicon required to 4.6 m^2. A prototype of a small section of an alpine stave is shown in Figure 6.15 which has been used to demonstrate the low material and required thermal performance are achievable.

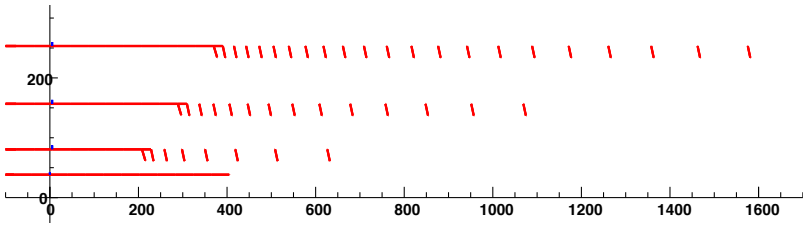


Figure 6.14: The pixel sensors in the alpine layout. Whereas the barrel region is not changed, the default end caps are replaced by modules placed perpendicular to the long stave direction.

Other concepts under consideration include a greatly extended forward region either with more disks, or, for the alpine layout, a much larger number of inclined modules. Preliminary physics studies indicated that it may be advantageous to extend the pixel coverage $|\eta| \sim 4$ and this option is also being considered. For all the alternative layouts, any final choices will require detailed simulation to compare their performance with that of the baseline, as well as evaluation of differences in complexity of engineering and installation, along with estimation of their costs.

6.3 The pixel system

The Phase-II pixel system presented in this LoI is largely based on existing solutions. Before the start of production of the detector there will be several more years of R&D, addressing the requirements of HL-LHC physics, in particular finer granularity, higher bandwidth and reduced material. This effort should allow the use of more performant technologies as they become available.

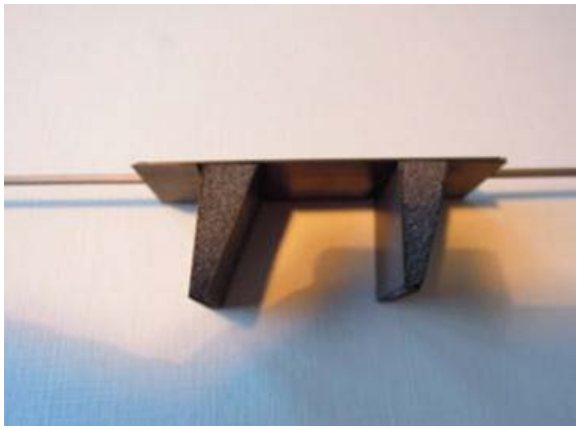


Figure 6.15: Prototype of an alpine stave

As a baseline, the outer pixel barrels and the disks will use the already existing planar technology, which is adequate to meet the radiation and rate requirements in this region. Important experience of this technology will be gained from the IBL [36]. For the two inner layers it is assumed that on-going developments will allow new electronics and sensors to be used. This technology division matching the physical division of the layout is advantageous because the outer layers and disks, which comprise most of the silicon area, will require a relatively long construction time and therefore need mature technology components relatively soon for large scale production. The inner layers in contrast are a relatively small system that can start construction later with a slower ramp-up of production. Since these layers are insertable they can be integrated last. Optionally, if the new technology needed for inner layers is available soon enough and is affordable, it could also be used for the outer layers.

Although the LoI baseline is to replace the full tracker in LHC Long Shutdown 3, the staging of pixel elements by a replacement of the present pixel system during the Long Shutdown 2 is considered as an option [2]. This would allow ATLAS to benefit earlier from an improved IBL technology, which may improve performance and/or recover from degradation and failures of the present pixel detector. This option was already discussed in the LoI for Phase-I [2] and will be revisited in 2013.

6.3.1 Pixel sensor choices

The choice of sensors depends mainly on the requirement that the detector has to withstand an expected fluence of 3000 fb^{-1}. This is particularly challenging for the innermost layer, which after the high-luminosity running is exposed to an estimated fluence of 1.4×10^{16} $n_{\mathrm{eq}}/\mathrm{cm}^2$ before allowing for any safety factor. For the outer pixels, the expected maximum fluence is 1.7×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$, and the sensor technology is well established. To enhance radiation tolerance, minimize material and reduce the cluster sizes, a sensor thickness of less than 150 μm is assumed.

At this stage three possible sensor types are considered for the inner pixels:

- planar sensors [50, 51],
- 3D sensors [52],

- diamond sensors [53].

A considerable R&D effort on all three options is under way. For the IBL [36], which is designed to work for an integrated luminosity of 850 fb^{-1} and will be installed into ATLAS in the 2013/14, both planar n-in-n sensors and 3D sensors are being used. Diamond detectors will be used for beam monitoring. The IBL will provide substantial operating experience which will be important for the Phase-II tracker. Planar sensors have the advantage of long experience of high quality mass production by multiple vendors, low costs and high yields. The 3D sensors require a relatively low depletion voltage even after high irradiation dose but are currently more expensive. Diamond detectors, promise lower capacitance and a high radiation hardness in terms of dark current and so do not require cooling, but are still expensive at the moment.

Although no baseline sensor choice has been made at this stage, planar sensors (n-in-n or n-in-p) are assumed for this LoI. Their radiation hardness has been proven and mass production capabilities are well established with multiple international vendors.

For the innermost two pixel layers n-in-n sensors are assumed. From test-beam results they have been shown to be operational at fluences of more than 10^{16} $n_{\mathrm{eq}}/\mathrm{cm}^2$. After 2×10^{16} $n_{\mathrm{eq}}/\mathrm{cm}^2$, a bias voltage of 1.5 kV is required for a signal of 6000 electrons. This is well above the threshold of the front-end electronics and a large factor above the noise. This implies a hit efficiency of close to 100% after being irradiated with the expected HL-LHC dose. The hit efficiencies for different doses as a function of sensor bias voltage are shown in Figure 6.16. The large bias voltage puts constraints on the power lines and cooling. However, this voltage is not significantly larger than the one needed for the IBL. These sensors are already being produced with thickness of 150 μm and thinner sensors are under development, which should be considered as an option. For this LoI a thickness of 150 μm is assumed. Recent studies [54] have shown that the inactive edge can be reduced to 100 μm or even less.

For the two outer layers, n-in-p planar sensors are assumed for the LoI which are somewhat cheaper, and guarantee a larger number of available vendors. A sufficient radiation tolerance has been demonstrated for test sensors for 5×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$ and 1kV bias voltage, sufficient for the conditions expected at the higher radii. Using longer pixels at the edge, the inactive region can be limited to 100 μm.

A final decision on the sensor technology will only be taken after evaluating the results of future R&D efforts. These are mainly directed towards reducing the voltage for planar sensors, reducing the thickness of all sensor technologies and reducing the cost, especially for 3D and diamond sensors. ATLAS groups are also involved in the R&D effort on HV-CMOS technology [55]. Here particularly the ability of high speed readout in a high radiation environment has to be shown.

6.3.2 Pixel FE electronics

The FE-I4 chip [56], presently being used for the IBL, forms the basis of the LoI pixel design. A chip with similar characteristics (size, power, I/O) as the FE-I4, but with smaller pixel size is assumed for the inner two layers. R&D aimed at such a chip is described below. The development of the FE-I4 was dual purpose - to enable a new inner layer in the near term (IBL), and to demonstrate low cost, efficient module production for large area pixel layers in the longer term future. It

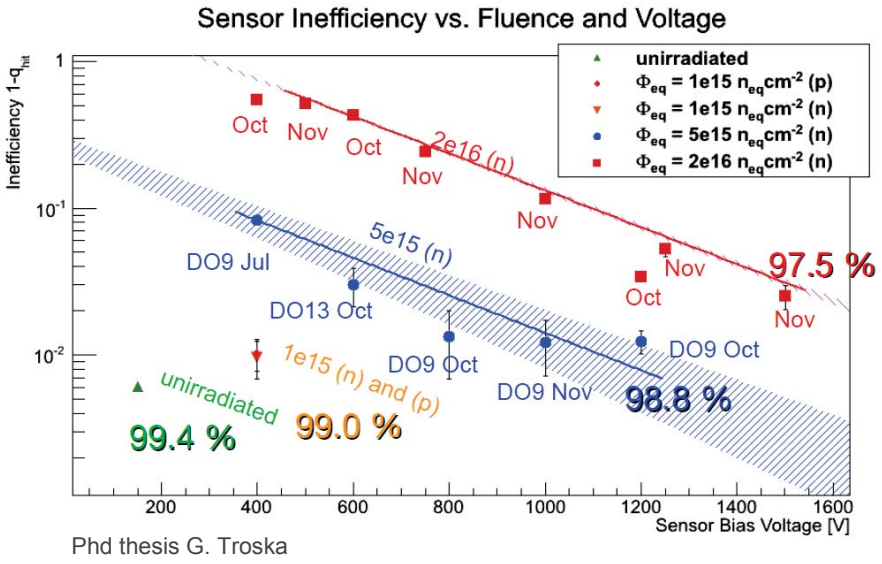


Figure 6.16: The hit inefficiency is shown for planar n+ - in - n sensors as a function of the bias voltage. The red and blue lines interpolate the results of measurements of sensors suffering fluences of 2×10^{16} $n_{\mathrm{eq}}/\mathrm{cm}^2$ and 5×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$, respectively. The percentage numbers indicate the hit efficiency for the different fluences achieved at the highest applied voltage.

was recognized early on that the IBL rate, radiation, and granularity requirements were essentially the same as those for Phase-II outer layers. The requirement of low cost is largely driven by the Phase-II detector, since the IBL silicon area is very small. This motivated making the FE-I4 chip as large as possible, since the cost of module bump bonding scales with number of chips, not with surface area. However, the FE-I4 version B chip being used for the IBL production could not be used as-is for the Phase-II outer layers. A version C would be needed to add compatibility with multi-chip modules and with DAQ and trigger protocols still to be defined. The 50×250 $\mu\mathrm{m}^2$ FE-I4 pixel size is assumed for the FE-I4-C outer layers.

R&D on a new generation of pixel readout chip beyond FE-I4 is in progress along two parallel directions: 3D electronics integration and 65 nm feature size conventional CMOS.

A pixel granularity of 25×150 $\mu\mathrm{m}^2$ is being targeted in order to have high resolution on the track impact parameters, avoid cluster merging for better two-track resolution and reduce occupancy for high efficiency and low fake rates. This granularity is assumed for the inner pixels of the baseline layout. This small pixel size must be coupled with thinner than 150 μm sensors in order to take full advantage of the granularity, because thicker sensors will produce too large clusters.

A charge measurement in every pixel with 8-bit resolution is being targeted. The present ATLAS pixel detector has such a capability, but in the FE-I4 chip the dynamic range has been

reduced to 4-bits in order to be able to fit the required logic in the desired pixel size. The higher logic density of the 65 nm process is expected to allow 8-bit resolution to be restored in spite of the smaller pixel size.

The front-end electronics are the main contributors to the power consumption of the pixel detector. To limit the cooling requirement, the total power including sensor bias should not exceed 450 mW/cm^2, as in the IBL.

R&D is ongoing to find the best solution to meet these requirements. ATLAS groups have made first prototyping efforts towards a new front end chip. The choice of 25×150 μm^2 is also motivated by prototyping in 65 nm technology, where an analog front-end chip [57] has been produced for a small matrix of 16×32 pixels (Figure 6.17) using staggered bump bonds. The analog noise below 100 electrons and threshold dispersion of 350 electrons for a mean threshold of 2700 electrons (Figure 6.18) are very satisfactory results. The 65 nm node is the smallest feature process still using silicon dioxide gate dielectric, which is intrinsically radiation hard due to quantum tunneling. By placing bump bond pads staggered in the z coordinate, a bump pitch of 50 μm is maintained, while allowing a Rϕ pitch of 25 μm.

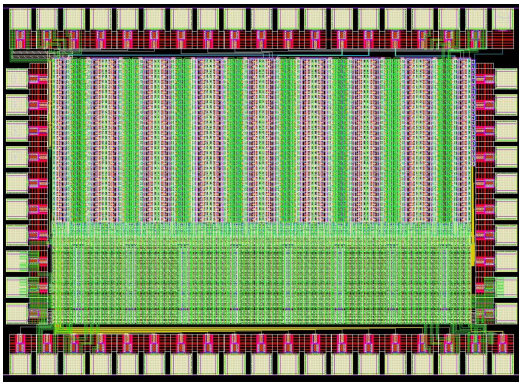


Figure 6.17: Test chip with an array of 16×32 pixels in 65 nm CMOS technology with pixel size of 25×125 μm^2. Scaling to realistic size of digital logic and memory gives the final pixel size estimation of 25×150 μm^2.

Another option for reducing the pixel size is to use 3D integration by having multiple tiers of electronics within the pixel footprints [58, 59]. The electrical interconnections between the tiers, to the sensors and with the external wire bond pads are made by Through Silicon Vias (TSV) and metal to metal surface interconnections. A Test chip with separate analog and digital tiers, FE-TC4-P1 (Figure 6.19), was produced with 130 nm CMOS technology. Pixels of size 50×166 μm^2 contain the full analog tier similar to FE-I4 and two variants of digital tiers. The analog tiers show good performance after been thinned to 10 μm thickness of Si bulk. The radiation hardness of the analog and digital parts has been demonstrated in a 24 GeV proton beam up to a dose of 500 Mrad. The communication between the analog and digital tiers has been shown to work without any loss of performance. The noise and threshold dispersion using readout by both tiers are the same as for a classical 2D readout. The next prototype of this 3D chip will have a pixel size of 50×125 μm^2.

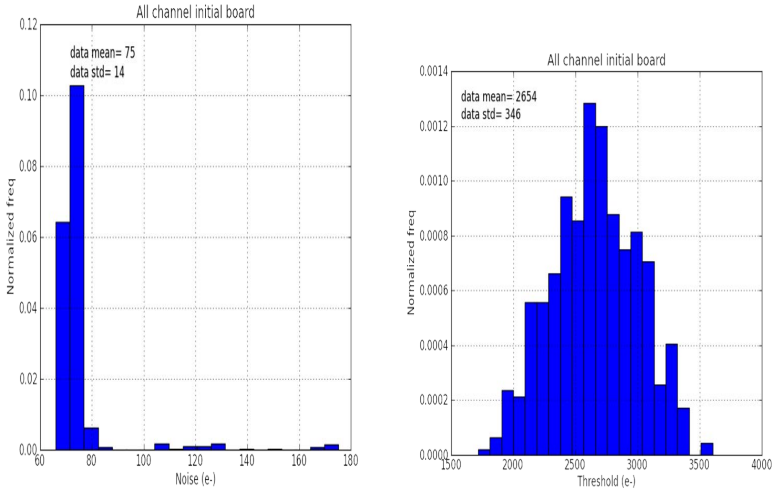


Figure 6.18: Noise and threshold dispersions in the test chip in 65 nm CMOS technology.

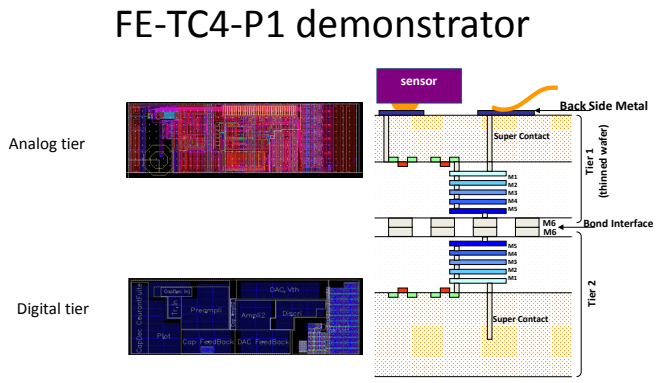


Figure 6.19: 3D electronics test chip FE-TC4-P1 with two tiers in 130 nm CMOS technology with TSVs and surface bonding pads. Analog tier on the top of the figure contains amplifier and discriminator and is connected with digital tier by surface metal Cu-Cu bonds. Analog tier is thinned to 10 μm thickness with amplifiers connected to sensor bump-bonding metallisation pads by 2 μm diameter TSVs. Wire bond pads (on the top left corner of the figure) are also connected with input/output drivers via TSVs.

6.3.3 Pixel modules

The basic electrical unit of the pixel detector is a module. There will be three types of modules: 2-chip-modules (2 chips bump bonded to a sensor, 4×2 cm^2) for the innermost layer to accommodate the limited space, quad-modules (4 chips bump bonded to a sensor, 4×4 cm^2) for the outer layers and hex-modules (6 chips bump bonded to a sensor, 6×4 cm^2) to be used in the disks. Details of the module dimensions are given in Table 6.8. All chips in the module will be controlled in parallel

(shared clock and command inputs), but the module may have one or several dedicated data outputs. The multi-chip output data multiplexing functionality may be incorporated into the next generation pixel chips being developed and in revision C of the FE-I4, or it could be implemented with a separate chip on the module as is done in current detectors. The key step in the classical hybrid module integration is the flip-chip bump bonding process which connects each pixel sensor with the corresponding FE readout cell. Bump bonding requirements for the Phase-II layout will be similar to those of the IBL currently in production:

- bump pitch 50 μm;
- staggering of bump pads to preserve 50 μm pitch (which has been demonstrated for 25 $\times$ 125 μm^2 pixels);
- defect rate $< 10^{-4}$.

The relevant parameters for barrel and disk modules are listed in table 6.8. The fractions of inactive regions are kept low by having longer pixels at the edge and in the regions between chips, and by minimising the edge region while still preventing voltage break down. External barrel pixel

parameter	layer 1	layer 2	layers 3 + 4	disks
chips (ϕ x z)	FE-5 2x1	FE-5 2x2	FE-I4 2x2	FE-I4 2x3
pixel size (μm^2)	25 x 150	25 x150	50x 250	50x 250
nb of pixels in ϕ	672	1348 +12 ganged	672 + 6 ganged	672 + 6 ganged
nb. of double columns in z	68	68	40	60
length of sensor at gap (μm)	300	300	450	450
distance to module edge (μm)	150	150	500	500
distance active to cut edge (μm)	100	100	100	100
active size (mm^2)	16.8 x 41.1	34.0 x 41.1	33.9 x 40.4	33.9x60.8
physical size (mm^2)	18.8 x 41.3	38.0 x 41.3	38.0 x 41.1	38.0x61.3
power (W)	0.9	1.8	1.8	2.7

Table 6.8: Basic parameters of the pixel modules. Note that for disks in addition to the hex-modules also the quad-modules of barrel layers 3+4 are used.

layers will be made of quad-modules shown in Figure 6.20. Disks are made of mostly hex-modules in high radii rings to minimize the number of data and power lines, but at low radii rings also quad-chips modules are used to meet the requirements for the inner radii.

ATLAS groups are pursuing an extensive prototype program on quad-modules [60], as well as high volume bump bonding using FE-I4-B wafers with several vendors. More than 70 wafers have been added to the IBL production order for such R&D efforts world-wide (the total order of 96 FE-I4B wafers contained more pixels than the present ATLAS and CMS detectors combined). A first prototype quad-module with the FE-I4-A chip is shown in Fig 6.21.

Four side buttable dual and quad modules with via last TSVs (when micro holes are produced after transistors and in a different factory) are under development in the ATLAS pixel institutes.

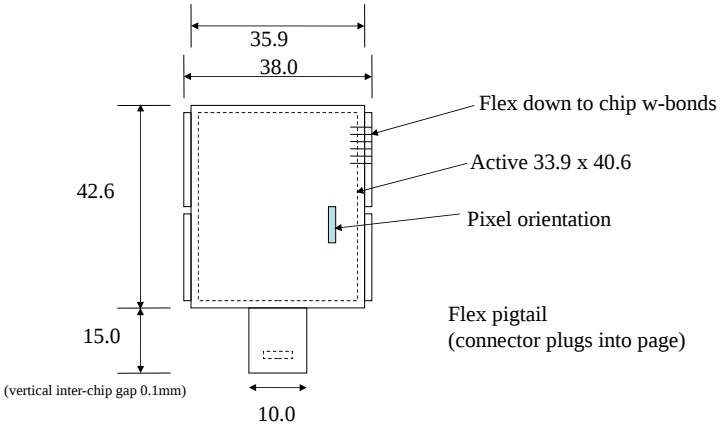


Figure 6.20: Drawing of the quad-module with 4 FE-I4 chips

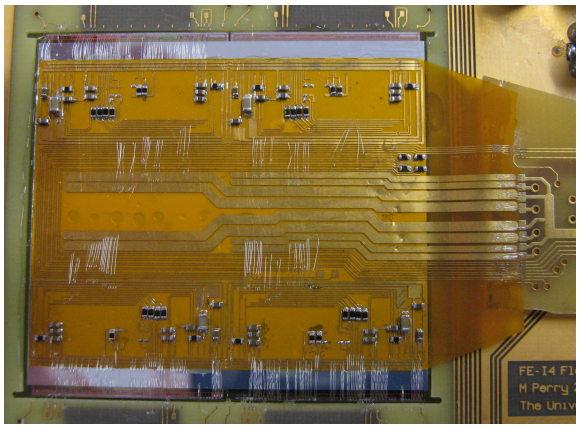


Figure 6.21: Prototype of the quad-module with four FE-I4 chips.

Such modules will relax the constraints on the mounting of the modules on the stave and potentially reduce the material budget.

6.3.4 Pixel barrel supports

An I-beam solution is considered as the baseline for the innermost two layers. Figure 6.22 shows the concept of the I-beam staves. At the outer and inner radius, thin carbon-fibre laminates provide stiffness and a mechanical support for the pixel modules. These are connected by a light weight carbon element, made possible by the different module sizes in the two layers. Because of its inherent stiffness, the I-beam solution does not need an external support structure. The total material is 1.21% X_0, while the bare stave contribution to the material budget is only 0.43% X_0 per layer at normal incidence (see Table 6.10). In addition to the low material budget, an I-beam solution allows fast replacement. In particular, an important further constraint for the innermost two layers is that they can be mounted as “clam shells” for extraction without breaking the LHC vacuum.

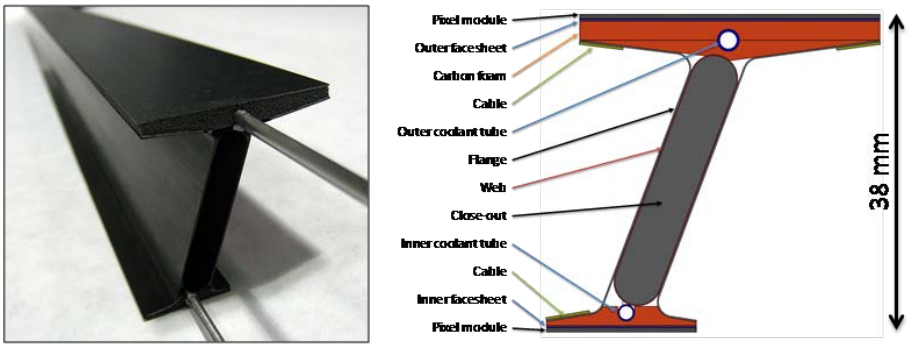


Figure 6.22: I-beam concept prototype and cross-section diagram for two inner pixel layers

The core of the mechanical support will be made of extremely light weight carbon foam, which still provides a very high thermal conductivity. The heat from the electronics is cooled with CO_2 evaporating in very thin titanium pipes. Two cooling tubes serve both layers in a common carbon structure. A circular arrangement of I-beams creates the two innermost pixel layers as shown in Figure 6.23. It can easily accommodate the clam shell installation requirement.

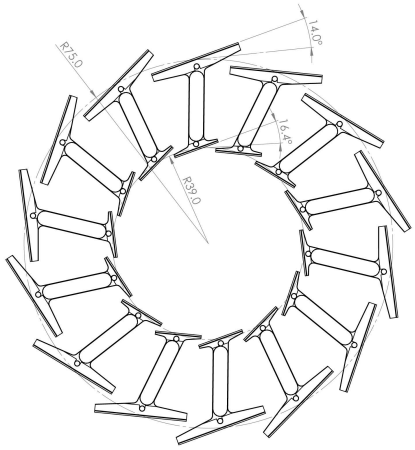


Figure 6.23: End view illustrating layout of I-beams for two inner pixel layers.

A 1000 mm I-beam prototype was built with carbon foam, carbon face-sheets and stainless steel tubes. By changing stainless steel tubes into titanium tubes and optimizing the glue, the material can decreased from actual 0.87% X_0 to 0.43% X_0. Several tests on the mechanical stability such as bending and vibrational properties showed an excellent performance.

The baseline for the outer two barrel layers, and an option for the innermost layers, is an improvement of the IBL stave [36]. These also use carbon foam as core material and embedded hard bonded titanium cooling pipes. The stiffness of the structure is provided by an omega shaped carbon fibre laminate, which is bonded to the foam. The IBL bare staves have a material budget

of 0.62% X_0. Figure 6.24 shows the components of the stave and a stave that will go into the IBL. Improvements could be made in the glue attaching the modules to the staves. The improved

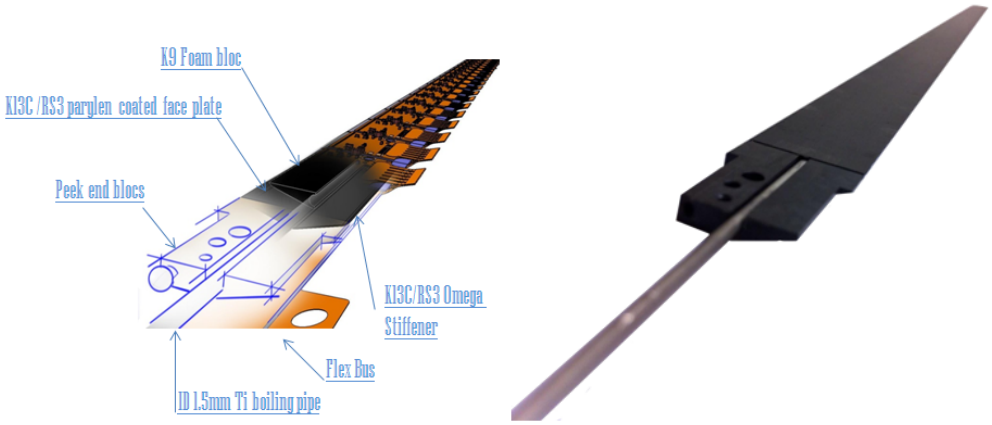


Figure 6.24: Sketch of the various components of the IBL stave (left) and an actual stave ready for module loading (right)

IBL staves benefit from considerable development work in local mechanics, cooling, connectivity, module loading, electrical low and high voltages and signal connectivity.

A tilt angle of -14 degrees was chosen to compensate the cluster size increase due to the Lorentz angle shift in planar sensors, to assure a minimum overlap of 40 pixels between staves in the $R\phi$ direction for alignment purposes and to guarantee sufficient clearance between two adjacent staves. The details of the overlap region are shown in Figure 6.25.

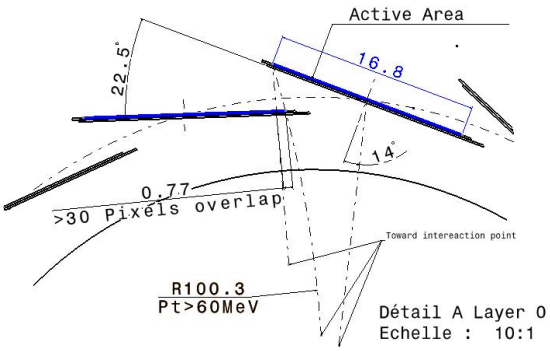


Figure 6.25: Details of the overlap region of the pixel barrel with IBL stave option.

An option being considered for the outer pixel layers are flat staves, similar to the staves in the strip region described in section 6.4.5, but with smaller dimensions. The module placement on such staves is similar to the IBL module placement. On prototypes of the outer pixel flat stave, good cooling quality and stiffness have already been demonstrated.

I-beam staves are very stiff and designed to be end-supported. The IBL-like and flat staves need support shells. I-beam and flat staves were already realised as prototypes, while IBL-like

stave is the extrapolation from the extensive development work an testing of those used for the IBL.

The main parameters of the pixel barrel detectors and the number of staves and modules are given in the Table 6.9.

Detector:	Radial position [mm]	Half Length [mm]	Tilt angle degrees	Staves	Modules/stave
Layer 1	39	456.5	−14	16	22
Layer 2	78	747.0	−14	16	36
Layer 3	155	722.8	−14	32	35
Layer 4	250	722.8	−14	52	35

Table 6.9: The main parameters of the pixel barrel detector.

In order to minimize the material budget and simplify module placement, tightly neighbouring module placement with gaps of 200 μm, as for the IBL, is chosen as the baseline. This leads to a total inactive fraction of 0.97% per stave. Options with shingled modules and two sided modules are studied as alternatives which allow the inactive area to be reduced. In Table 6.10 the global break-down of radiation lengths X_0 of a stave in sensors, electronics, module and mechanics is given. Cables with flexes give the second largest contribution after the bare staves.

Component	I-beam (% X_0) (inner baseline)	IBL-like (% X_0) (innermost option)	IBL-like stave (% X_0) (outer baseline)	Flat stave (% X_0) (outer option)
Sensor 150um	0.16	0.16	0.16	0.16
FE chip 100 um	0.11	0.11	0.11	0.11
Flex	0.19	0.19	0.19	0.19
Cables	0.29	0.29	0.16	0.16
Glue	0.03	0.03	0.03	0.03
Bare stave	0.43	0.42	0.43	0.64
Support	0.00	0.27	0.25	0.25
Total	1.21	1.47	1.33	1.54

Table 6.10: Global break-down of material X_0 of a stave: sensors, electronics, modules, mechanics.

6.3.5 Pixel disks

The forward pixel detector layout consists of six disks on each side of the detector, with each disk supporting 2-3 rings of rectangular multi-chip modules.

The main parameters of the pixel disks and the number of modules per disk are given in Table 6.11. A possible arrangement of modules on a disk [61] is shown in Figure 6.26. The reverse side of the disk has modules in the gaps, providing hermetic azimuthal coverage and near-hermetic radial coverage.

Detector:	Rin [mm]	Rout [mm]	z position [mm]	Modules/type Ring out	Modules/type Ring middle	Modules/type Ring in
Disk 1	150.1	315.0	877	60/hex	48/hex	36/quad
Disk 2	150.1	315.0	1059	60/hex	48/hex	36/quad
Disk 3	150.1	315.0	1209	60/hex	48/hex	36/quad
Disk 4	150.1	315.0	1359	60/hex	48/hex	36/quad
Disk 5	170.6	315.0	1509	60/hex	48/quad	40/quad
Disk 6	212.4	315.0	1675	60/hex	48/quad	-

Table 6.11: The main parameters of the pixel disks. Quad-modules are made out of 2×2 chips bump bonded to a sensor, while the hex-modules are made of 2×3 chips bump bonded to a sensor.

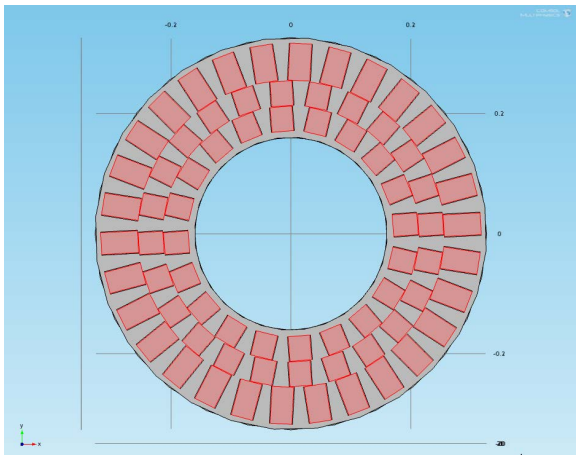


Figure 6.26: Layout of modules on pixel disk 5, with the module positions shown in red. This particular disk has two rings of "quad" modules (modules build on four FE-I4 chips) and one ring of "hex" modules (modules built on six FE-I4).

Disks will be made of a carbon-foam core with an embedded cooling pipe and power and data cables and with a carbon-fibre skin glued to each side. The estimated non-silicon material in the active region is 2% per disk.

6.3.6 Pixel powering

On-detector power conversion is the baseline in the LoI for reasons of affordable cable modularity and material budget. The reuse of the present detector outer cables requires a strong reduction of modularity of the outer services. In the present pixel system the inner cable material is dominated by the low voltage supply wires to ensure a small enough resistance to avoid excessive voltage drops. Each FE chip will have a nominal operating current close to 0.6 A and an expected final delivery voltage of 1.2-1.5 V at the front-end. There is an ongoing ATLAS R&D activity to reduce the material budget by reducing the number of cables and increasing the modularity per stave. This requires higher supply voltage, lower current and especially the need to regulate the voltages for each module. Both DC-DC and serial powering are considered as powering options, the latter being

the baseline.

Internal DC-DC converters would reduce the supplied current by a factor of four. Rudimentary on-chip DC-DC converters were already implemented and tested inside the FE-I4-A [62]. This option is for the moment on-hold in FE-I4-C development.

Research and development on the baseline option of serial powering has been going on for some time. The FE-I4 chip being used for IBL already contains shunt-linear regulators suitable for serial powering (these are being used in IBL for regulation to manage cable voltage drops, but not for serial power). Prototyping of outer Phase 2 staves with FE-I4 modules and serial power distribution is on-going [63]. Extensive prototyping of the serial powering was done with FE-I3 modules [64].

Voltage and current monitoring should be available at the external power supplies with a selectable value of the output current. SPP ('serial power protection') chips of one serial power chain share one control line, which is used to send commands to the individual SPPs to switch on/off individual modules. Serial Powering has been assumed for the service count and the calculation of the mass of services in the baseline layout. The final decision will rest on studies of noise injection and pick-up sensitivity, as well as the additional material represented by the on-detector powering components.

All operations on high voltage channels are foreseen to take place at the level of power supplies. This includes monitoring the HV current and voltage, changing the set voltage and switching on/off modules.

6.3.7 Data transmission

Bandwidth needs for data transmission are driven by a trigger requirement to allow an L0 accept rate of 500 kHz with a latency of 5 μs and an L1 accept rate of 200 kHz with latency of 20 μs. A 500kHz trigger rate readout of the full pixel detector is compatible with the L0/L1 scheme without the need for a 2-level trigger complication. In this case the full pixel detector will be readout at L0.

The output data rate per 2-chip module of the innermost layer is foreseen to be 2.56 Gb/s. The data rate for the external pixel barrel layers will be 640 Mb/s per module.

All modules of a half stave will be connected by data input and output lines, low voltage, high voltage and slow control lines to an End of Stave (EoS) card. All electrical connections are routed through this card. In addition, the EoS card includes some active components, a Giga Bit Transmission (GBT) style data multiplexer and a DCS chip. Because they are not sufficiently radiation hard and because of space constraints, the optical interface opto-board will be installed outside the tracker volume, where it can be accessed more easily. Data from the modules are transmitted to the EoS card by flex LVS lines, which are sufficient for the outer layer up to 320 Mb/s based on current prototyping. The transmission lines from the chips to the EoS card on the inner layers need more development. The data between the EoS card and the opto-boards will be transmitted via Micro Twinnax cables. Cables rather than fibres also allow more routing flexibility to reach pixel elements. The GBTx [31] chip will be used to drive the data over microtwinax cables from the EoS to opto-boards over a distance of 6-7 m.

The layout of the readout chain in the detector is shown in Figure 6.27.

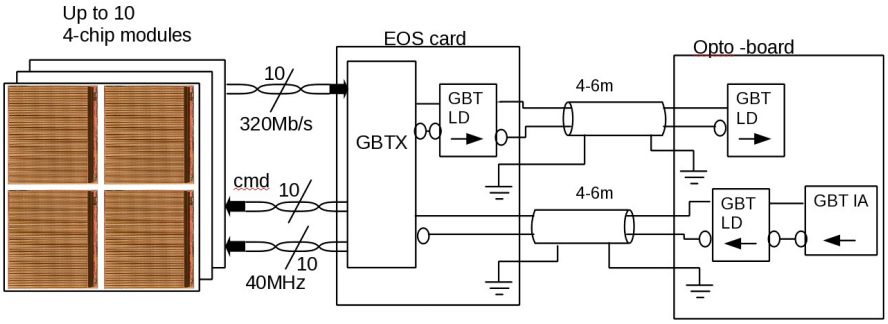


Figure 6.27: Input and output data connections between modules, EoS card and Opto-board.

Depending on the number of modules associated with the EoS card, more than one GBT chip might be located on the same EoS card. The GBTs can be powered by the same low voltage as DCS chips.

Opto-electrical transmission will be used to get the data out of the detector. The basic elements are the on-detector multiplexer combining the signals from several modules, the GBT chip, the opto-boards for electrical to optical conversion and the receiver card at the outside of the detector reconverting optical to electrical signals. A first version of the GBT has been produced. However, for Phase-II operation the baseline assumes a next generation high bandwidth GBTx in 65 nm technology, called LpGBTx. Where very safe transmission with Forward Error Correction (FEC) is in use, this provides a bandwidth of 6.4 Gb/s. Turning off FEC, which should be possible in the direction of the pixel output data flow, improves the bandwidth up to 8.96 Gb/s. The technology of the off detector part of the readout is discussed in Section 6.5.2.

The data rates per module are derived by conservatively assuming 200 pileup events, full 500 kHz L0 pixel readout and the baseline active front-end area and pixel pitches.

The expected modularity of the components is driven by the data rate per module. Given a data rate of 2.56 Gb/s per module in the innermost two layers, only two modules can be handled by one GBT. These will be electrically transmitted with twinnax cables and via the opto-board into an optical fibre. The output data rate of 640 Mb/s for the outermost layers can be accommodated by connecting 10 modules to one GBT. This leads to the modularity listed in Table 6.12 [1]

The transmission of 2.56 Gb/s along the inner staves from modules to the EoS card requires R&D on low material cables (picocoax or twisted pairs). A high rate of 5.76 Gb/s from EoS card to opto-board is achievable with a combination of a short ($\sim$2 m) section of micro-Twinnax cable and a longer section ($\sim$4 m) of macro-Twinnax cable with lower electrical losses.

Current estimates give 812 links in the pixel barrel and 288 links in the pixel disks.

6.3.8 Pixel global support and service routing

There will be one support frame for the two inner barrel pixel layers and a frame for the outer pixel parts comprising the outer barrel layers and the disks [65]. During integration and for removal

[1] This table assumes 51 staves in layer-4, disk-5 with hex/hex/quad modules and disk-6 with hex/hex modules

Detector:	Module type	Rate/module [Mb/s]	GBT/stave side	Rate/GBT [Mb/s]	Links/side	Total links
layer 1	dual	2560	6	5120	96	192
layer 2	quad	2560	9	5120	144	288
layer 3	quad	640	2	5760	64	128
layer 4	quad	640	2	5760	102	204
Total barrel	dual/quad				406	812
disks 1-5	hex/quad	960/640	6/10	5760	60	120
disks 6	hex	960	6	5760	12	24
Total disks	hex/quad				144	288
Total pixels					550	1100

Table 6.12: The main parameters of the pixel readout.

these structures will slide inside support tubes, an Inner Support Tube (IST) for the inner structure at a radius of 110 mm, and the Pixel Support Tube (PST) for the outer structure at a radius of 345 mm. These need to be hermetic to prevent humidity and other contamination entering the outer sections of the tracker during the removal or insertion process. In the latter case the barrier could be provided by the inner cylinders of the barrel and EC strip structures, with rails mounted for the support of the pixel package. It is likely that there will be separate frames for the barrel and disk sections of the outer structure, which will be populated separately with their local supports and then joined together during the pixel integration. The limited space in front of the end-cap calorimeter in a standard opening scenario also means that the (clam shell) inner pixel package either needs to be able to split at $z = 0$, or is designed so that all type-1 services are routed in one direction.

The I-beam staves will only be supported at their ends, which for the outer barrel will come from an end-plate connecting the frame with the IST on the inside of the package. The pixel disks will be supported between the frame on the outside and the IST on the inside. To estimate the amount of material, a minimum face-sheet thickness (CFRP) of no less than 350-500 μm is assumed. With frame and support tube constructions similar to the current ATLAS pixel detector the structural material is estimated to be 3.06% X_0 for the disks, and 1.68% X_0 for the combined barrel (inner and outer).

The service routing for the pixel detector will be developed along the following lines:

- Minimize the amount of material in the active tracker volume and for the outer detectors.
- Allow a replacement of the inner two barrel layers in a short (standard) shutdown.
- Allow a removal of the whole pixel detector in a long shutdown without interfering with the strip system,
- Where possible, minimize the particle fluence induced by the pixel services within the tracker volume.

Because of the high radiation environment, removal of old and installation of new cables should be as fast as possible. In addition, work on these highly irradiated cables would lead to large challenges on infrastructure, personnel and time. Therefore all services up to PP2 will be removed and new ones will be installed.

In total 4152 LV wire pairs are estimated for the barrel and 2880 for the disks, where the low voltage cables will be be organized in groups for one type of modules assuming serial powering.

The services of the two innermost layers will be routed axially within the outer radius of the inner pixel package. To allow for a replacement in a short shutdown all type-I services will run towards one end of the inner pixel package, leading to a length of detector and services of 3.7 m [66]. Alternatively, if the inner pixel package is split at $z = 0$, services could run towards both ends.

For the outer two layers the type-I services will constrain the pixel disk dimensions. It is estimated that a space of 40 mm between IST and the inner disk radius is sufficient for the type-I services from the two outer barrel layers and the disks, provided that no space for disk EoS cards is required in this volume. After passage through the disk system the services would run at increasing radius, outside of the tracker acceptance limit at $|\eta| = 2.7$.

It is expected that work which would be needed on the current pixel type-II services (external services from PP1 to PP2) to re-use them for the Phase-II tracker will be very challenging due to the radiation environment. Therefore it is planned to remove all type-II services and to install new ones. Quick removal is vital and the design of the new type-II services must be optimized for fast installation.

6.4 The strip system

The upgrade strip detector consists of a central barrel region between $\pm$ 1.3 m and two end-caps that extend the length of the strip detector to $\pm$ 3 m. They cover $\pm$ 2.5 units of rapidity. The strip barrel consists of five full length cylinders that surround the beam-line and a short "stub" barrel that covers the loss of acceptance between the end-cap and barrel. The layers are populated with 472 full length staves (236 on each side of Z = 0) and each stave with 26 modules (13 on each face). There is a small gap between staves at Z = 0. The strips on the inner three cylinders are 23.820 mm long (short-strips) and those on the outer two cylinders and stubs are 47.755 mm long (long-strips).

The stave is the basic mechanical element of the barrel. It consists of a low mass central stave core that provides mechanical rigidity, support for the modules, and houses the common electrical, optical and cooling services. The interface to the stave is the End-Of-Stave (EOS) card. Figure 6.28a shows the main components.

The strip end-caps have seven disks on each side. Each disk is populated with 32 identical petals. The petal is the basic modular mechanical unit analogous to the stave. The petals have nine modules on each face organized in six rings. The outer three rings have two modules side-by-side and the three inner rings have one module each. Covering the complex geometry over such a large area requires six different sensor geometries and 14 hybrid variations. Figure 6.28b shows the components of the petals. Because it extends beyond the required acceptance it is possible to eliminate the innermost ring of the outermost disk.

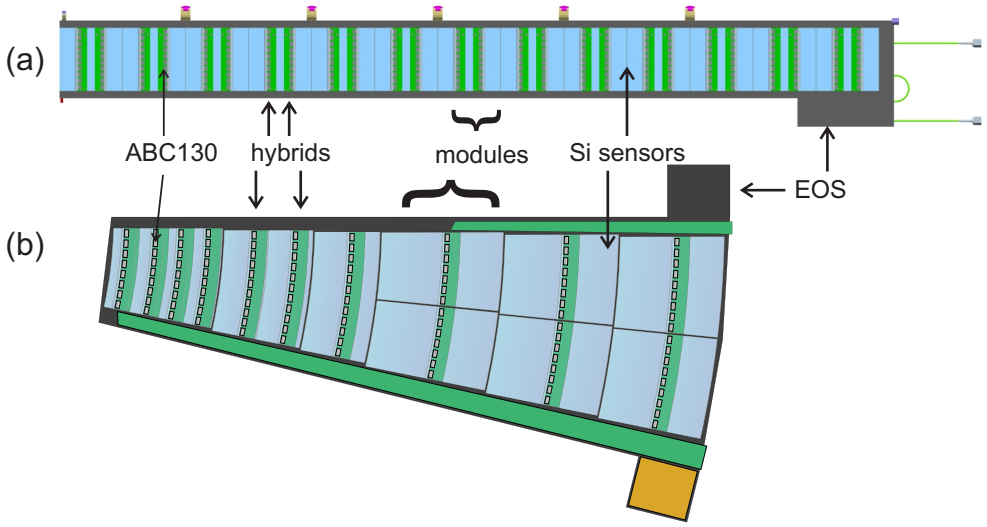


Figure 6.28: a) Barrel stave components. b) Petal stave components

The basic electrical unit of a stave or petal is a module. Application specific front-end chips ("ABC130") are mounted on kapton circuits to make a hybrid. One or two hybrids (four for the innermost ring of a petal) are glued to the surface of a silicon detector to form a module.

Many stave components have already been prototyped and a complete evaluation of their performance is ongoing. So far the results of the evaluation are very encouraging and the development is continuing.

6.4.1 Sensors

Strip sensors are AC-coupled with n-type implants in a p-type float-zone silicon bulk (n-in-p FZ). This type of sensor collects electrons and has no radiation induced type inversion. The sensors are 320 ± 15 μm thick; the sensor thickness could be reduced to 250 μm but this will incur additional costs.

The size of the barrel sensors is 97.54 x 97.54 mm^2 to maximally utilize the area of a 6" wafer and reduce cost. There are 1280 strips across a sensor, giving a strip pitch of 74.5μm. The strips are parallel to the sides of the sensor. On a stave, one side will be axial, with a 40 mrad stereo angle being achieved by rotating the sensors on the other side. There are two variations of barrel sensors. One has four rows of short strips (23.820 mm) to be used on the three inner cylinders, and one has two rows of longer strips (47.755 mm) for the outer two cylinders. This choice matches the track density at HL-LHC giving reasonable occupancy as well as low strip capacitance.

The petal sensors need radial strips (i.e. pointing to the beam-line) to give an accurate measurement of the $r\phi$ coordinate. As a result the petal sensors have a wedge shape. The dimensions of the sensors have been chosen to use as few 6-inch silicon wafers as possible [67] with 32 petals per disk and fully covering the radial range required by the layout. A 40 mrad stereo angle between strips on opposite sides of a petal is achieved by rotating the strips 20 mrad within the sensors.

The petal sensors are divided into pairs of rows of strips as the ABC130 chip is designed to read out two rows of strips. The inner-most ring is in a region of very high track density and radiation damage, and so needs very short strips. It has four pairs of rows, with lengths chosen to

accommodate four hybrids. The next ring has two pairs, and the outer four rings have one pair. The number of chips on a hybrid, and hence the number of strips, is chosen to keep the strip-pitch at the bond pad region as close to the barrel pitch (74.5 μm) as possible.

The key requirements for the strip sensors are to withstand the expected maximum fluence of 8.1×10^{14} $n_{\mathrm{eq}}/\mathrm{cm}^2$ and to operate up to 500 V. To allow for uncertainties in fluence calculations, a specification of 2×10^{15} $n_{\mathrm{eq}}/\mathrm{cm}^2$ is imposed. Prototype sensors [68] that featured short strips, with one half of the detector having axial strips and the other half having stereo-angled strips, have been designed and fabricated. Miniature sensors ($1 \times 1\,\mathrm{cm}^2$) from the same wafers were subjected to extensive radiation testing. Figure 6.29 shows the charge collected in a number of sensors fabricated by HPK after standard annealing of 80 minutes at 60°C and at 500 V bias.

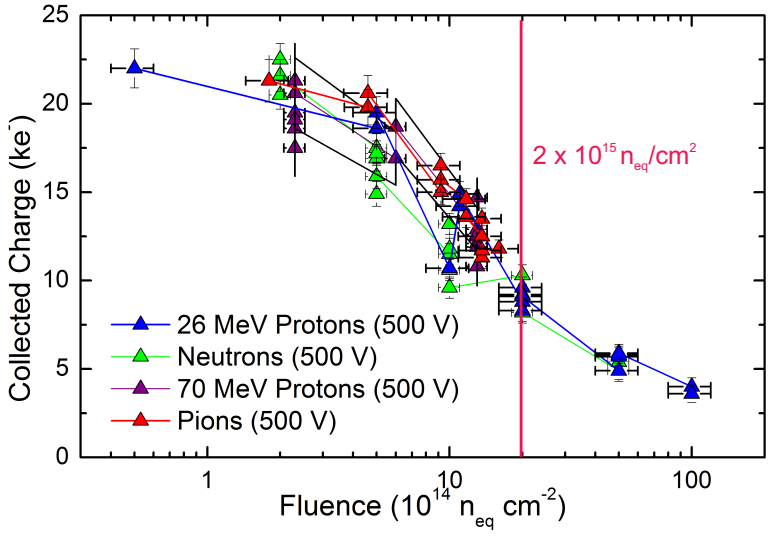


Figure 6.29: Collected signal charge at 500 V bias voltage for minimum ionising particles as a function of 1MeV $n_{\mathrm{eq}}/\mathrm{cm}^2$ fluence for various types of particles.

The measured collected charge and leakage currents of these miniatures are used to estimate the signal-to-noise at the highest fluence expected at the end of HL-LHC operation. The different sensor geometries and temperatures as well as the expected noise performance of the ABC130 readout ASIC are taken into account to estimate the noise. The signal-to-noise is estimated to range between 23–25:1 and 17–26:1 at the end of HL-LHC operation for the different regions of the strip barrel and end-cap, respectively. For all sensor types and regions, the performance at the end of operation is significantly better than the signal-to-noise of 10:1 that is needed for efficient tracking.

In addition, single-sided and double-sided modules have been irradiated with 24 GeV/c protons at the CERN PS to HL-LHC doses. With both, the measured increase in the noise was consistent with shot noise expectations. This indicates that gluing the hybrids onto the sensor face does not cause significant noise problems after irradiation, and that the 250 nm CMOS technology used in these prototypes is radiation tolerant. The 130 nm CMOS technology to be used in the upgrade is expected to be more radiation tolerant to total ionizing dose than the 250 nm technology [69].

6.4.2 Strip FE electronics

Figure 6.30 shows the main electrical components on a short-strip stave. Two hybrids, each containing 10 ABC130s (130nm CMOS ASICs) [70], are mounted on each silicon sensor. Each hybrid has a Hybrid Controller Chip (HCC) [70] that interfaces the ABC130s to the EOS. Trigger, Timing, and Control (TTC) signals are sent from the EOS to each HCC via the TTC/DATA/DCS bus. Each HCC has a unique address. The HCC sends its data along its dedicated link (called an e-link) to the EOS so that for the stave there are 26 data e-links. The TTC consists of a 40 MHz system clock, a serial command/L0 trigger, and a R3/L1 trigger that are sent to each HCC in parallel. The TTC/DATA/DCS bus and power bus are integrated into a single copper/aluminum/kapton bus tape that is laminated onto the stave core.

The EOS has a GigaBit Transceiver (GBT) [31] that interfaces with the HCCs and a Versatile link (Vlink) [71]) fibre optic driver. LV and HV power connect to the EOS and are distributed to each hybrid via a power bus. A power interface connects each hybrid to the power bus. The nature of the power interface will depend upon which power architecture is chosen, either serial power or DC-DC conversion. There may be several detector slow control (DCS) links that communicate with each power interface from the EOS. DCS interlocks will also connect to the EOS from the overall ATLAS DCS system.

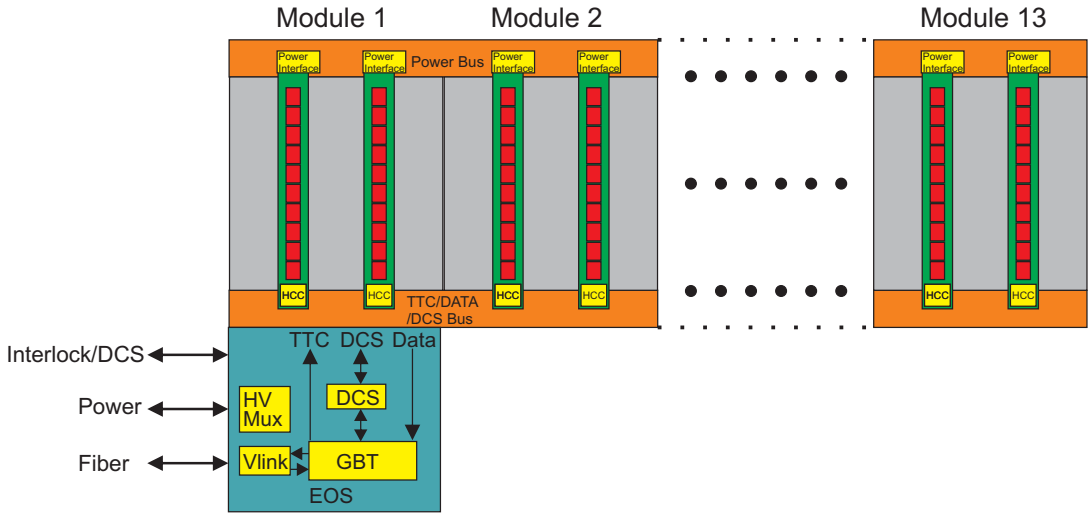


Figure 6.30: Overview of main electrical components.

To date all module prototyping has been done with the ABC250, a 128 channel, 250 nm CMOS successor to the ABCD chip used in the present ATLAS SCT. To save mass, the ABC130 has 256 readout channels that will read out strips from two adjacent sensor rows. Like its predecessors it is a binary readout chip that provides only hit information and no charge information. It differs from its predecessors in that it has two buffers that permit implementation of a two level trigger scheme consisting of a L0 trigger and a L1/R3 trigger.

All hybrids for the barrel are identical and will contain ten ABC130s. For the petals there are 13 variants. The hybrids are fabricated on kapton circuits to minimize mass. In addition to the ABC130s, the hybrids contain passive components and the HCC interface chip. As for the ABC130, the HCC will be fabricated in 130 nm technology. The HCC serves to interface each

hybrid to the EOS. It is envisioned it will incorporate some DCS functions such as reading hybrid temperatures and sensor bias currents.

Modules [72] are constructed by directly gluing kapton flex hybrids to silicon sensors with electronics-grade epoxy. The design focuses on material reduction and so uses no substrate between the kapton and the sensor. The design uses the large cross-sectional area of the hybrid and sensor to the stave core for cooling.

The power required by the strip tracker is dominated by the ABC130s. The ABC130 will require at most 0.7 mW/channel according to the most recent estimates [70]. To account for additional neglected power from components such as the HCC or the sensors a total power dissipation $\sim$1mW/channel is assumed. Total channel count is shown in Table 6.6; we therefore expect total power dissipations of approximately 47 kW for the barrel strip detector and 27 kW for the endcaps. A short-strip stave requires approximately 133 W, a long-strip stave approximately 67 W, and a petal approximately 60 W. The average power dissipation per unit area is approximately 39 mW/cm^2.

6.4.3 LV power

Because of the limited space for cables in the ATLAS detector and the need to save material in the tracker volume, the low voltage power will be multiplexed. Both serial powering and DC-DC powering are being considered. These distinct approaches are schematically shown in Figure 6.31. The relative merits of each are discussed in [73]. Provisions have to be made to isolate individual hybrids that fail in either system to permit operation of the remaining hybrids.

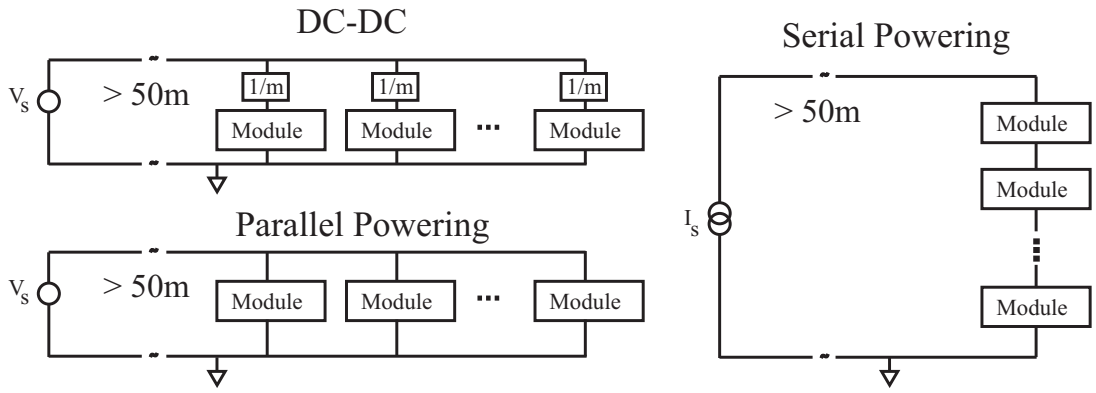


Figure 6.31: Sketches illustrating low voltage powering with the DC-DC, serial power, and parallel power architectures.

DC-DC converters are being developed for the DC-DC architecture [74]. The converter will provide up to 3A of current and provide a voltage ratio m $\approx$ 8. A four module stave (stavelet) has been built with ABC250s and is powered with DC-DC converters built with commercial components. It is pictured in Figure 6.32. Its noise performance was comparable to that of a single module powered directly from a low noise power supply.

Serial powering uses shunt regulators on each module to generate the voltages needed by the front-end ASICs. A Serial Power Protection ASIC(SPP) [70] is being designed that in combination with pass transistors built into the ABC130 will act as a shunt regulator. A serial power stavelet

has been built with ABC250s using commercial shunt regulators and a custom bypass circuit. It has been successfully tested and its noise is comparable to that of a single module operating from a low noise power supply.

For the barrel strips, determination of which architecture will be used will depend upon results of the prototyping program with ABC130s and the custom power ASICs. Power efficiency, mass, noise, and reliability of the systems will be compared. The petal strips have a preference for DC-DC conversion as the power efficiency should be higher since serial power must deliver the same power to each module. Due to the variation in the number of ABC130s on the petal hybrids, power must be wasted (dissipated in the shunt regulators) on the hybrids with a lower number of ABC130s. In principle, however, serial power can work for the petals with acceptable efficiency.

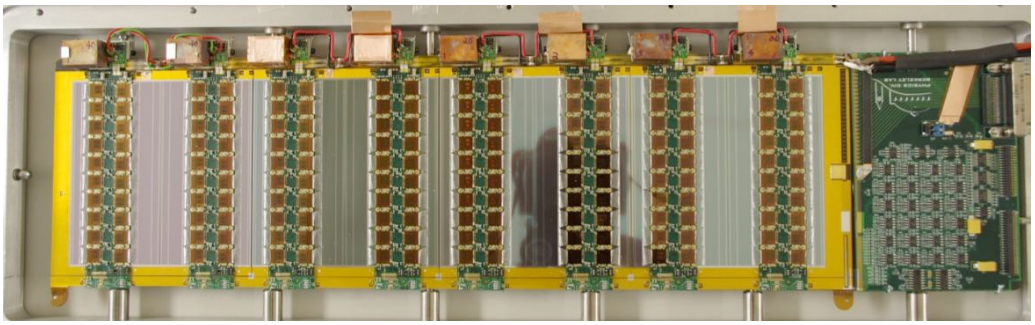


Figure 6.32: DC-DC powered stavelet with ABC250s.

6.4.4 HV power

Lack of space for cables and traces will not permit each sensor to have its own HV bias. For example, if existing cables are re-used, groups of 4 sensors or more will need to be connected in parallel. Should one sensor fail due to high current it may result in the loss of the other modules on the same bias line.

To avoid such losses, an approach is being investigated in which all 13 modules on one side of a stave are powered with a single line. Each sensor, however, can be disconnected from the bias line with a DCS controlled HV switch. Such switches must be radiation-hard and operate above 500 V. MOSFETs and BJT transistors that operate at such voltages are not sufficiently radiation-hard. Two emerging technologies that feature transistors made from gallium nitride or silicon carbide have the potential to produce high voltage radiation hard switches and are under investigation.

6.4.5 Stave and petal design

The stave core [75] is a carbon composite structure that provides both mechanical support and cooling for 26 modules (13 on each side) in the barrel strips. Figure 6.33 (a) shows a drawing of the stave core with its main components labeled. A core of carbon fibre honeycomb and carbon foam with embedded cooling pipes is sandwiched between two carbon fibre facings. The facings are made from several layers of carbon fibre. The carbon fibre has both high tensile modulus and high thermal conductivity. The honeycomb provides shear stiffness so that the facings along with the honeycomb form an I-beam like stiff structure. The carbon foam provides little shear stiffness; its function is to provide a low thermal impedance path between the facings and the cooling pipe.

Figure 6.33 (b) shows a prototype stave core with a copper/aluminum/kapton electrical bus tape co-cured into the facings.

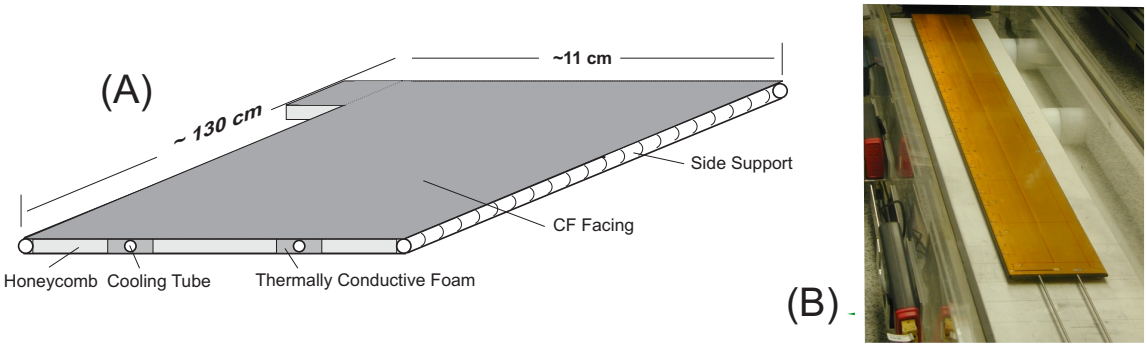


Figure 6.33: a) Drawing of a stave core. b) Photo of a stave core with a kapton bus co-cured into facing.

The thermal requirements on the stave performance are set by the need to comfortably avoid thermal runaway in the sensor. Thermal runaway occurs when increases in detector self-heating, due to radiation induced leakage current, creates a situation where there is essentially positive thermal feedback and the detector current "runs away". It is a function of the coolant temperature, thermal impedance of any location on the detector to the coolant, and sensor and ABC130 power. The design of the stave core is intended to minimize the thermal impedance. Simulations have shown [75] that current designs have a large safety margin ($> 20^{\circ}$C coolant temperature headroom) to prevent thermal runaway.

A requirement on the stave is that it be insertable in the z-direction rather than the radial direction. In principle, this permits the replacement of a stave in all but the last stages of testing on the barrels. One approach, in which the stave is end-inserted onto five carbon fibre/peek brackets, is shown in Figure 6.34. A second approach using cantilevered support to permit smaller stave tilt angles is also under development. A program to compare and choose between the two options is in place.

A number of prototype stave cores have been built and tested [75]. The intention is to optimize fabrication techniques, minimize mass, validate simulations, and to develop and test quality assurance techniques. The stave shown in Figure 6.33 (b) is an example of a prototype that would satisfactorily meet the needs of the upgrade. However, further mass minimizations continue to be explored.

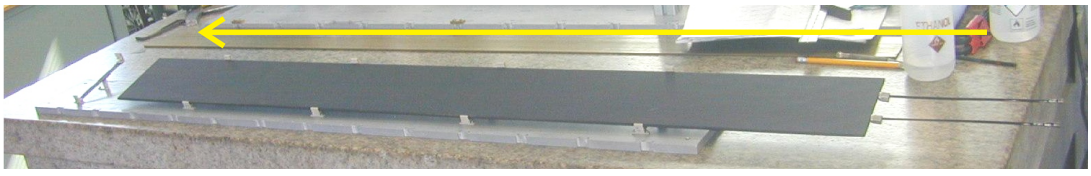


Figure 6.34: Stave being end inserted onto five carbon fibre/peek brackets

The petal core supports and cools end-cap modules. It uses the same materials as stave cores. Here, the main differences are outlined.

The petal core is wedge shaped, with the cooling tube approximately V-shaped as shown in Figure 6.35. Each module in the outer three rings is cooled by a single length of tube, while the inner three modules – which have higher power densities – are cooled by two lengths of tube, similar to a stave.

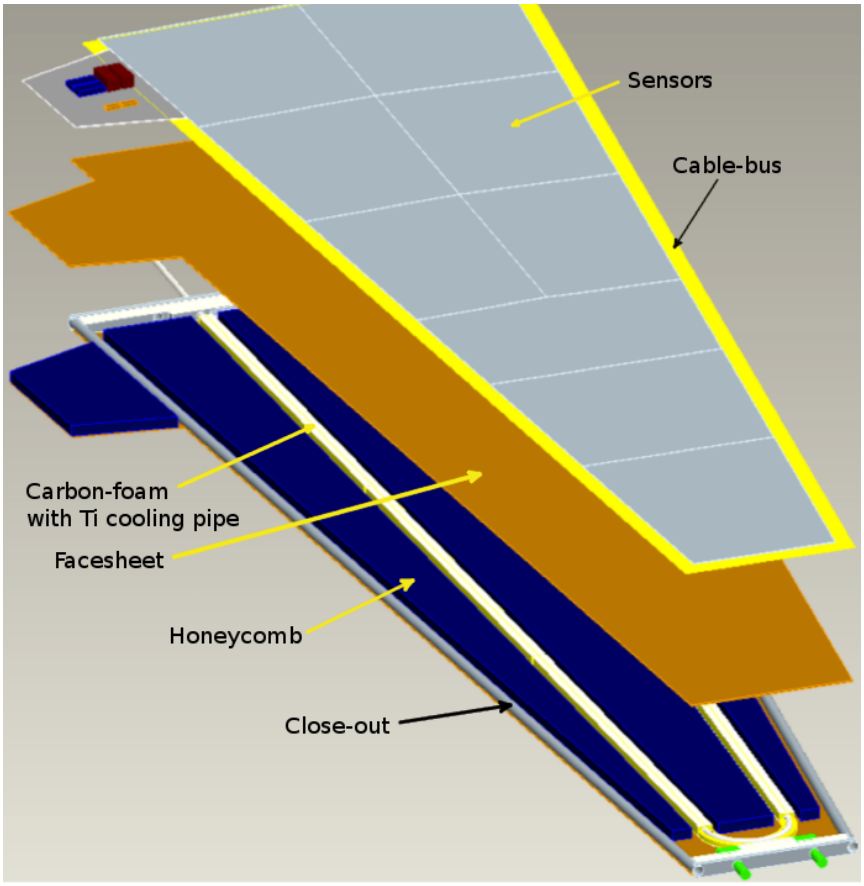


Figure 6.35: Petal core showing the various components.

A petal has fewer modules than a stave (18 cf 26) and most of these (14) have a single hybrid. This reduces the power load and servicing needs compared to a stave. Furthermore the wide-end of a petal is much wider than a stave. Placing the cable-bus under the modules adds significant unnecessary material. Instead, a cable bus will run down each outer edge of a petal, with the modules glued directly to the core face-sheets. To ensure HV insulation the face-sheets can be coated with Parylene-C as used on SCT end-cap module spines. On each side of the petal is an EOS that connects to the cable bus on that side.

The cooling parameters are quite different between staves and petals. Gluing sensors directly to the core-facing slightly reduces the thermal impedance. The outer modules have lower power density than short-strip barrel modules, but with only one tube-section for cooling so that the heat path-length is roughly doubled. The inner most modules have higher power-density than short-strip barrel-modules, but have two lengths of cooling pipe giving more area and shorter heat-path. Initial studies show sufficient margin to thermal runaway with CO_2 cooling at -30 °C. Several prototypes have been made and measured, as documented in [76].

Staves are arranged in concentric cylinders centred on the beam-line as illustrated in Figure 6.36. The staves are rotated a small angle (the tilt-angle) away from the tangent to allow an overlap in the ϕ direction. The amount of overlap is arranged to make the layer hermetic down to 1 GeV/c tracks, by choosing the number of staves in the cylinder, the cylinder radius, and the tilt-angle. The tilt-angle is $\geq 10^{\circ}$, tuned to allow a 15 mm thick envelope around each stave to prevent interference between adjacent staves. The direction of the rotation is chosen to minimise the amount the Lorentz-angle spreads the charge among strips. The overlap is sufficient for software alignment, with at least 2 % of tracks passing through the edges of two neighbouring staves. The number of staves in a cylinder is chosen to be a multiple of 4 so that each quadrant is identical; this simplifies the routing of services and the design of structure supports between cylinders.

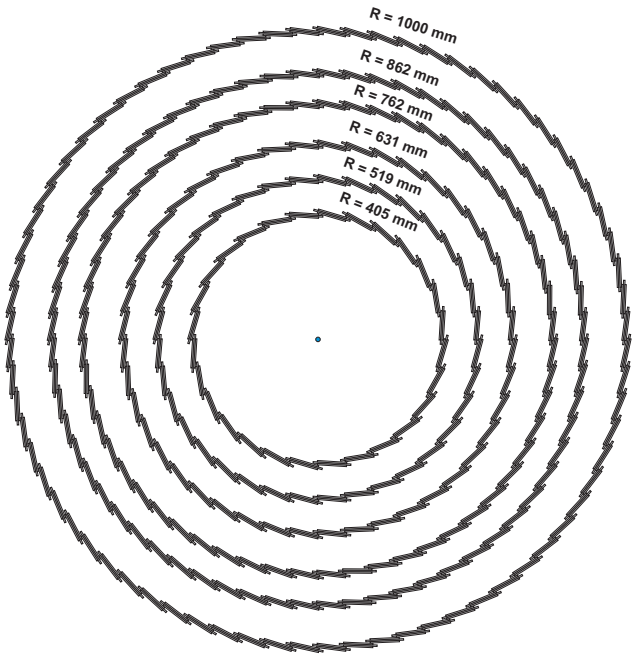


Figure 6.36: Arrangement of staves in barrels. Staves are tilted 10 degrees

The petals are arranged to make the end-caps hermetic down to 1 GeV/c. The first disk is the most critical as it has the highest low-p_T acceptance. The overlap is achieved by increasing the sensor active area so a petal covers slightly more than $1/32^{nd}$ of 2π. The optimum layout is being evaluated, with two options: the turbofan (figure 6.37) and a castellated layout (figure 6.38). The turbofan is akin to the barrel stave arrangement. However, in the end-cap region the geometry means that the gap between neighbouring petals increases with radius. A tilt-angle of about 13° is needed to give a 15 mm thick envelope at the inner radius, ending with about 45 mm centre-to-centre at the outer radius. The castellated version currently also has a large gap all the way along the petal (about 45 mm) but it may be possible to reduce this by moving the support disks out from between the petals and by moving all EOS connectors onto one side of a petal.

Table 6.13 shows estimates for the radiation length of a single short-strip barrel stave as well as for a petal. These numbers represent best estimates based upon stave and petal prototyping and extrapolation to the use of expected lower mass components (e.g. ABC130 hybrids and titanium

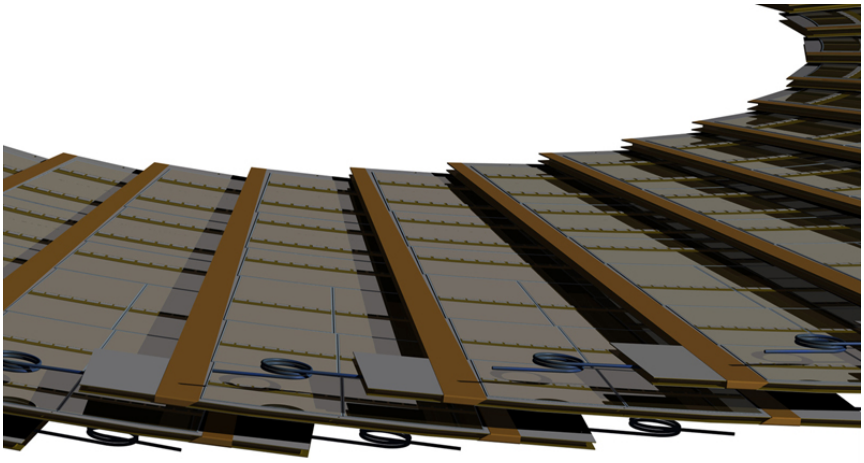


Figure 6.37: Turbofan arrangement of petals.

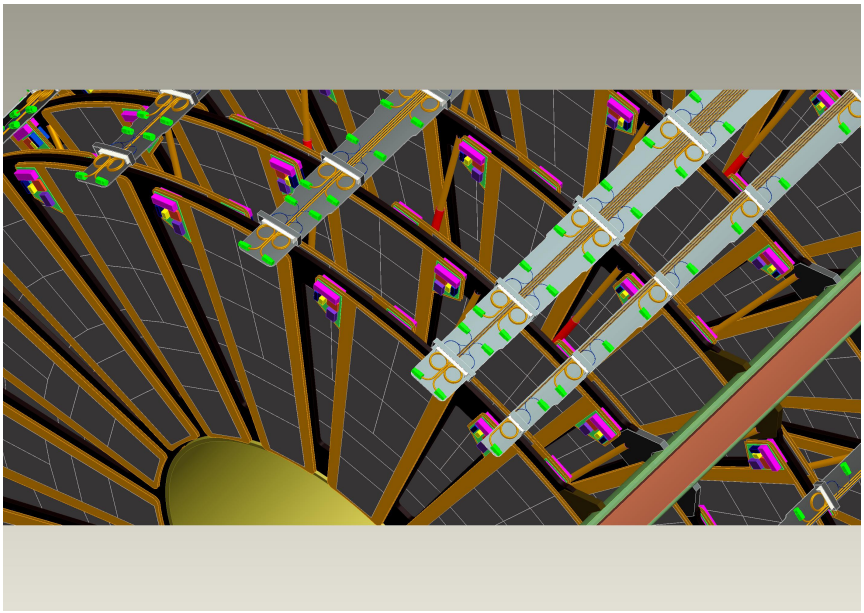


Figure 6.38: Castellated layout: Petals arranged on either side of a disk, with services on one ear only.

cooling pipes). For comparison, the current SCT radiation length excluding barrel or disk support is 2.48% for the barrels and 3.28% for the end-caps [24]. The lower expected radiation lengths are due to the high degree of sharing of support, power and services in the upgrade designs.

6.4.6 Supermodule

As an alternative to the stave concept, "supermodules (SM)" [77] are considered as an option. This is an evolution of the demonstrated merits of the existing SCT design (mechanical and thermal stability, true stereo space-point reconstruction, overlaps to mitigate weak alignment modes, etc) while addressing some existing design limitations. In this design, double-sided modules are attached to a light, stable carbon fibre local support (LS) that is end-inserted into the overall tracker support structure. Figure 6.39 (a) shows the design of a prototype short-strip module, and Fig-

Barrel		End-Cap	
Element	% Radiation Length	Element	% Radiation Length
Stave Core	0.55	Petal Core	0.47
Bus Cable	0.30	Bus cables	0.03
Short-Strip Modules	1.07	Modules	1.04
Module Adhesive	0.06	module adhesive	0.06
Total	1.98	Total	1.60

Table 6.13: Radiation Length estimates for the barrel stave and end-cap petal. Power ASICs and the EOS are not included.

ure 6.39 (b) shows the SM frame supporting the modules. Figure 6.39 (c) shows the design of a prototype LS that allows stress-free end insertion to the structure.

The module design [78, 79] allows a relative sensor alignment at the level of $\pm 1 \mu$m. The hybrid and sensor thermal paths are independent and the use of low CTE and high thermal conductivity materials minimizes deformations during temperature cycling (detailed FEA calculations indicate a longitudinal deviation of only 1.4 μm between room temperature at the design coolant temperature of -35°C, and a sensor temperature of 11-13°C above the coolant temperature). Eight modules have been fabricated using the ABC250 chip. The noise level for directly powered modules is uniform, between 560 and 590 ENC, at 250 V bias and a module temperature of 35°C. The modules remain stable for bias voltages well beyond 500 V. The module design can evolve naturally to the ABCN130 prototype and the material is estimated in the range 1.6-1.68% X_0 for short strip modules assuming 320 μm thick sensors and a hybrid width in the range 12-15 mm.

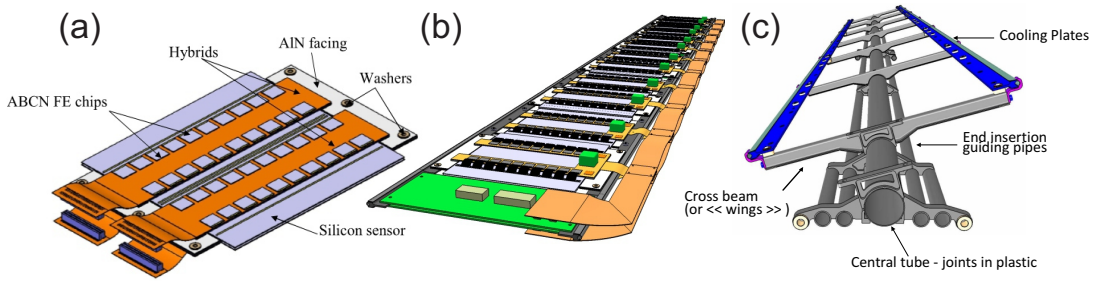


Figure 6.39: a) Schematic of the double sided short strip modules using the ABC250 FE chip b) Supermodule frame supporting double-sided modules. c) Supermodule support allowing end insertion.

The SM concept of Figures 6.39 (a) and (b) foresees 13 modules attached to a light, stable local SM support. This concept has been (is being) verified mechanically, thermally and electrically. It has been developed with the following guidelines: all SM components are decoupled from the module design and can be separately optimized, with precise reproducible module connections, minimal thermal deformation and full hermeticity; the assembled modules, the cooling pipes and the service bus are thermo-mechanically separate, minimizing mechanical stress; full re-work is possible beyond individual components until the commissioning step; and the procurement, construction and QA steps can be factorized.

Modules of the SM are attached to the light (0.18% X_0), stiff carbon fibre (CF) local support (LS) made of a central tube supporting CF "wings". High thermal conductivity cooling plates (0.18-0.2% X_0), attached to the wings with $\pm 10\mu$m precision, provide a mechanical and thermal contact for the modules with similar precision. A multi-layer flex in contact with the blocks, made of Al for the power planes and Cu for the signal and HV (as for the IBL), ensures minimal material (0.11% X_0), at the same time decoupling the thermo-mechanical stresses from the LS and modules. Figure 6.40 (a) shows the photograph of a prototype SM with dummy modules and the LS. A radiation hard grease contact with the cooling pipes limits the distortion to the $\pm 1\mu$m level during temperature cycling. A "pre-production" mechanical prototype is foreseen before Summer 2013.

A separate 8-module electrical prototype SM has also been constructed and equipped with 8 electrical double-sided modules (Figure 6.40 (b), demonstrating the advantages of modular assembly for the SM. The electrical components and prototype results are discussed in [80]. Using individual high-voltage connections to each module, and prototype DC-DC converters for the digital voltage, uniform noise is achieved. Noise levels of 580 to 625 ENC are achieved and when powered on one side the noise is at the level of individual modules in test boxes. Figure 6.41 shows the noise for each hybrid with 4 modules assembled. Work is ongoing to understand the approximate 40 ENC noise excess with respect to individual modules (related to grounding and shielding) when all 8 modules are powered.

Apart from the design evolution of modules and SM, as well as the choice of DC-DC or serial powering and high voltage multiplexing, the most important design issue is the electrical and mechanical optimization of the cable bus (this can be developed independently of other design criteria). The second issue is to optimize the material budget, which is currently estimated to be in the range 2.15%-2.25% X_0.

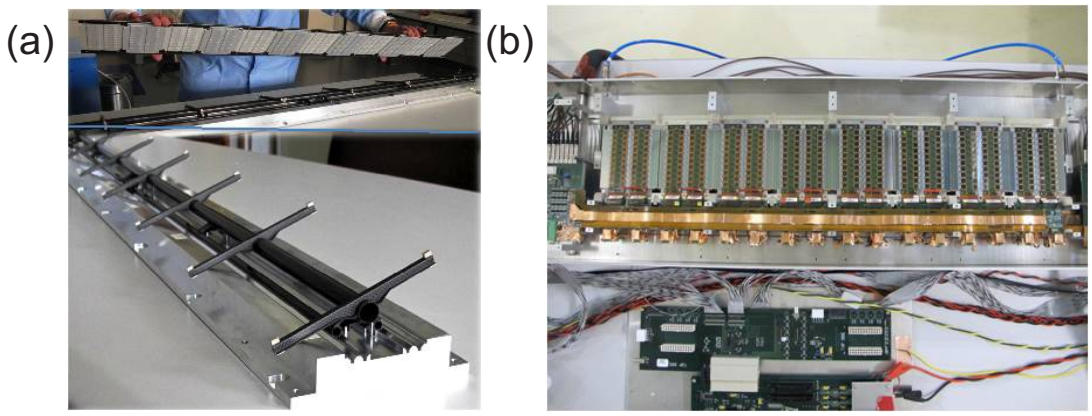


Figure 6.40: (a) Photograph of mechanical SM prototype loaded with 8 mechanical double-sided modules (upper) and the local support with an aluminium frame replacing the CF support cylinder (lower). The cooling plates and modules are mounted on the CF wings. (b) Photograph of the electrical SM prototype loaded with 8 electrical double-sided modules read via a cable bus to the SMC card.

6.4.7 Barrel and End-cap Structures

Five support cylinders will support the full length barrel staves. The cylinders will be approximately 2.6 m long with radii between approximately 400 mm and 1000 mm. They will be fabri-

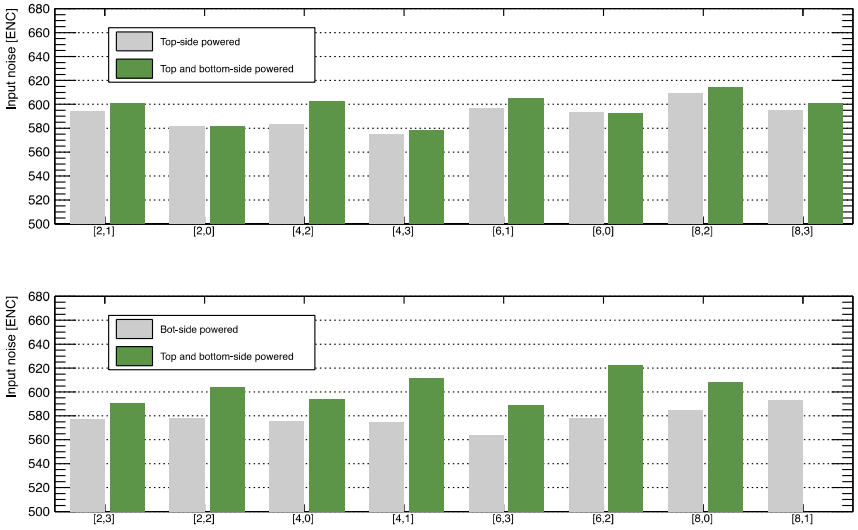


Figure 6.41: Average noise measured for each of the upper and lower hybrids with 4 modules mounted, with either the single side powered, or with both sides powered.

cated with carbon fibre reinforced plastic in order for the cylinders to have high beam and torsional stiffness.

Two options for the design of the support cylinders are being considered. The first option, similar to that used in the current SCT, uses a carbon fibre skin-honeycomb-skin sandwich structure. The second option uses single skin cylinders with reinforcing rings to provide hoop stiffness and to improve local rigidity at critical locations, i.e. at the bracket locations for the stave local supports. The first option should be stiffer for a given amount of mass, while the second option should be simpler (and cheaper) to fabricate.

The end-cap support will consist of an inner cylinder, an outer frame, and structural elements to connect the two. The outer frame will hold the services in 16 axial rows, each serving 14 petals. The petals are stiff enough that 3 supports, one at the inner radius and two at the outer radius, will be sufficient.

6.4.8 Type-1 services

The services to and from staves connect between the EOS (End of Stave) card and PP1 (Patch Panel 1). They convey low and high voltage power, timing/trigger/control-commands/data on fibres, coolant, detector control (DCS) signals, and grounding and shielding needs. The services represent a considerable amount of non-sensitive material, which can adversely affect both inner tracker and calorimeter performance. Considerable effort is on-going to reduce the services: by not putting in excessive detector layers, by multiplexing, and by careful choice of components with low radiation-length materials and small sizes. Another important consideration is assembly time, reliability, and repairablility during assembly.

Whilst the new inner tracker will have many more channels than the current one, the total cross-sectional area of services per channel will be much less. This is achieved through multi-plexing. Staves and petals will have one LV supply and return per side, with serial or DC-DC

powering. The HV power granularity will be at least 4 modules per HV supply, and if the HV switching scheme is successful, will also be one HV per stave or petal side. One GBT per side will receive timing, trigger and control signals and send data back (at 5 Gbit/s compared to the current 40 Mbit/s links). Most of the DCS data will also be routed by the fibres, with only a very limited number of copper wires remaining, which are needed for protection from over heating. Coolant is one inlet capillary and one small exhaust line per stave/petal, with some further multiplexing foreseen before exiting the inner tracker volume. Grounding and shielding needs are minimal. The cross-sectional area of type-1 electrical plus fibre services in the new tracker, per readout channel, is a factor 24 smaller for short strip staves than the current barrel SCT.

The barrel-stave services pass from PP1 around the outside of the end-caps and inside the Outer Cylinder, then bend radially inwards in the gap between the barrel-to-end-cap gap to reach the EOS cards. The hope is to fit all the barrel PP1s outside the outer radius of the end-caps, allowing the end-caps to be inserted without displacing them. If this cannot be achieved, a scheme with barrel PP1 on hinges is under investigation; during end-cap insertion, they are folded out of the way and once the end-cap is in place they are rotated back in place. To speed up the barrel assembly, the services will be arranged in 32 pre-fabricated and tested service assemblies at each end, which will be placed axially around the inside of the Outer Cylinder.

Two schemes are investigated for the end-caps, with either 16 or 32 axial runs of services. The choice is linked to the end-cap assembly method and choice of turbo-fan or castellated layout. The intention with both schemes is to be able to assemble the support structure, attach all type-1 services to it, and then later insert the petals. This has the advantage that if during testing a problem is found with a petal or a type-1 service, then the faulty item can be removed without displacing any other components.

For both staves and petals, the connection method to cooling lines is very important. Several thousand CO_2 connections are needed, with extremely high reliability and low mass. Several options are being investigated, including direct orbital welding of thin-wall Ti tubes; brazing stainless end-pieces to the Ti and welding to standard stainless welding connectors; and developing low-mass in-house connectors that can be disconnected and re-used for the connection of staves and petals to type-I tubes. For in-situ welding, tests are under way to ensure the front-end electronics are not damaged by the process.

6.5 Common readout, data acquisition and controls for pixels and strips

The readout and control system outside the EOS card will be as much as possible common to both pixel and strip systems and adopt the standards of ATLAS at large. Its key elements are the optical data and control links, the data acquisition system and the detector control. A generic overview of the pixel and strip detector's electronic systems is given in Figure 6.42.

6.5.1 Optical links

Given the challenges of the HL-LHC of higher data rates and increased radiation, the main direction of the optical link R & D is to increase significantly the bandwidth compared to the current detector while improving the radiation hardness of fibres and especially electronics. These are common challenges faced by ATLAS and CMS for the optical links in the Phase-II upgrade. Therefore, for HL-LHC, groups from ATLAS and CMS are collaborating to develop a Versatile Link (VL), which

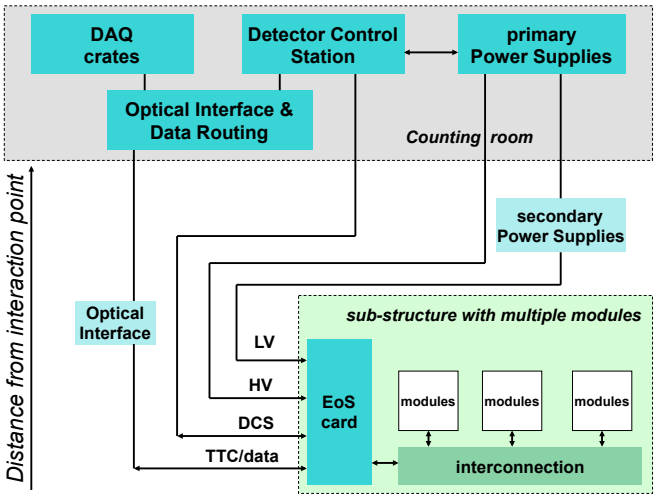


Figure 6.42: Overview of the electronic system of pixel and strips

could be used in all the sub-systems. The strategy adopted by the VL project is to make minimal modifications to existing commercial components, thus benefiting from the extensive reliability testing of large scale manufacturers as well as the extent to which these components are qualified by the customers. The VL is based on bi-directional optical links. The VL being developed is based on the SFP+ standard and will be used in conjunction with the GBTx [31]. Prototypes of the basic building block, Versatile Transceiver (VTRx) have been produced which give a very encouraging "eye diagram" [81]. The future aim is to have a new version in 65nm technology, which would consume less power and be able to operates at 9.6 Gbits/s, a significant improvement compared to the currently used 80 Mbit/s.

As discussed before in sections 6.3.7 and 6.4, the maximum data rate per link will be 5.6 Gbits/s for the pixels and 4.16 Gbits/s for the strips. These numbers include the requirements from the RoI track trigger. As a backup solution strips can work with present 130 nm GBTx by turning off the Forward Error Correction (FEC).

The radiation hardness of VCSEL lasers in transmitters (VTx), p-i-n diodes in receivers (VRx), optical fibres, connectors and optical couplers have been found satisfactory for Phase-II application. Small performance degradation can be mitigated by adjusting operating parameters.

The feasibility of the VL for use in the Phase-II tracker has been established and can therefore provide a baseline solution for the optical links.

6.5.2 DAQ

Several technological platforms are considered for the DAQ in ATLAS. For the pixel and strip detectors the system will be as common as possible, and will be integrated into the general ATLAS solution. A generic view of the DAQ system is shown in 6.43

The readout requirements for the pixel and strips systems discussed in the previous sections leads to a total of 3196 readout links as listed in Table 6.14.

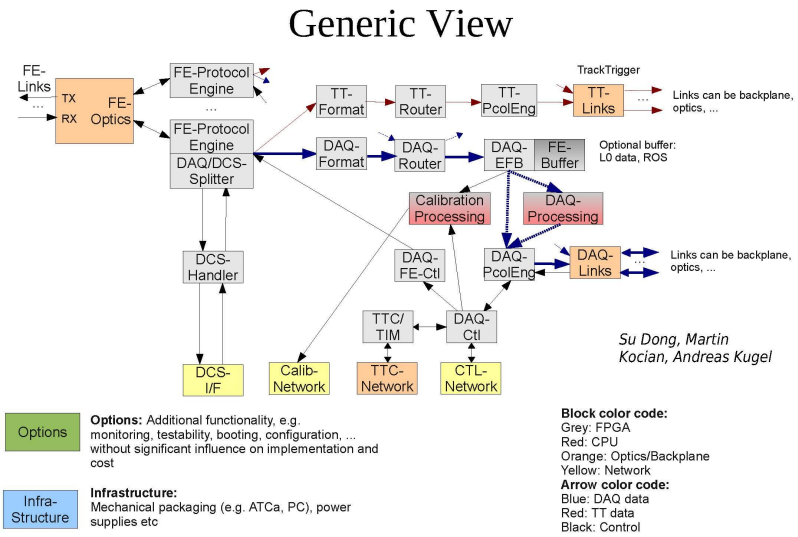


Figure 6.43: Generic Phase-II tracker DAQ diagram.

Detector:	Number of GBT links
Pixel Barrel	812
Pixel Disks	288
Strip Barrel	1200
Strip Disks	896
Total	3196

Table 6.14: Estimate of the number of Phase-II tracker readout links

For the strips, a readout rate of 500 kHz at L0 and 200 kHz at L1 was assumed, whereas for the pixels it was assumed that the full detector would be read out at the L0 rate of 500 kHz. The proposed system has sufficient bandwidth for the expected data rates for the maximum luminosity considered of 7×10^{34} cm^{-2}s^{-1}.

The currently considered technological platforms for VME ROD replacement are ATCA and PCI. Both platforms can be used for the Inner Detector. LoI baseline is according to ATLAS wide decision to use ATCA standard.

6.5.3 Detector Control System

The detector control system (DCS) for the Inner Tracker provides diagnostics, monitoring and an interlock functionality. The components for pixel and strips will be largely identical. The basic philosophy is that the control path should be independent of the operational status of the detector and the interlock system will be purely hardware based. The diagnostics will be routed via an optical link. Figure 6.44 shows a generic overview of a common pixel and strip DCS system.

The control of the modules will be managed by a DCS chip on the EOS card. The circuits incorporated at each module will permit the measurement of: temperature and humidity, low volt-

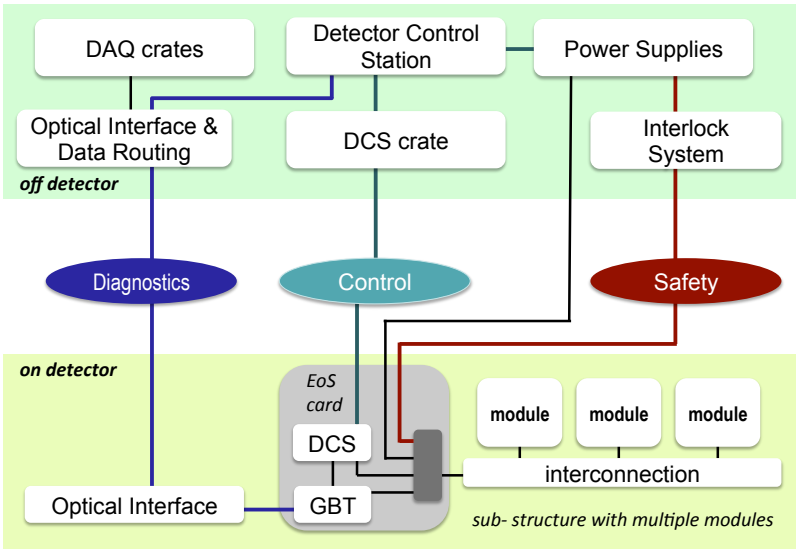


Figure 6.44: Architecture of the detector control system

age and current, sensor leakage current and possibly voltage. The exact implementation of this functionality will depend upon the choice of powering scheme (DC-DC or serial power). In either case it will be possible to disable a module under DCS control. The DCS monitoring data will be transmitted off detector using the same optical fibres carrying module data. Optical interface in the counting room will contain a data routing system that will separate the physics data from DCS data (see Figure 6.42). This will allow the DCS system to continue to operate even when the DAQ system is down.

6.6 System engineering

The system engineering tasks include the design of the overall support structures, the integration of large sub-assemblies to complete the inner tracker, installation, all external services, the neutron moderator, and the cooling system. The design has to take into account the need to install in a short time (total shutdown 18 months) in a significant radiation field, as well as allowing for maintenance scenarios [82].

The tracker will be assembled above ground as two units: the inner pixel layers as one, and the rest, including outer pixels, barrel and end-cap strips, as the other. The latter assembly will be housed in a very large carbon fibre cylinder (2.2 m diameter x 7 m long). These assemblies will be tested above ground before installation. They include patch panels at the ends for fast connection to services mounted on the cryostat end flange.

Radiation levels in the inner tracker region of ATLAS due to activiation will be significant during installation and even higher during repair scenarios after running of the HL-LHC. After

considering the importance of good vertex reconstruction performance, the harsh environment and past experience, the following requirements for the tracker have been arrived at:

- The inner two pixel layers should be removable in a standard shutdown. This limits their length, including services, to 3.5 m.
- The outer two pixel layers and the pixel discs should be removable during a long shutdown.
- The strip end-caps should be removable while leaving the barrel and outer cylinder in place.

These requirements have led to the design described in the following sections.

6.6.1 Global support and tracker integration

The entire tracker will be housed inside the Outer Cylinder [65] made with carbon fibre skins over a honeycomb core. This will be attached via two mounts at each end to the cryostat. As well as supporting the tracker; the humidity barrier, thermal barrier and grounding and shielding elements are all integrated within it. Large external flanges will be used during assembly to maintain circularity. After assembly, these will be replaced by internal disks.

The inner bore of the cryostat will be covered in a polyethylene or similar neutron moderator about 2.5 cm thick. The inner wall of the barrel region of the OC will house a further cylinder of polymoderator about 2.5 cm thick. The end-cap region omits this extra layer to make way for services.

The outer cylinder supports the strip barrels through radial interlinks and the strip end-caps through semi-kinematic mounts. The inner strip barrel holds the Pixel Support Tube (PST). The outer pixels in turn support the Inner Support Tube (IST) which holds the inner pixels. Figure 6.45 illustrates the support scheme, and more details are available in [65].

The entire inner tracker will be assembled above ground, and tested before insertion [66]. Services from the staves, disks, and petals will be connected to patch-panels (PP1) mounted at the two ends of the structure. Each of the major sections (inner pixel, outer pixel, barrel strip, two end-cap strips) can be assembled in parallel. The strip barrel will be assembled directly in the Outer Cylinder, and so will take place at CERN. Other parts can be assembled at CERN or at outside institutes.

6.6.2 Cooling

The pixel and inner parts of the strips will suffer from very high hadron fluxes, leading to large leakage currents unless the sensors are kept cold. The entire tracker will be cooled by a two-phase system, with liquid boiling in the staves, disks and petals to cool the modules and remove all heat produced in the tracker volume. The cooling needs, including heat-load, thermal runaway calculations, thermal models, and an initial distribution scheme with tube sizes, is available in [83] and references therein. The overall power to be removed from the tracker volume is currently estimated at 180 kW nominal, 240 kW with safety-factors, at a coolant temperature of -35 °C.

Liquid CO_2 is chosen as the coolant baseline, because of its high heat transfer coefficient at the required temperature. This allows for very small cooling-tube diameters ($\leq$ 2 mm), which

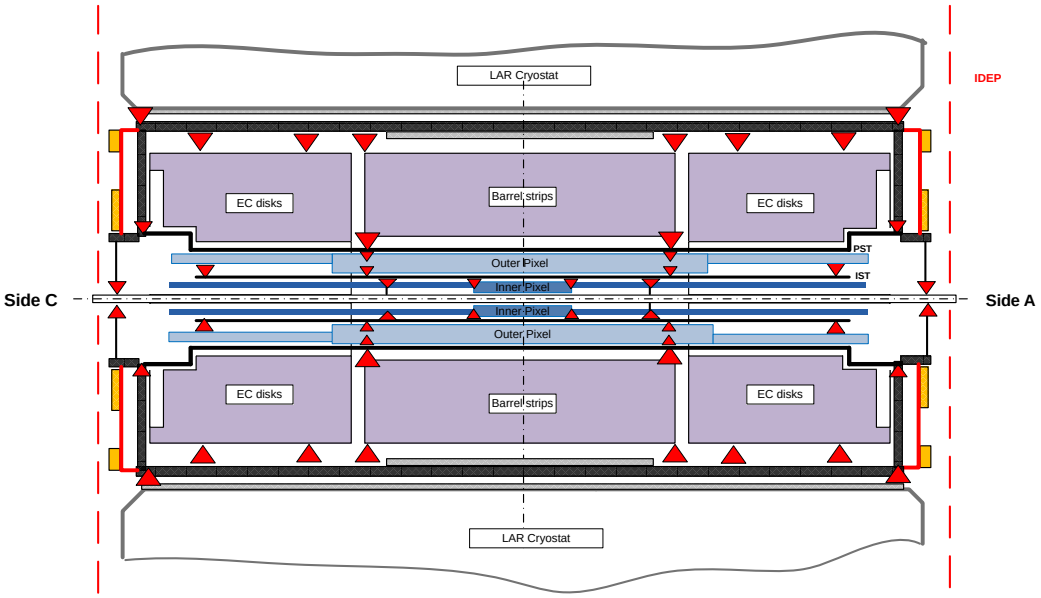


Figure 6.45: Support system for the tracker, showing how loads of each major component are transferred to the outer cylinder.

reduce radiation lengths, are easier to handle, and reduce forces due to thermal expansions and contractions. ATLAS is building up experience with CO_2 cooling in the IBL project.

There is a large development needed to go from the IBL scale (3 kW) to the 200 kW needed for the tracker. The aim is to scale up an IBL-like design to about 20 kW and then install 10 or so identical copies in the space in USA15 vacated by the current C_3F_8 compressors. In case of unforeseen problems, the option of using fluorocarbon cooling is kept as a back-up. This will require a mixture of C_2F_6 and C_3F_8 to achieve the required termperatures.

6.6.3 External services

The Phase-II tracker, like the current ID, will be connected by long service lines to equipment in USA15 and US15. Services include low and high voltage power, data fibres, DCS lines, CO_2 coolant tubes, dry nitrogen tubes, and grounding and shielding elements. The services pass through a series of patch panels PPn, with different types of cables between: type-1 to PP1, type-2 to PP2, and so on. PP1 will be fixed on the end walls of the tracker. Services outside of PP1 are considered external.

All services of the current ID travel radially out from PP1, across the barrel calorimeter ends, to $r \approx 4230$ mm. From there, most of the SCT services are spliced from type-2 to type-3 without a patch panel, and pass in the radial gap between the barrel calorimeter and the BIL muon chambers to $z = 0$. Some TRT and pixel servcies also follow this route - particularly cooling tubes. These

services cannot be removed or replaced without removing and later re-installing all BIL chambers, which would take prohibitively long. Most of the pixel and TRT electrical services pass radially between BIL chambers to five platforms (Figure 6.46) at each of ends A and C, positioned radially between the inner and middle muon chambers and in z between about 1 and 3 m. The platforms house racks with PP2. From these racks, type-3 cables continue to $z = 0$. All services then go radially out through the gap in the muon chambers at $z = 0$.

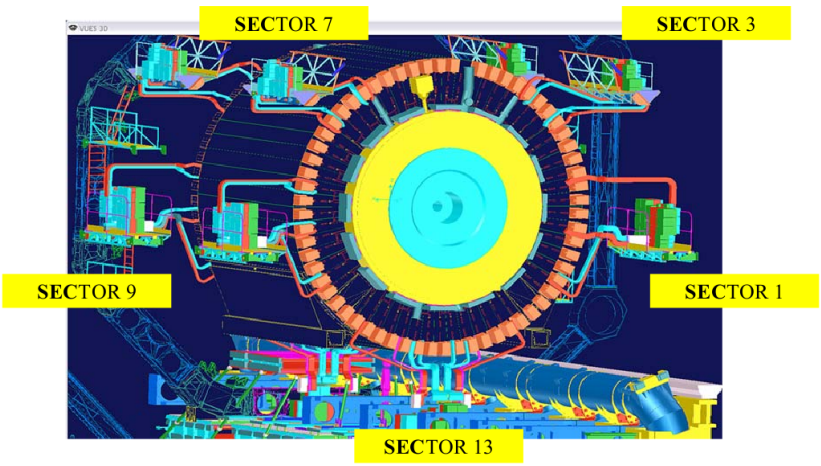


Figure 6.46: Services routing of current inner detector via PP2 platforms.

It would take too long to remove and replace the type 3 and 4 services, even those pixel and TRT services that are not buried below muon chambers. Instead they will be left in place. The intention is to re-use a large fraction of these so as to minimise the amount of new services to be installed.

The current ID fibre-optics transmit data at 40 Mbit/s or small multiples of that, using slow Step Index Multi-mode (SIMM) fibres, with each module having its own fibres. The future tracker will multiplex many modules into one fibre with the GBT chip and Versatile Links [81], greatly reducing the number of fibres needed. However, the much higher transmission rates (4.8 Gbit/s or more) require high-speed fibres; we will lay a small number of new Graduated Index Multi-mode (GRIN) fibres for this.

The CO_2 cooling system transports cold fluids ($\approx -40\,^{\circ}$C) and therefore needs insulated lines. The current ID cooling system transports room-temperature fluids, and so are not insulated. Furthermore the CO_2 tubes need to be tested up to about 100 bar, beyond the safe operation of the current ID tubes. Hence new CO_2 lines will need to be installed. The number though is very small: 10 lines are planned from USA15, one to each PP2 platform. Each line has two concentric tubes, the inner as inlet and the outer as return line; these are surrounded by a few cm of insulation. At PP2 each will be distributed into many more lines (about 45), with a final splitting inside the tracker of about one to four [83].

The LV, HV and DCS cable needs and N2 distribution can be met with current services. In fact, those pixel and TRT services passing through the PP2 platforms are sufficient, without the need to use the SCT services; clearly beneficial since re-use of the SCT services would require cutting the

type-2 cables near the outer rim of the ends of the barrel calorimeter, and re-connectorising them – a considerable amount of work in a radiation environment.

The presence of significant radiation levels during installation will require careful design and planning to minimise radiation exposure. Tests with mock-ups, addition of shielding and development of remote-handling equipment and robotics will all be investigated. The path to be followed by the new tracker from the surface to the cryostat bore has been investigated and checked for clearance. Considerable preparation of platforms and modifications are needed for this, although only one end of ATLAS needs to be fully opened.

The inner pixel design is a significant departure from the current ATLAS pixel in that it can be removed and re-installed without breaking the LHC vacuum. The beampipe will not be equipped with heaters permanently, reducing its radiation length and allowing better vertexing performance – very important in the high vertex density at the HL-LHC. In the event of a vacuum failure leading to saturation of the getter inside the VI beam-pipe, beampipe bake-out will be required. This will need a standard opening allowing the inner pixel to be moved to a park position away from $z = 0$. Heaters will be installed around the beam-pipe and the bake-out performed. Then the heaters are removed and the inner pixel reinstalled.

In view of the critical nature of the first pixel layers for ATLAS, these are designed to be replaceable in a standard opening. Bringing all its services to one end is considered, so access is only needed on one side. It is partially retracted to allow type-1 services to be disconnected from PP0, followed by full retraction and opening of the clam-shell. The detector and type-1 services can then be removed to the surface via the lift shafts. While some minor repair such as welding to fix a leak may be possible, any major damage will require replacement with new detectors, taking advantage of technology advances available at the time.

7. Upgrades of computing and Software for high luminosities

There are two major drivers behind the upgrade requirements in the computing and software for ATLAS. Both require R&D in the short term, partial solutions in the medium term and complete solutions by Long Shutdown 3.

The first is that the computing and software must respond to the needs of the upgraded detector systems and increased luminosity. This means it must not only provide simulations and reconstruction code that takes into account the upgrades to the detectors, but the code must also deal with the increased event complexity. Furthermore, a mixture of increased selection, compression, algorithmic efficiency and new workflows will be required to deal with both the increased instantaneous rate of new events and the rapidly growing integrated data volume. The simulation and initial reconstruction code must be developed in the short term, to meet the design needs for the TDRs, and then iterated to meet the longer-term design and efficiency requirements. The effect of the increased pile-up on the current reconstruction is illustrated in Figure 7.1.

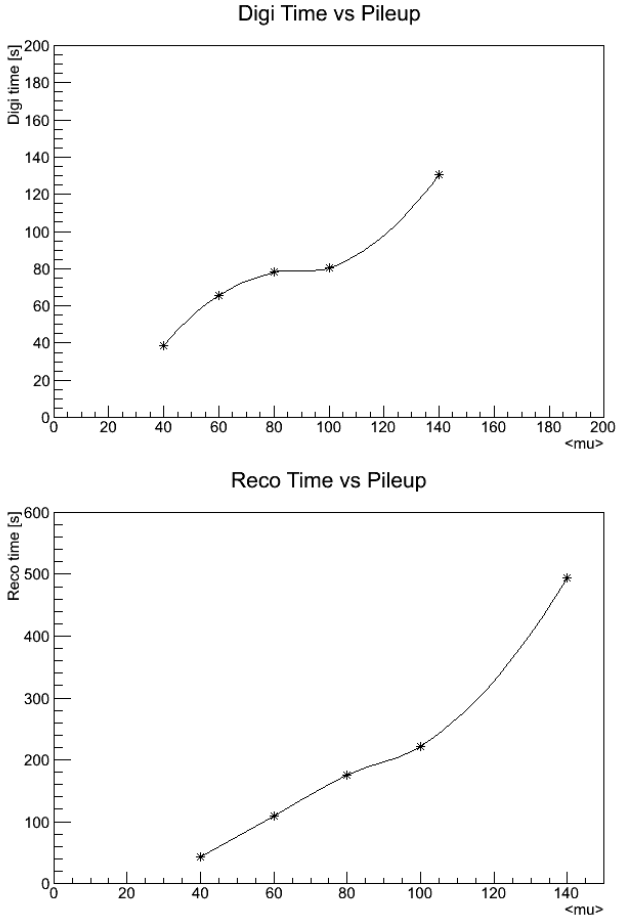


Figure 7.1: The impact of increasing pile-up on the existing simulation digitisation step (upper plot) and for the reconstruction step, lower plot.

The second major driver is the changing computing architectures that will be available at

affordable prices over the next decade. The current improvement gains in processing power have come from an increased number of cores, a trend that appears to be continuing. There are also developments towards co-processing on a single chip, which have much in common with those in programming on general purpose Graphical Processing Units. It is now evident that our code must use multi-threading, vectorization and parallelization to be efficient on future architectures; keeping on with the current code with minor changes for the new detectors will neither deliver the best physics outputs nor be economically sustainable. This implies work on the algorithms, on the data representation and on the framework, with implications for the distributed computing environment.

Given the times scales involved, ATLAS must have a flexible computing infrastructure to adapt rapidly to changing conditions. For example, by 2020, there will be increased use of commercial clouds, general science High Performance Computing (HPC) centres or other "opportunistic" facilities not owned by ATLAS. This requires that our Software Infrastructure Team (SIT) develop flexible, modular build and testing software to be able to quickly optimise the entire ATLAS software suite for each computing environment.

The target rate from trigger to offline chosen for our planning is 5kHz; rates up to 10kHz have been discussed in some trigger scenarios. The target rate chosen would allow high precision studies of boson pair production, the Higgs sector and new physics searches. This represents an order of magnitude increase from the current rate during recent running.

The data size for the output of reconstruction will be a crucial issue with such high pile-up and trigger rates. Various strategies are possible, which will need evaluation in terms of their physics impact. For planning purposes a 300-400kB average AOD size for real data (500kB including truth and/or extra trigger information) is assumed. The raw data size is expected to be 2MB/evt, which presupposes compression and optimization for size.

7.1 Computing Model Considerations

The numbers discussed above imply computing model changes. For the raw data size considered, compression should be applied online where possible. The AOD size considered for planning is aggressively small. One way of achieving this would be to reduce the per event size for all objects by, for example, increasing the tracking p_T cut; where the target for the size of an AOD for 14TeV data with a pile-up of 200 would be 300-400kB/evt.

An alternative strategy would be to vary the event data stored based on the trigger signature, allowing more detailed studies of expected signatures at the risk of bias against new physics. Such approaches would have physics costs, which will have to be evaluated.

The issue of the overall data volume will also require changes to the data production and distribution model, including changing which reconstruction output formats are kept on disk, and what is stored in those outputs. Increased use will be made of data access across the wide area network at an event level (rather than the file level), reducing the number of instances required on disk.

The balance of compute resources in each activity will also require adjustment. The reconstruction share on the Grid will grow again with high rates, while simulation will have to make more use of fast simulation modes and transient samples will also have to be employed if the simulation volume is to be greater than the real data. Our system of job submission and data distribution

must be able to quickly adapt to changing grid infrastructure, and Tier-1s will increasingly act like the current Tier-0.

7.2 Frameworks

A key to the required software developments will be the software framework. This will serve both the trigger and offline requirements, allowing the cross-porting of code between the two. The framework will have to serve the need of efficiency in speed, memory footprint and memory throughput (through the internal workflows), the IO and persistency, and the requirements of vectorization, parallelization, multi-threading and concurrency at the event and sub-event level. It will also need to serve both offline and HLT needs. One route under evaluation is the planned version of the current Gaudi framework, on which Athena is based, and which is built with concurrency support at the sub-event level.

The framework developments will have several goals. It must make the most of the prevailing commodity technology, which implies flexibility and the avoidance of over-specific solutions. Lock in to particular technologies must be avoided, keeping options open. This will require continuing R&D that will inform design decisions and allow ATLAS to profit from short term benefits. The upgrade plan for frameworks begins with a phase of R&D focussed on using the latest current software technologies and practices to make optimal use of current commodity computing platforms and their short term evolution as planned in manufacturer road maps. The purpose of this will be two-fold: first to understand the potential and be prepared if these directions are continued in the longer term, and second to yield short term benefits which can be implemented through evolution of the existing Athena/Gaudi framework.

Profiling tools will be key, and will evolve with time and with the technologies considered. This profiling will help at each stage inform the areas most needing development and the coding tools required.

The need for concurrent programming requires the framework to support multi-process event-level implementations in the short term, with Copy on Write memory sharing, and also new programming languages, support libraries and paradigms. Memory management will become even more important in Phase-II, and tools such as Software Transactional Memory(STM) show the potential to provide better access control to shared memory with traditional difficulty of programming locking strategies.

The framework must support virtualization and containers as an extension of the process forking concept, partly to give better memory sharing, but also as a way of abstracting underlying hardware. It is almost certain that best performance at reasonable cost will require the support of on-board co-processors. To maintain flexibility, a hardware abstraction client/server architecture to facilitate R&D with the growing range of hardware available, coupled with effective means of inter-process communication is proposed.

All of the above imply that the existing framework needs to be adapted in the short term to prototype some of the features and maintain moderate efficiency. In parallel, a significant new framework version will be required at the time of the Phase-I upgrade for development and in case it is needed due to computing hardware and/or software technology advances. By Phase-II, a second framework version supporting the successful development strategies needs to be in place.

This framework needs to be ready about one year before LS3 so that it can be tested by, for example integrating and testing HLT and reconstruction within the new framework.

7.3 I/O, persistence, and increasingly many-core architectures

As data volumes and rates increase substantially in Phase-II, and as computing technologies evolve, how ATLAS represents, writes, reads, organizes, stores, distributes, and accesses its data must change. ATLAS expects to be required to use a variety of data hosting, storage, and computing technologies, and so the I/O and persistence infrastructure will be designed for flexibility and agility in the face of heterogeneous and evolving computing environments.

Input and output are often points of serialization, for example file merging from multiple jobs, and this is increasingly true as node-level parallelism increases with increasing core counts. I/O is a fundamental challenge to efficient exploitation of emerging computing platforms, and I/O bandwidth has not scaled with core count. It is unlikely that every process or thread can handle its own I/O and communicate with storage independently; on the other hand, an architecture that concentrates input and output into separate processes or threads is susceptible to serial bottlenecks. This is evident even today with relatively low process counts for applications that do not involve large per-event processing times. An I/ O infrastructure capable of supporting increasingly many-core processing and evolving computing platforms is an essential component of the ATLAS framework upgrade effort.

The Phase-II data volume will require the development of new, sophisticated strategies and infrastructure for selective access to data. One should read only the events required, and only those event components that are needed for processing or analysis. This has implications not only for I/O components but also for persistent data organization. Increasing core counts will require that as little unneeded data as possible be read. When data must be transferred in order to be accessible, the development of sub-file caching capabilities will obviate needless data movement; these capabilities must be content-aware in order to deliver complete events or specific event components.

New storage technologies an the Phase-II timescale, such as solid-state devices, will have implications for I/O and persistence. Differences in latencies, in random access times, in writing-versus-reading tradeoffs, in the need to erase versus merely to mark as free, and so on, will dictate the I/O framework prefetch strategies, the write and read granularities, the buffering, the sequential reading versus direct selection and navigation, and the commit strategies. ATLAS must plan for a heterogeneous storage deployment using an array of technologies, and its I/O framework and persistent data organization must be flexible enough to be configurable, and to allow optimization for the range of storage implementations encountered.

State representation of data objects and the ability to stream them are important considerations in persistence, but exploiting multi-core architectures also often requires sharing or transferring objects. This will be made simpler with appropriately data-oriented designs, but what is required in any case is the ability to represent and transfer or share the object state. The requirement is very similar when a process or task needs to retrieve an input object from or send an output object to an I/O server task or process on the same node or elsewhere. The current approach used by ATLAS TDAQ software to serialize object states and write them to a byte stream is directly related.

Persistence technology requirements are not, for the most part, unique to ATLAS, and will need to be addressed by all the LHC experiments. ATLAS will look to common projects, such as

ROOT, to provide the required persistancy technology.

7.4 Simulation

The physics programme at the HL-LHC will require accurate simulations of the upgraded detectors and triggers to prepare samples of signal and backgrounds for a wide range of physics analyses. Given the very large data samples expected, even larger simulation samples will be required. Thus the simulation must be optimised for maximise efficiency in terms of processing time, memory footprint and storage. To do this, the code must adopt the tools and paradigms introduced with the new software frameworks, and make full use the profiling and timing tools. The simulation must also be configured so that it can be run in a direct chain with the reconstruction, avoiding the storage of events on disk.

Production of the large data samples will not be sustainable using the full simulation for all detectors in all cases. Accordingly, the simulation framework must support various approximations, allowing some tasks to be completed in a very speedy manner when appropriate. The framework must allow for the mixing of full and fast simulation and match this to the specific analysis .

Building upon upon the ATLAS simulation packages Atlfast-II and FATRAS, a new highly configurable ATLAS Integrated Simulation Framework (ISF) is being developed. Furthermore, depending on the required accuracy, different full and fast simulation approaches can be mixed within a single event to allow an optimal balance between precision and execution time; this is illustrated in graphic form in Figure 7.2. The ISF is built into the ATLAS event data processing framework Athena and is designed to allow for future extension and the application of parallel computing techniques.

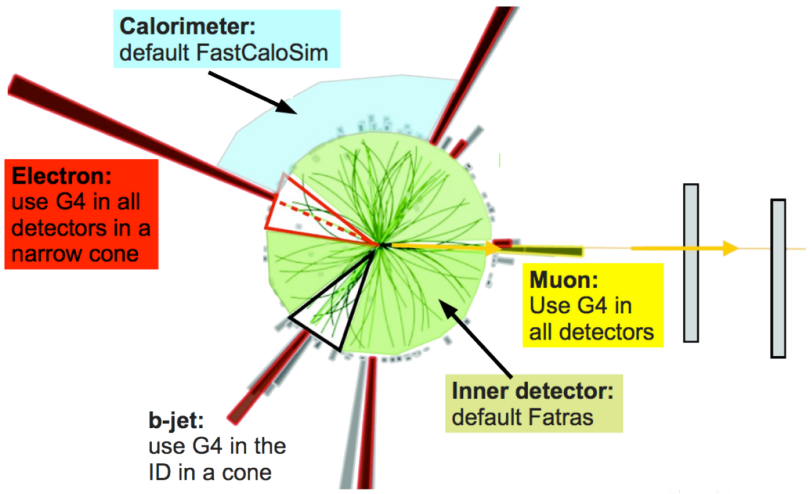


Figure 7.2: The Integrated Simulation Framework vision, indicating how its flexible nature will allow efficient choices to be made for different regions and objects.

Phase-II will introduce a regime where there will be a substantial number of multiple proton-proton collisions within each filled LHC bunch-crossing, and also multiple filled bunch-crossings within the sensitive time windows of the ATLAS detectors. This will increase with increased luminosity. Including these effects in Monte Carlo simulation poses significant computing challenges.

It is crucial to manage the conflicting demands of keeping the background dataset size as small as possible while minimizing the effect of background event re-use. The memory footprint of these digitization jobs will be minimized, to keep them within the grid limit, despite combining the information from thousands of simulated events at once.

An alternative approach, known as "overlay", will be investigated. In this approach, the actual detector conditions are sampled from raw data using a special zero-bias trigger and the simulated physics events are overlaid on top of this zero-bias data. This gives a realistic simulation of the detector response to physics events. The overlay simulation runs in time linear in the number of events and consumes memory proportional to the size of a single event, with small overhead. There are a number of computational issues and challenges that must be addressed to run overlay in upgrade production mode on the grid. In addition, to do this effectively data must be overlaid onto data without double counting the noise. This will take careful development. There are also the computational issues that will arise due to generating large amount of luminosity weighted zero-bias data and making it available on the grid.

7.5 Reconstruction

The reconstruction code is a consumer of the all the changes to the upstream data representations, and must deal in an efficient manner with the pattern recognition and reconstruction of events in the highly-complex pile-up conditions. This will require several parallel strands of development. First, the new detector geometries must be implemented and algorithms created to evaluate the designs. Given the timescales involved for the framework developments, this must, at least in part, be based on a modification to the current framework and paradigm. The second strand must structure the reconstruction to support the new coding techniques, and to re-implement (or possibly replace) the initial algorithms. At all stages, these developments must be continually optimized to improve their speed, memory footprint and physics performance. In this work, they will be aided by the same profiling and timing tools developed or implemented for the framework development. A third strand of activity is in the design of the data objects. Clearly, the event size at all stages of processing is an issue, and an effective design allowing speedy access with minimal use of storage is essential. This ongoing activity is essential, with very large potential savings in terms of storage costs.

The current Event Data Model is based on the DataVector, a vector of pointers to EDM objects. This has very bad performance on modern CPUs a re-modelling will be undertaken to utilize modern CPUs. Redeveloping the Event Data Model further, the reconstruction algorithms should internally use highly optimized special EDM objects tailored for each algorithm. Memory allocation of new objects will be reduced as much as possible, and existing objects modified and updated with new information used instead. This reduces the overhead of allocating memory from the operating system as well as improving the data locality.

Underpinning this work will be a necessary programme of benchmarking the core software in ATLAS for modern software technologies and optimal performance. The job configuration will also have to be re-worked. At present, all sub-systems configure algorithms in their own way, introducing inefficiencies and inconsistencies. With the new job configuration, all subsystems will use the same framework and the same underlying technology to configure any job, improving performance and tailoring the job to the target platform.

The vast increase in workload will mean that the jobs need to be handled more efficiently. An example of an inefficiency inherent in the current system stems from the error handling. At present, errors or crashes in the reconstruction chain result in several re-tries of the same job in the same configuration to exclude hardware related problems like out of physical memory of worker nodes, file system related problems, etc. With more concurrent reconstruction jobs the penalty to do these retries will increase due to increased number of events processed by these jobs in terms of CPU time. A more robust way of dealing with this will be implemented and the overall efficiency be increased. Single events causing repeated failures could be tried on special hardware allowing for more memory allocation.

To exploit the big improvements in simulation brought about by the ISF, the reconstruction must also become more modular, with some considerable speedups, trading off precision. A careful evaluation of the precision required will guide the changes to the reconstruction of the simulated events.

7.6 Databases and Database technologies

A nontrivial volume of critical data is needed to record events, deploy, execute, and monitor event processing and understand results is stored in databases. There is a significant amount of software and infrastructure needed to enter and maintain the data within databases and make it available to many diverse applications, including the above-mentioned event processing but also including applications which do not necessarily process events.

The full consequences for the database systems of a ten-fold increase in the online event rate will depend on the level of skimming to be performed in the initial reconstruction. However, the number of events stored in event-wise metadata will increase proportionally to the online rate (if all events are stored) or proportionally to the acceptance rate from skimming (if only skimmed events are stored). There may be a reduction in the luminosity block time duration, which would increase the volume of some Conditions Database folders, but not to the level which cannot be handled assuming needed partitioning is performed during long shutdowns. There will certainly be an increase in the number of data files per dataset. This will result in an increase in the offline task rates and number of files to be distributed and their associated bookkeeping (which must be tracked in database domain systems).

7.7 Software Infrastructure

There are many implied additional requirements on the software infrastructure for the upgrade. Most obviously the ATLAS Nightly System will need to support new branches of upgrade software releases. These branches must also include more platforms, new language tools. New compilers will be required to exploit both the hardware and new programming techniques. The ATLAS Nightly System Upgrade is a central part of ATLAS Infrastructure update and will provide the improved monitoring of nightly build results, new tools for automation of offline release shifters tasks, brings the modern database and web technologies into the Nightly System. The long term plans include the creation of distributed nightly system in which releases are created using the GRID resources and then validated on the same GRID sites where ATLAS Production tasks run.

The database infrastructure and subsystem specific implementations will be designed to operate well below capacity so that the system can handle spikes in usage so that no process in ATLAS

is limited by a reasonable database access request. Particular areas of development that are required specifically for the upgrade will mainly concern the detector control system database, PVSS, and the conditions database. Both require increased flexibility and optimisation to handle the increase in stored data, transactions and more complex queries. In addition, work will be needed to extend or replace the underlying database technologies.

7.8 Distributed Computing

Until now, HEP has taken a leading role in Grid development through the WLCG. In addition to this generic software infrastructure, Atlas Distributed Computing has developed the required ATLAS-specific components, such as: Panda, DQ2, AGIS and monitoring. All this has been used successfully in ATLAS data processing and physics analysis. In particular, it enabled the quick processing and analysis of 2012 data leading to the "new boson" discovery.

For the next few years, it is assumeed that distributed computing facilities will still be accessed primarily through Worldwide LHC Computing Grid (WLCG) components, but with an increasing use of Infrastructure As A Service (IAAS) and other technologies. It is expected that WLCG will concentrate its effort on few topics: specific to HEP or where it has the leading role, while integrating industry standard products in other areas. These will then also be incorporated into the ADC toolkit. Cloud computing is one example. Where possible, common solutions will be sought between the LHC experiments and the WLCG, with effort mutualized as much as possible.

The increased load from Phase-II will require the ADC to optimize the usage of computing resources within the LHC environment (CPU, storage, network, maintenance manpower), and must evolve to adapt to the changing ATLAS workflows, the changing hardware, and in particular to the increased dataflows and processing volumes implied by the upgrades. Changes in technology will imply development that far exceed the normal M&O, for instance work on Clouds and virtualization, or new storage techniques. A major areas that will require development to cope with the Phase-II data flows and volume is in networking. Next generation advanced networked applications (*i.e.*, Cloud based services) will require a set of network capabilities and services far beyond what is available from networks today. A new class of intelligent network services must be developed in order to satisfy additional application-specific requirements and to feed the co-scheduling algorithms that will search for real-time and scheduled resources that span the network and application spaces associated with large volume, worldwide distributed data analyses.

7.9 Analysis

Experience suggests that the current analysis system, based around the D3PD maker framework and its attendant ecosystem, provides a good basis for the future. However, there are reasonable concerns about scalability in the Phase-II era, beginning with the file and user dataset volumes. A common analysis framework where state of the art code for efficiently harnessing hardware resources on current architectures, but written to exploit many-cores, is required. It must be very easy to install on a laptop on all three major platforms (Linux, MacOSX, Windows), either from source or from binaries. It must support many-core architectures, using the embarrassingly parallel nature of the typical event analysis, either via multiple processes or multiple threads, and should also support the 'analysis train model', whereby large analysis tasks are run together in a single

centralised processing to save CPU and I/O resources. It should be based on a simple data model, easily serializable to various file formats and easily accessible from various languages. and must also deliver good throughput performances with low overhead, as well as good and automatic scalability on laptops and on the cloud.

7.10 Data Preservation

It is an increasing obligation of all science projects to put into effect a data management and access policy that extends beyond the confines and lifetime of the experiment. ATLAS has an expected lifetime of decades, but effective long term preservation requires action be taken early in the design of the data objects and software, such that it will be reusable long after the last collision is recorded. Furthermore, due consideration and effort must be made so that at least some form of the data can, after a reasonable embargo period, be used by non-collaboration members. ATLAS is currently evolving a policy on these matters. The policy must inform the other software and computing efforts, and will require effort over and above the normal exploitation of the data.

8. Physics Goals

8.1 Introduction

The physics programme of the LHC is only just beginning with pp collisions at 7 and 8 TeV centre-of-mass energies. A new boson with a mass around 126 GeV, consistent with the Standard Model Higgs particle, was recently discovered [84, 85]. This marks the start of a new era of physics discovery at the LHC, and opens a new chapter in the study of the mechanism of electroweak symmetry-breaking. This chapter presents a few highlights of the physics case for operation of the LHC at high luminosities [86–92].

The studies presented here are based on generator level simulation with a parameterisation of the detector response. The parameterisations are the best estimate of the performance of the upgraded detector based on experience of the present detector [92]. The parameterisations of detector response describe the energy and momentum resolutions, particle identification and fake rates. In general, the parameterisations were based on the current performance of the detector and extrapolated to higher pile-up using high pile-up samples with up to 69 pile-up events. In most cases, this was based on the assumption that the upgraded detector together with improved analysis algorithms would maintain the performance of the current detector in the presence of high pile-up.

An offline threshold of $p_T > 25$ GeV was used for single electrons and muons based on a L1 threshold of 18 GeV and 20 GeV for electrons and muons, respectively, and association with a hard track to simulate the track trigger. This corresponds to the thresholds assumed for the trigger upgrades presented in chapter 2 and takes into account the additional efficiency due to the proposed track trigger.

The b-tagging parametrisation used in the physics studies is based on current performance and extrapolated to a pile-up of 140 giving a light jet rejection factor of $\sim$100 for a b- tagging efficiency of 70%. However, performance studies of the upgraded tracker using full simulation, presented in chapter 6 [46], show that the expected light jet rejection is $\sim$200 for the same b-tagging efficiency. Based on this, it seems reasonable to assume that even complex performance criteria such as b-tagging can be maintained or perhaps even improved in the presence of pile-up.

Detailed studies of boosted topologies have not been made, however these will be a focus of future studies.

A wealth of measurements of the new 126 GeV boson is needed to probe its compatibility with the Standard Model Higgs boson. The nature of its spin and parity must be determined, its couplings will be explored to as many final states as possible, and it is possible that the Higgs-like self-coupling will be manifest through interference effects.

Measurements of weak-boson scattering are also a sensitive probe of electroweak symmetry breaking phenomena beyond the Standard Model Higgs mechanism.

A wide range of searches have been carried out at the LHC, extending far beyond previous experiments in sensitivity. To date no evidence has been reported for physics beyond the Standard Model (BSM) in collisions at 7 or 8 TeV, allowing limits to be placed in the TeV range for strongly-produced SUSY particles, and in the range from hundreds of GeV to 2-3 TeV for other new particles. The sensitivity scales rather well with centre-of-mass energy, but the dependence on luminosity is more complex. The possibilities for extended reach for several SUSY and exotic

scenarios are presented. An additional case study searching for flavour changing neutral currents in top decays is also included.

8.2 Measurements of the Higgs-like boson

With the discovery of a new Higgs-like boson [84, 85], a major goal is to establish the nature of this particle by determining its quantum numbers and by increasingly precise measurements of its couplings to fermions and vector bosons. A Standard Model Higgs boson with a mass of 126 GeV is particularly suited for studies at the LHC since it decays to many final states that can be experimentally reconstructed.

Models with an extended Higgs sector, like Supersymmetry, predict deviations of the Higgs couplings from Standard Model predictions that can be large, but can also be arbitrarily small, for example in Supersymmetry when the other Higgs states are very heavy. The goal is therefore to measure couplings as precisely as possible and in parallel look for other (heavier) particles of the spectrum of the new theory.

For a final confirmation of the Higgs mechanism, the observation and measurement of the triple (and quartic) Higgs self-coupling is important. While the quartic Higgs boson coupling is not accessible within any currently planned collider program, the triple Higgs coupling could be observable as an interference effect in the Higgs boson pair production.

In the studies presented here of the precision with which its properties can measured with HL-LHC luminosities, it is assumed that the recently discovered particle is the Standard Model Higgs boson.

8.2.1 Higgs boson couplings

While Higgs boson coupling measurements have already started at the LHC, the luminosity of the HL-LHC will provide substantially improved statistical precision for already established channels and allow rare Higgs boson production and decay modes to be studied. From the combination of the observed rates in all channels, detailed measurements of the Higgs boson coupling strength can then be extracted.

For an estimate of the precision with which the SM Higgs boson couplings to other particles can be measured at the HL-LHC, the following Higgs boson decays, that are already addressed in the current 7 and 8 TeV analysis, are considered:

- $H \to \gamma\gamma$ in the 0-jet and the 2-jet final state, the latter with a vector-boson fusion (VBF) selection. The analysis is carried out analogously to [84].
- Inclusive $H \to ZZ^* \to 4\ell$ following a selection close to that in Ref. [84].
- $H \to WW^* \to \ell\nu\,\ell\nu$ in the 0-jet and the 2-jet final state, the latter with a VBF selection. The analysis follows closely that of Ref. [84].
- $H \to \tau^+\tau^-$ in the 2-jet final state with a VBF selection as in Ref. [93].

For all channels, changes to the trigger and the photon/lepton selections needed to keep misidentification rates at an acceptable level at high luminosities are taken into account. For the VBF jet

selection, the cuts were tightened to reduce the expected fake rate induced by pile-up to be below 1% of the jet activity from background processes.

In addition to these channels, final states targeted to the measurement of couplings with high luminosities have now been studied:

- $WH/ZH, H \to \gamma\gamma$ and $t\bar{t}H, H \to \gamma\gamma$: these channels have a low signal rate at the LHC, but one can expect to observe more than 100 signal events with the HL-LHC. The selection of the diphoton system is done in the same way as for the inclusive $H \to \gamma\gamma$ channel. In addition, 1- and 2-lepton selections, dilepton mass cuts and different jet requirements are used to separate the WH, ZH and $t\bar{t}H$ initial states from each other and from the background processes. The $t\bar{t}H$ initial state gives the cleanest signal with a signal-to-background ratio of $\sim$20%, to be compared to $\sim$10% for ZH and $\sim$2% for WH.

  The $t\bar{t}H$ initial state is of special interest, as it yields a precise measurement of the square of the top-Yukawa coupling, which is otherwise not directly accessible. Figure 8.1(left) shows the expected signal in the $t\bar{t}H$ 1-lepton final state and Figure 8.2(a) shows the expected measurement precision.

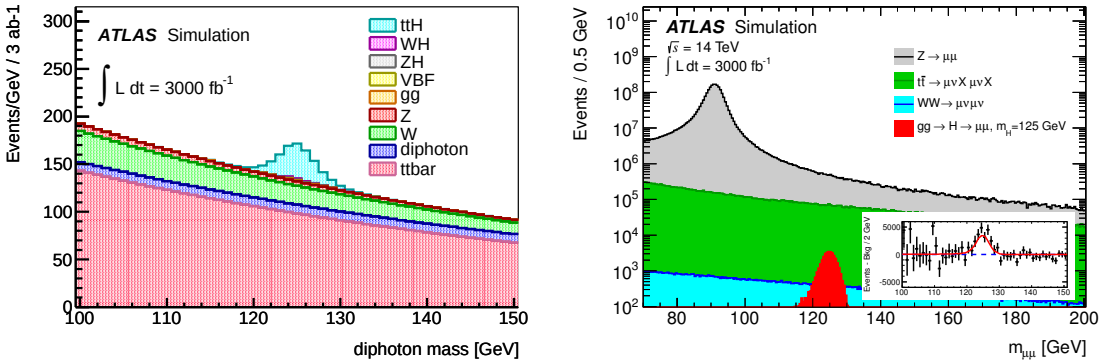


Figure 8.1: (left) Expected $\gamma\gamma$ invariant mass distribution for the $t\bar{t}H, H \to \gamma\gamma$ channel in the 1-lepton selection for an assumed integrated luminosity of 3000 fb^{-1} at $\sqrt{s}$ =14 TeV. (right) Expected invariant mass distribution for the inclusive $H \to \mu\mu$ channel, for an assumed integrated luminosity of 3000 fb^{-1} at $\sqrt{s}$ =14 TeV. The inset shows the expectation for the $H \to \mu\mu$ signal after the subtraction of the fitted background.

- $H \to \mu\mu$: this channel has also a low signal rate at the LHC with a signal-to-background ratio of only $\sim 0.2\%$. However, the expected narrow signal peak allows a signal extraction at very high luminosities, resulting in an expected signal significance larger than $6\,\sigma$ with 3000 fb^{-1} for the inclusive channel. The analysis follows Ref. [94] with changes to maximise the sensitivity for an inclusive $\mu\mu$ signal. Figure 8.1(right) shows the expected signal compared to the large continuous background and Figure 8.2(a) shows the expected measurement precision.

  Also the exclusive $t\bar{t}H, H \to \mu\mu$ channel was studied. While the expected signal rate is only $\sim$30 events at 3000 fb^{-1}, a signal-to-background ratio of better than unity can be achieved and hence this channel gives information on both the top- and μ-Yukawa coupling with a precision on the total signal strength of $\sim$25%.

An overview of the expected measurement precision in each channel for the signal strength μ with respect to the Standard Model Higgs boson expectation for a mass of 125 GeV is given in Figure 8.2(a) for assumed integrated luminosities of 300 fb^{-1} and 3000 fb^{-1}.

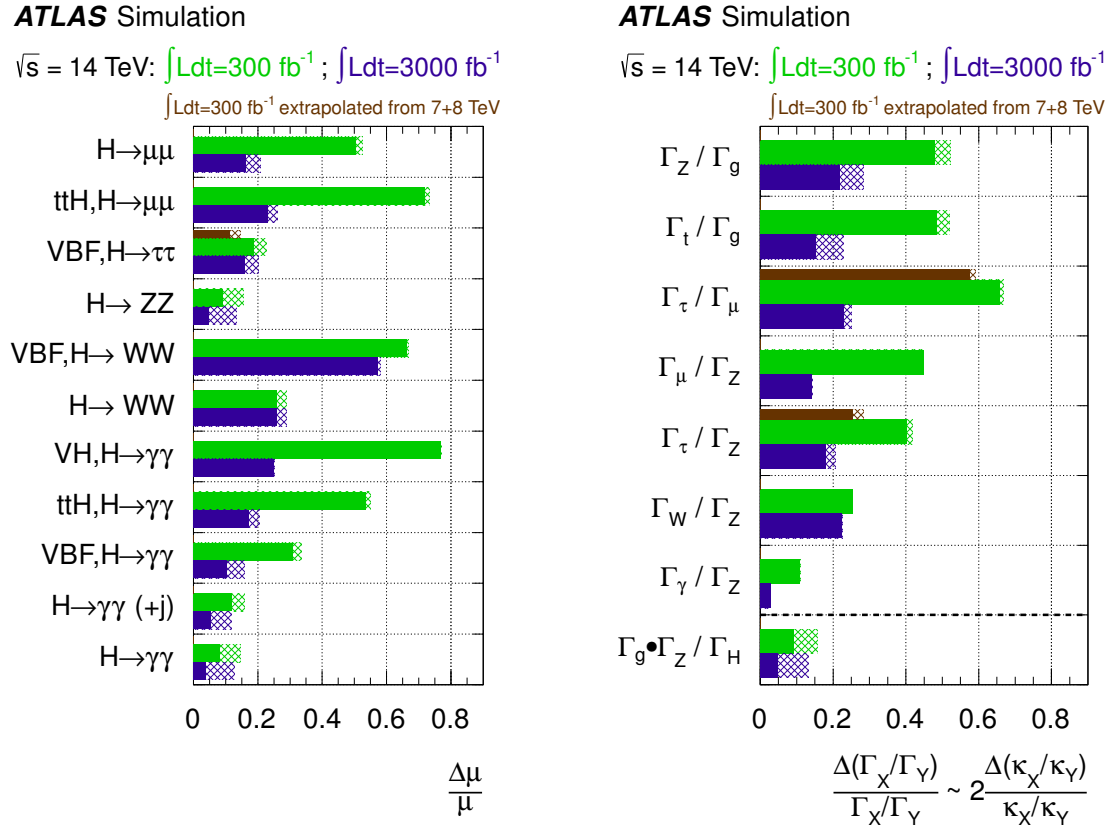


Figure 8.2: (a): Expected measurement precision on the signal strength $\mu = (\sigma \times \mathrm{BR})/(\sigma \times \mathrm{BR})_{\mathrm{SM}}$ in all considered channels. (b): Expected measurement precisions on ratios of Higgs boson partial widths without theory assumptions on the particle content in Higgs loops or the total width. In both figures, the bars give the expected relative uncertainty for a Standard Model Higgs boson with a mass of 125 GeV (the dashed areas include current theory uncertainties from QCD scale and PDF variations [95]) for luminosities of 300 fb^{-1} and 3000 fb^{-1}. For the $\tau\tau$ final state the thin brown bars show the expected precision reached from extrapolating all $\tau\tau$ channels studied in the current 7 and 8 TeV analysis to 300 fb^{-1}, instead of using the dedicated studies at 300 fb^{-1} and 3000 fb^{-1} that are based only on the VBF $H \to \tau\tau$ channels.

The $\gamma\gamma$ and ZZ^* final states profit most from the high luminosity, as both statistical and systematic uncertainties (which are dominated by the number of events in the sideband) are reduced considerably. The $\gamma\gamma$ final state is especially important, as this final state can be used as a clean probe of all initial states and associated couplings accessible to the LHC.

In the $\tau\tau$ channels, dedicated studies were done only for the VBF production of the $H \to \tau_{lep}\tau_{lep}$ and $H \to \tau_{lep}\tau_{had}$ final states, where τ_{lep} (τ_{had}) denotes a leptonically (hadronically) decaying τ. However, especially for the $\tau\tau$ channels, the combination of dedicated analyses targeted

at all accessible initial and final states is needed to reach the best sensitivity. This is illustrated by scaling the expected sensitivity at 7 and 8 TeV with $\sim 5+10\,\text{fb}^{-1}$ to 14 TeV, $300\,\text{fb}^{-1}$. From the scaling of all $\tau\tau$ channels one expects a roughly 50% more precise measurement than from the VBF channels alone (see Figure 8.2(a)). A similar improvement is also expected at $3000\,\text{fb}^{-1}$, however a more precise result is currently not available.

For the WW^* channels the signal rate is ultimately not the limiting factor as the background systematic uncertainties rapidly become dominant. The results presented here include as realistic as possible background systematic uncertainty expectations. However, sizeable uncertainties on these expectations remain. Further improvements on the W-coupling are expected from the inclusion of other channels with initial and final states that are sensitive to this coupling, but more detailed studies are needed.

The $WH/ZH, H \to b\bar{b}$ final state is not included in the current estimates, as both jet energy resolution and rejection of light jets with b-tagging algorithms are crucial for this channel and suffer from the high instantaneous luminosity conditions. More careful studies and a well understood upgrade design are both needed to finalise estimates of $H \to b\bar{b}$ measurements.

With respect to Ref. [86], the $\gamma\gamma$ results improved due to a better determination of the background shape systematic uncertainties and the addition of the WH and ZH initial states, but were degraded for the VBF and $t\bar{t}H$ initial states because of more realistic estimates of backgrounds with multiple jets. For $\tau\tau$ the VBF results improved mainly from the inclusion of the $H \to \tau_{lep}\tau_{had}$ final state and a better optimisation of the $300\,\text{fb}^{-1}$ analysis for the expected lower pile-up conditions compared to the $3000\,\text{fb}^{-1}$ scenario.

All measurements are combined in a general coupling fit, where no assumption about the particle content of the $H \to \gamma\gamma$ and $gg \to H$ loops is made. Furthermore, no assumption on possible BSM Higgs boson decay modes and hence on the total width Γ_H is made, which allows only the measurement of ratios of coupling parameters. This scenario represents the most general case. For a given production mode i and decay channel j the cross-section $\sigma_i \cdot \text{BR}_j$ is assumed to be proportional to $\Gamma_i \cdot \Gamma_j / \Gamma_H$ with $i = g, W, Z, t$ and $j = W, Z, \gamma, \mu, \tau$. The coupling fit parameters are chosen as the ratios

$$\frac{\Gamma_W}{\Gamma_Z}, \frac{\Gamma_\gamma}{\Gamma_Z}, \frac{\Gamma_\tau}{\Gamma_Z}, \frac{\Gamma_\mu}{\Gamma_Z}, \frac{\Gamma_t}{\Gamma_g}, \frac{\Gamma_Z}{\Gamma_g} \text{ and } \frac{\Gamma_g \cdot \Gamma_Z}{\Gamma_H}\,. \tag{8.1}$$

Using the nomenclature of Ref. [96] one finds: $\Gamma_X/\Gamma_Y = \kappa_X^2/\kappa_Y^2 = \lambda_{XY}^2$ and for the relative uncertainty $\frac{\Delta(\Gamma_X/\Gamma_Y)}{\Gamma_X/\Gamma_Y} \approx 2 * \frac{\Delta\lambda_{XY}}{\lambda_{XY}}$. The chosen parametrisation corresponds to the one in Table A.1 in Ref. [96].

Figure 8.2(b) shows the expected relative uncertainties on the determination of these coupling parameters assuming an integrated luminosity of $300\,\text{fb}^{-1}$ and $3000\,\text{fb}^{-1}$. The experimental uncertainties are reduced by a factor of two or more for almost all ratios with $3000\,\text{fb}^{-1}$ compared to $300\,\text{fb}^{-1}$ and reach <5% for the best cases. The ratios Γ_γ/Γ_Z and Γ_t/Γ_g, which provide constraints on new physics contributions to the $H \to \gamma\gamma$ and $gg \to H$ loops, will be measured at the ∼5–15% level per experiment. For the derived ratio Γ_τ/Γ_μ, that gives insight into the coupling relation between the 2nd and 3rd fermion generation, a precision of ∼25% per experiment is reachable.

In a minimal coupling fit, where only two independent scale factors κ_V and κ_F for the vector and fermion couplings and no additional BSM contributions are allowed in either loops or in the total width ($\sigma_{V,F} \sim \Gamma_{V,F} \sim \kappa_{V,F}^2$, see Section 4.2 in Ref. [96]), experimental precisions of ∼2% on

κ_V and $\sim$3.5% on κ_F are expected with 3000 fb^{-1}. This is a significant reduction compared to the 300 fb^{-1} expectation, which gives $\sim$3% on κ_V and $\sim$9% on κ_F. Table 8.15 summarizes these numbers and also shows the expectations including current theory systematic uncertainties.

	300 fb^{-1}	3000 fb^{-1}
κ_V	3.0% (5.6%)	1.9% (4.5%)
κ_F	8.9% (10%)	3.6% (5.9%)

Table 8.15: Expected precision for the determination of the coupling scale factors κ_V and κ_F for 300 fb^{-1} and 3000 fb^{-1}. Numbers in brackets include current theory uncertainties.

8.2.2 Sensitivity to the Higgs self-coupling

In order to establish the Higgs mechanism as being responsible for electroweak symmetry breaking, the measurement of the Higgs self-couplings and subsequent reconstruction of the Higgs potential are important. A direct analysis of the Higgs boson trilinear self-coupling λ_{HHH} can be done via the detection of Higgs boson pair production. At hadron colliders, the dominant production mechanism is gluon-gluon fusion, and for centre-of-mass energies of 14 TeV, the production cross-section of two 125 GeV Higgs bosons is estimated[2] to be $34\ ^{+6}_{-5}$ (QCD scale) ± 1 (PDF) fb. Due to the destructive interference of diagrams involving $gg \to HH$, the cross-section is enhanced at lower values of λ_{HHH}; cross-sections for $\lambda_{HHH}/\lambda^{SM}_{HHH} = 0$ and $\lambda_{HHH}/\lambda^{SM}_{HHH} = 2$ are $\sigma_{\lambda=0} = 71$ and $\sigma_{\lambda=2} = 16$ fb respectively.

A Higgs boson mass $m_H \approx 125$ GeV implies a number of potential channels to investigate, due to a wide spectrum of decay modes. Sensitivity studies at the generator level[3] for the HL-LHC upgrade were performed on just two channels, $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}W^+W^-$, chosen for their clean signature and high branching ratio, respectively[4] [89].

The final state of the $HH \to b\bar{b}W^+W^-$ channel is identical to $t\bar{t}$-production, which gives a huge background to this decay mode. As a consequence, no constraints on the Higgs self-coupling can be obtained from this channel.

The $HH \to b\bar{b}\gamma\gamma$ channel has a branching ratio of 0.27%, resulting in a predicted yield of 260 events in 3000 fb^{-1} of 14 TeV pp collisions. Several main backgrounds are considered; the irreducible $\gamma\gamma b\bar{b}$, $b\bar{b}H(H \to \gamma\gamma)$, $Z(Z \to b\bar{b})H(H \to \gamma\gamma)$, $t\bar{t}H(H \to \gamma\gamma)$, and $t\bar{t}$ (with two electrons faking photons) which have $\sigma \times BR$ of 111, 0.124, 0.04, 1.71 and 5×10^5 fb respectively, compared to 0.087 fb for the signal.

The energies of the final state particles and jets are smeared based on parameterisations extrapolated to the upgraded detector and high luminosity pile-up. For photons, a smearing of the direction is also applied. The expected photon identification efficiency is around 80% and the b-tagging efficiency is between 70 and 80%.

[2] Cross-sections at NLO calculated using the HPAIR package [97]. Theoretical uncertainties provided by Michael Spira in private communications.

[3] Event files produced by Dolan, Englert and Spannowsky as described in [98].

[4] The $b\bar{b}b\bar{b}$ final state has the highest branching ratio, but is expected to be too difficult to extract from the huge background

Events are selected that contain two b-jets and two photons with $50 < M_{b\bar{b}} < 130$ GeV and $120 < M_{\gamma\gamma} < 130$ GeV respectively, where the following object definitions are used. For b-jets: b-tagged and $p_T > 40/25$ GeV for the leading/sub-leading jet. For photons: $p_T > 25$ GeV, and fulfilling an isolation requirement. Additionally, cuts are applied on the angles between $b-b$, $\gamma-\gamma$ and $b-\gamma$ pairs, based on those described in [99]. Finally, a lepton veto and a jet multiplicity cut are applied.

Following this selection, a signal yield of approximately 11 events is obtained, with the irreducible $\gamma\gamma b\bar{b}$ background sample of initially more than 300,000 events suppressed to a contribution of around 1 event in 3000 fb^{-1}, the $b\bar{b}H(H \to \gamma\gamma)$ and $t\bar{t}$ (assuming the $e \to \gamma$ fake-rate to be 1%) giving approximately 1 event each, and approximately 3 events coming from $Z(Z \to b\bar{b})H(H \to \gamma\gamma)$. The contributions of the reducible backgrounds $\gamma\gamma jj$ and $jjjj$ are also small since these events populate mainly the low $m_{\gamma\gamma}$ region. The most significant remaining background is $t\bar{t}H$, contributing approximately 14 events. The overall background contribution is approximately 19 events. This corresponds to a S/B ratio of around 0.6 (or $S/\sqrt{B} \sim 2.4$).

In addition to the SM value $\lambda_{HHH} = 1$, the study was repeated using signal samples with $\lambda_{HHH} = 0$ and $\lambda_{HHH} = 2$. Yields of approximately 18 and 6 events were obtained for the $\lambda_{HHH} = 0$ and $\lambda_{HHH} = 2$ cases respectively. With this decay mode alone first evidence of double-Higgs boson production can be obtained but the Higgs self-coupling cannot be established.

In summary, preliminary studies of the $HH \to b\bar{b}\gamma\gamma$ channel indicate that a sensitivity for double Higgs boson production of $\sim 3\sigma$ per experiment is within reach if small improvements in the S/B ratio can be achieved. Additional channels such as $HH \to b\bar{b}\tau^+\tau^-$, the subject of a promising recent phenomenological study [98], are currently under investigation. Assuming that other channels contribute a combined significance comparable to our current estimate for $HH \to b\bar{b}\gamma\gamma$ and combining the two experiments, a $\sim 30\%$ measurement of λ_{HHH} may be achieved at the HL-LHC. Further studies with more realistic simulations are needed and will be made to consolidate and expand this preliminary result.

8.3 Weak boson scattering

A major reason for new physics to occur at around the TeV energy scale has been the prediction that an untamed rise of the weak boson scattering (WBS) cross-section in the longitudinal mode would violate unitarity at this scale. In the SM it is the Higgs particle which is responsible for its damping. Alternative models such as Technicolour and little Higgs have been postulated which encompass TeV-scale resonances and a light scalar particle. Other mechanisms for enhancing WBS at high energy are possible, even after the SM Higgs mechanism is established. The measurement of the energy dependence of the WBS cross-section is therefore a task of principal importance, which may also lead to unexpected new observations.

WBS can be parameterised by an effective Electroweak Chiral Lagrangian (EWChL), which includes two new operators that conserve the custodial SU(2) symmetry [100]. These new operators, which induce anomalous quartic couplings that are not strongly constrained by precision electroweak measurements, are scaled by numerical coefficients a_4 and a_5. The experimental sensitivity to weak boson scattering has been calculated using two separate approaches to unitarising the scattering amplitudes from the EWChL.

In previous studies [86], the sensitivity to new high-mass resonances was investigated using calculations based on the inverse amplitude method (IAM) of unitarisation. The model of Dobado *et al.* [101], implemented in the PYTHIA [102] generator, parameterises the anomalous WBS in terms of a_4 and a_5 and uses the IAM to induce a single new resonance to unitarise the scattering amplitude. This is a Higgs-less model, where the resonance mass, width, and couplings are fully determined by the coefficients a_4 and a_5.

In this document, a second approach is pursued [88], based on a scheme that merges high-mass resonances with the low-energy behaviour of the EWChL using a minimal K-matrix unitarisation method [103]. This method is implemented in the WHIZARD generator [104], which was used to generate weak boson scattering mediated by a new high-mass resonance in addition to a 126-GeV Higgs boson.

The fully leptonic $ZZjj \to \ell\ell\ell\ell jj$ channel has a small cross-section but provides a clean, fully-reconstructible ZZ resonance peak. A leading jet-jet mass requirement of 1 TeV reduces the contribution from jets accompanying non-WBS diboson production. Figure 8.3 shows the jet-jet invariant mass distribution and the reconstructed 4-lepton invariant mass distribution. The high-mass resonance is easily visible in this simulated dataset normalised to $3000\,\text{fb}^{-1}$.

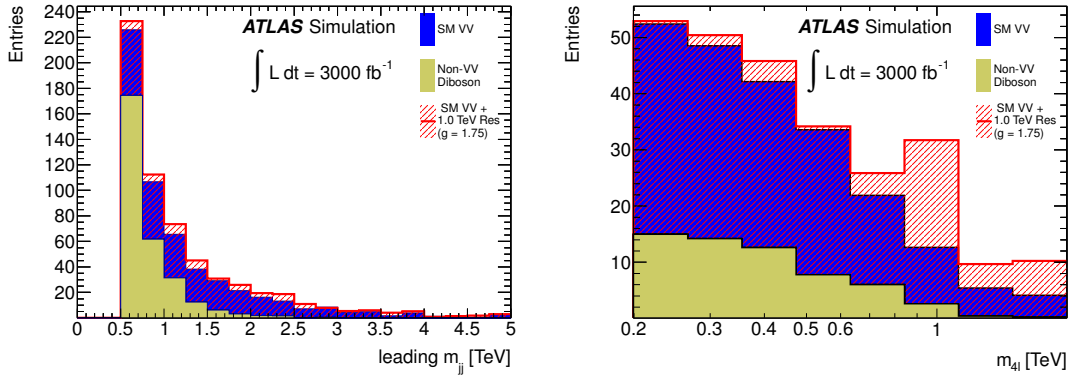


Figure 8.3: The leading jet-jet invariant mass (m_{jj}) distribution for simulated events in the $pp \to ZZ+2j \to \ell\ell\ell\ell+2j$ channel (left), and the reconstructed 4-lepton mass ($m_{4\ell}$) spectrum for this channel after requiring $m_{jj} > 1$ TeV (right). The WBS events are generated using WHIZARD without and with a ZZ resonance mass of 1 TeV and coupling $g = 1.75$, in addition to a Standard Model 126 GeV Higgs boson. The non-WBS diboson background is generated using MADGRAPH [105].

Table 8.16 shows the statistical significance of potential resonant signals given the background-only hypothesis, for a number of resonance masses and couplings in the $ZZ \to 4\ell$ channel. The increased luminosity of the HL-LHC is needed to push the statistical significance beyond the discovery threshold.

In terms of measuring the integrated cross-section for the purely-electroweak SM process $pp \to ZZ+2j \to 4\ell+2j$, a statistical precision of 10% is achievable with $3000\,\text{fb}^{-1}$, compared to 30% with $300\,\text{fb}^{-1}$ in the kinematic region of $m_{jj} > 1$ TeV and the 4-lepton invariant mass $m_{4\ell} > 200$ GeV.

Table 8.16: Summary of the expected sensitivity to anomalous WBS signal, quoted in terms of the background-only p_0-value expected for signal+background. The p_0-value has been converted to the corresponding number of Gaussian σ in significance. These results are given for the $pp \to ZZ + 2j \to \ell\ell\ell\ell + 2j$ channel at $\sqrt{s} = 14$ TeV. The increase in significance with integrated luminosity is shown for different resonance masses and couplings g.

model	$300\,\text{fb}^{-1}$	$3000\,\text{fb}^{-1}$
$m_{\text{resonance}} = 500$ GeV, $g = 1.0$	2.4σ	7.5σ
$m_{\text{resonance}} = 1$ TeV, $g = 1.75$	1.7σ	5.5σ
$m_{\text{resonance}} = 1$ TeV, $g = 2.5$	3.0σ	9.4σ

8.4 Supersymmetry searches and measurements

The discovery (or exclusion) of weak scale supersymmetry (SUSY) remains among the highest priorities for the LHC experiments. By introducing supersymmetric partners to all SM particles, SUSY removes the quadratic divergences that accompany a fundamental scalar Higgs boson. In the framework of generic *R*-parity conserving supersymmetric extensions of the SM, SUSY particles are produced in pairs and the lightest supersymmetric particle (LSP) is stable. The LSP can provide the massive thermal relic required to explain the cosmological dark matter. The lightest neutralino is often assumed as LSP, where neutralinos ($\tilde{\chi}_j^0$, $j = 1,2,3,4$) and charginos ($\tilde{\chi}_i^{\pm}$, $i = 1,2$) are the mass eigenstates originating from the mixture of the SUSY partners of Higgs and electroweak gauge bosons (higgsinos and gauginos). The scalar partners of right-handed and left-handed fermions can mix to form two mass eigenstates, for the studies presented here it is assumed that they are ($\tilde{q}$ and $\tilde{l}$), whilst possibly split in the case of bottom and top squarks (sbottom, $\tilde{b}$ and stop, $\tilde{t}$) and tau sleptons (stau, $\tilde{\tau}$). The lighter stop mass eigenstate can thus be significantly lighter than the other squarks and the gluinos ($\tilde{g}$, supersymmetric partners of the gluons). The world-leading constraints on the masses of many supersymmetric particles arise from LHC data at $\sqrt{s} = 7$ TeV and 8 TeV. Assuming a light LSP, the 1st and 2nd generation squark and gluino masses are currently excluded below about 1.4 TeV and 1.0 TeV, respectively [106]. Less stringent limits are placed on third generation squarks [107], gauginos [108] and sleptons (superpartners of leptons), where the mass constraints strongly depends on the assumed SUSY mass spectrum [108, 109]. The sensitivity to heavy SUSY particles will be increased significantly when the full LHC energy of $\sqrt{s} = 14$ TeV is reached. In the following, the mass reach for three types of SUSY searches, inspired by simplified standard analyses developed for 7 and 8 TeV data, is discussed [90]. Coarse systematic uncertainties of 30% on the background estimates are assumed to determine the reach. The sensitivity is expressed in terms of expected 5σ discovery reach and exclusion limits at 95% CL.

8.4.1 Squark and gluino searches

Generic searches for the production of squarks of the first two generations and gluinos are carried out in events with final states characterised by the presence of multiple jets and large missing transverse momentum. Figure 8.4(left) shows the discovery potential and exclusion reach achievable for $\tilde{q}$ and $\tilde{g}$ in an illustrative simplified model with a massless LSP. The HL-LHC improves

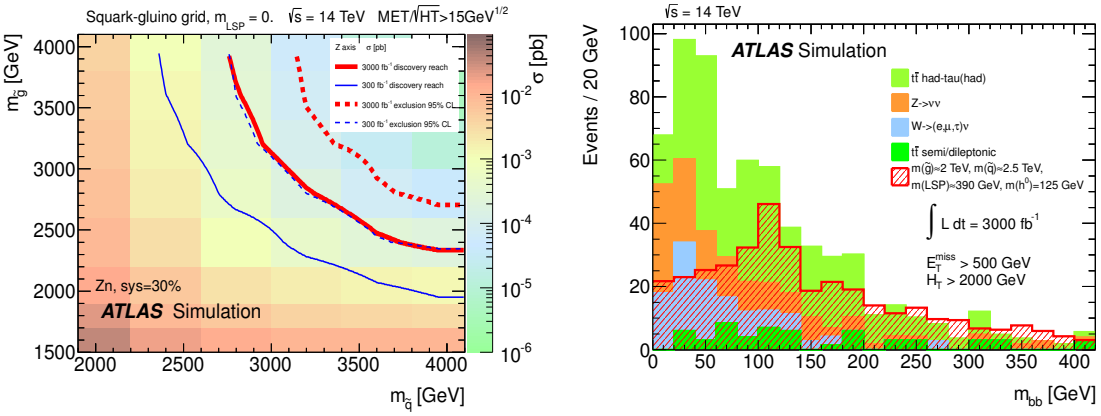


Figure 8.4: (left) The 95% CL exclusion limits (solid lines) and 5σ discovery reach (dashed lines) in a simplified squark–gluino model with massless neutralino with $300\,\text{fb}^{-1}$ (blue lines) and $3000\,\text{fb}^{-1}$ (red lines). The colour scale shows the $\sqrt{s} = 14\,\text{TeV}$ NLO production cross section calculated by `Prospino 2.1` [110]. (right) The m_{bb} invariant mass distribution for a benchmark SUSY model compared to the SM background processes for $3000\,\text{fb}^{-1}$ of integrated luminosity.

the sensitivity to both particles by approximately 400-500 GeV, a result confirmed in independent studies [111]. These results remain essentially unchanged for LSP masses up to about 1/3 of the mass of the strongly produced particle. If large deviations are observed with respect to the SM background estimates, dominated by $Z \to \nu\nu$+jets and $t\bar{t}$ production, the kinematic properties of the events can be studied and decay products in the decay chain of the SUSY particles can be identified. Figure 8.4(right) shows the m_{bb} invariant mass distribution for a benchmark SUSY model with squarks and gluinos decaying in complex final states including Higgs bosons, for which with $3000\,\text{fb}^{-1}$ the SUSY Higgs signal yield could be determined to $\approx$10%.

8.4.2 Third generation searches

Naturalness arguments [112, 113] require the top squark to be light, typically below 1 TeV. At $\sqrt{s} = 14\,\text{TeV}$ the direct stop pair production cross section for 600 GeV (1 TeV) stops is 240 (10) fb. The HL-LHC integrated luminosity will increase the sensitivity to stop or, if stop candidates are found, will allow their properties to be measured. Stops can decay in a variety of modes which typically includes top or b-quarks, W/Z or Higgs bosons, and an LSP. Pair production signatures are thus characterised by the presence of several jets, including b-jets, large missing transverse momentum and possibly leptons. Two studies are carried out targeting different decay modes and based on standard counting analyses. A one-lepton-based selection ($\ell \in \{e, \mu\}$) with stringent requirements on missing transverse momentum is employed to search for $\tilde{t} \to t + \tilde{\chi}_1^0$; in this case, the main SM background arises from $t\bar{t}$ production. A two-lepton-based selection ($e\mu$ only) is used for scenarios with $\tilde{t} \to b + \tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \to W^\pm + \tilde{\chi}_1^0$, where the dilepton m_{T2} [114, 115] variable is taken as the main discriminant against SM background processes dominated by top and W pair production. Figure 8.5(left) shows the discovery and exclusion potential versus the $\tilde{t}$ and $\tilde{\chi}_1^0$ masses in the two studies. The m_{T2} distribution for the two-lepton channel, useful to distinguish the SUSY signal from SM background processes, is shown in Fig. 8.5(right). A ten-fold increase in integrated

luminosity increases the sensitive stop mass range by up to 200 GeV. Further improvements in analyses techniques exploiting specific features that differentiate signal from SM background (*i.e.* missing transverse momentum shape, angular correlations, boosted objects) and taking into account additional final states can considerably extend the mass reach especially in case of heavy $\tilde{\chi}_1^0$.

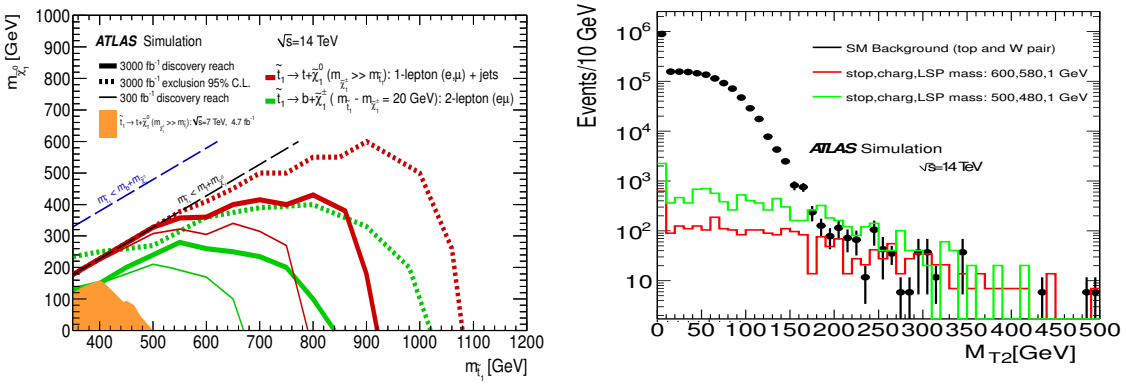


Figure 8.5: (left) The 95% CL exclusion limits for 3000 fb^{-1} (dashed) and 5σ discovery reach (solid) for 300 fb^{-1} and 3000 fb^{-1} in the $\tilde{t}$, $\tilde{\chi}_1^0$ mass plane assuming the $\tilde{t} \to t + \tilde{\chi}_1^0$ (red) or the $\tilde{t} \to b + \tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \to W + \tilde{\chi}_1^0$ (green) decay mode. (right) The m_{T2} distribution for two-lepton channel for SM background and 2 benchmark SUSY scenarios.

8.4.3 Electroweak Gaugino searches

Based also on naturalness arguments the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_j^0$ ($j = 1,2$) are expected to have masses in the hundreds of GeV range [112,113] and potentially be within the reach of the LHC. In scenarios with heavy squarks and gluinos, direct pair production of weak gauginos (and/or sleptons) dominates the SUSY production at the LHC. The cross-section of $\tilde{\chi}_1^\pm$-$\tilde{\chi}_2^0$ associated production ranges from 10 to 10^{-2} pb for masses between 50 and 600 GeV. The $\tilde{\chi}_1^\pm$ can decay as $\tilde{\chi}_1^\pm \to W^{\pm(*)}\tilde{\chi}_1^0$ whereas the $\tilde{\chi}_2^0$ decays as $\tilde{\chi}_2^0 \to Z^{(*)}\,\tilde{\chi}_1^0$, leading to final states with three leptons and missing transverse momentum where a pair of same-flavour and opposite-sign leptons has a mass consistent with that of a Z boson. In the analysis, a BR($\tilde{\chi}_1^\pm\tilde{\chi}_2^0 \to W^{(*)}\,\tilde{\chi}_1^0 Z^{(*)}\,\tilde{\chi}_1^0$) of 100% is assumed. The search is optimised using several kinematic variables to discriminate the signal from the two dominant background processes, namely top-pair and WZ production. With an integrated luminosity of 300 fb^{-1}, scenarios with chargino masses up to 500 GeV can be probed for $\tilde{\chi}_1^0$ mass values below 100 GeV. With a ten-fold increase in luminosity, the discovery potential is extended to scenarios with chargino masses values of ~800 GeV and $\tilde{\chi}_1^0$ masses below ~300 GeV (see Figure 8.6(left)). Figure 8.6(right) shows the missing transverse momentum for the SM background and two benchmark SUSY points for three lepton events surviving the final selection.

8.5 Exotics

The luminosity upgrade of the LHC substantially increases the potential for discovery and study of exotic new phenomena [116]. While the range of models and their respective parameters is quite large, their salient feature is the production of high-p_T leptons, photons, jets and missing E_T. Our

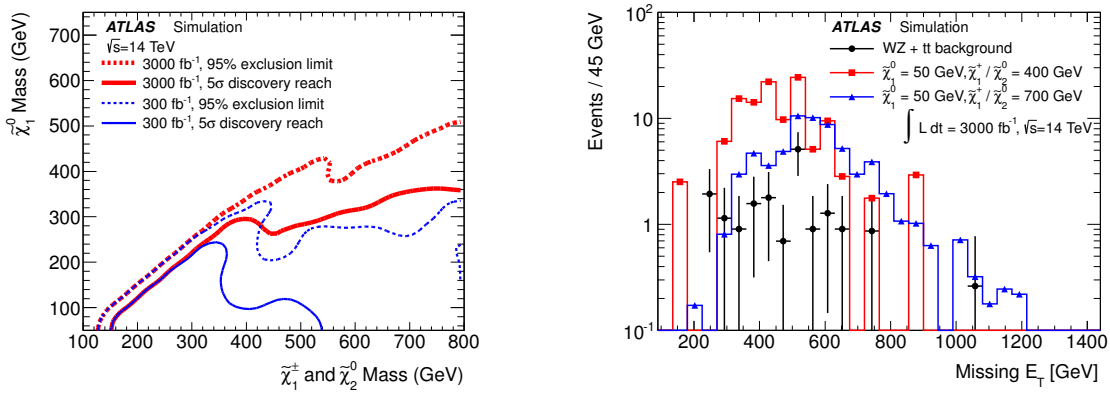


Figure 8.6: (left) The 95% CL exclusion limits (dashed lines) and 5σ discovery reach (solid lines) for charginos and neutralinos undergoing $\tilde{\chi}_1^{\pm}\tilde{\chi}_2^0 \to W^{(*)}\tilde{\chi}_1^0 Z^{(*)}\tilde{\chi}_1^0$ decays with BR=100%. The results using $300\,\text{fb}^{-1}$ and $3000\,\text{fb}^{-1}$ are reported. (right) The missing transverse momentum distribution in three-lepton events surviving the BDT-based selection.

goal is to ensure that the detector design maintains the sensitivity to any signature containing these characteristics. In this section, two benchmark models of new physics and the expected gain in sensitivity of the HL-LHC are discussed [91].

High mass resonances decaying into top quark pairs and into lepton pairs are studied to assess the discovery potential. The di-top resonances considered are a Kaluza-Klein gluon ($g_K K$) in the Randall-Sundrum model [117], [118], [119] and a Z' boson in the topcolour model [120]. The former interacts strongly and has a significant width while the latter interacts weakly and its width is narrrow compared to the detector resolution. In addition Z' decays to e^+e^- and $\mu^+\mu^-$ pairs are studied as an example of one of the simplest and most robust signatures of new physics [121].

8.5.1 $t\bar{t}$ resonances

Strongly- and weakly-produced $t\bar{t}$ resonances provide benchmarks not only for cascade decays containing leptons, jets (including b-quark jets) and missing E_T, but also the opportunity to study highly boosted topologies. Studies have been made of the sensitivity to the KK gluon via the process $pp \to g_{KK} \to t\bar{t}$ in both the dileptonic and the lepton+jets decay modes of the $t\bar{t}$ pair. These decay modes are complementary in a number of ways. The lepton+jets mode allows a more complete reconstruction of the final-state invariant mass, increasing the sensitivity of the search. This is particularly true for narrow resonances such as the $Z' \to t\bar{t}$ decay. On the other hand, this mode is more susceptible to W+jets background and to the loss of top quark identification when the top-jets merge at high boost. This mode also requires b-tagging to suppress a huge light-flavour background. The dominant background for the dilepton channel is the irreducible Standard Model (SM) $t\bar{t}$ production and b-tagging does not help to suppress the background, but it is less susceptible to the merging of top decay products because leptons, particularly muons, are easier to identify in close proximity to the b-jet. However the dilepton channel does not allow reconstruction of the resonance mass. By studying both channels, a conservative estimate of the sensitivity from the dilepton channel in comparison to the more sensitive lepton+jets channel is obtained.

Table 8.17: Summary of the expected limits for $g_{KK} \to t\bar{t}$ and $Z'_{\text{Topcolor}} \to t\bar{t}$ searches in the lepton+jets (dilepton) channel and of $Z'_{SSM} \to ee$ and $Z'_{SSM} \to \mu\mu$ searches in the Sequential Standard Model for pp collisions at $\sqrt{s} = 14$ TeV. All boson mass limits are quoted in TeV.

model	$300\,\text{fb}^{-1}$	$3000\,\text{fb}^{-1}$
g_{KK}	4.3 TeV (4.0 TeV)	6.7 TeV (5.6 TeV)
$Z'_{\text{Topcolour}}$	3.3 TeV (1.8 TeV)	5.5 TeV (3.2 TeV)
$Z'_{SSM} \to ee$	6.5TeV	7.8TeV
$Z'_{SSM} \to \mu\mu$	6.4TeV	7.6TeV

The search sensitivity for different integrated luminosity scenarios has been investigated. For the dilepton mode, the dominant background is from $t\bar{t}$ production, followed by Z+jets and diboson production. The analysis of current ATLAS data have shown that mis-identification backgrounds from W+jets and QCD multi-jets are not significant. The statistical analysis is performed by constructing templates of the H_T distribution for background plus varying amounts of signal at different resonance masses and cross sections. H_T is defined as the scalar sum of the transverse momenta of the two leading leptons, the two leading jets, and the missing E_T. The likelihood function is defined as the Poisson probability product over all bins for the pseudo-data given the expectation in each bin. The Z+jets and diboson background normalisations are given by the theory cross section, while the $t\bar{t}$ background is floated and is effectively constrained by the low-H_T region. The resulting expected limits in the absence of signal, using statistical errors only, are shown in Table 8.17.

A similar study for the lepton+jets channel where the reconstructed $t\bar{t}$ mass spectrum is used to search for the signal was performed. In this channel, the dominant background: $t\bar{t}$ and W+jets, is considered whereas the diboson background is not significant. The sensitivity obtained from the statistical analysis is also shown in Table 8.17. The distribution of reconstructed resonance mass and the resulting limits as a function of g_{KK} pole mass for the lepton+jets channel are shown in Fig. 8.7.

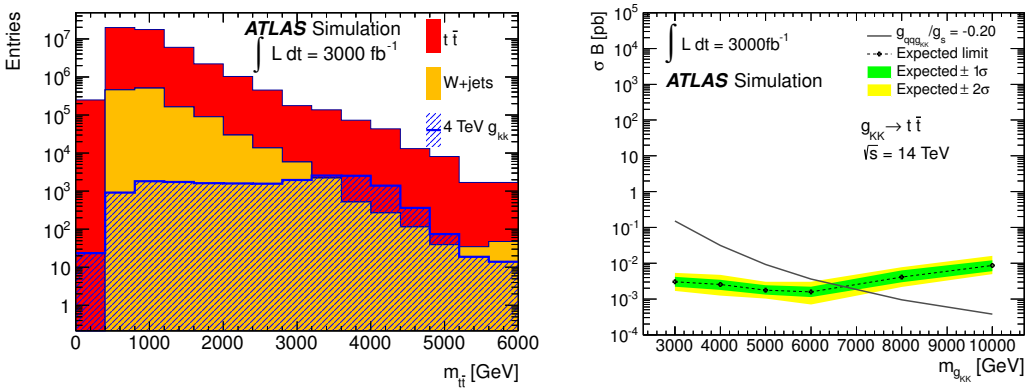


Figure 8.7: The reconstructed resonance mass spectrum in GeV (left) and limits (right) for the $g_{KK} \to t\bar{t}$ search in the lepton+jets channel with $3000\,\text{fb}^{-1}$ for pp collisions at $\sqrt{s} = 14$ TeV.

8.5.2 Dilepton resonances

The main issues in the detection and reconstruction of very high-p_T electrons and muons from heavy dilepton resonances such as Z' bosons and Randall-Sundrum gravitons are (i) the prevention of EM calorimeter response saturation for electrons due to the readout electronics, (ii) maintaining muon momentum resolution at high p_T, and (iii) maintaining sufficient angular coverage to measure the spin of the resonance. In the sensitivity studies of the Z' boson, the dielectron and dimuon channels were treated separately since their momentum resolutions scale differently with p_T and the detector acceptances are different. The sensitivity analysis uses the same methodology that is used for the $t\bar{t}$ study [122] and the Z' search [121] with the current ATLAS data, which is the template-based likelihood fit of the dilepton mass spectrum. The background is dominated by the SM Drell-Yan production, while $t\bar{t}$ and diboson backgrounds are substantially smaller. In the electron channel, there is an additional background from jet-misidentification which needs to be suppressed with good rejection of photon conversions. It is assumed that the required jet rejection will be achieved with the upgraded detector. The sensitivity study is based on the Drell-Yan background, and the resulting sensitivity is shown in Table 8.17.

8.6 FCNC in top decays

The HL-LHC will provide several hundred million top-quarks for a programme of precision top physics, one example: a study of flavour-changing neutral currents (FCNC) decays, is presented here.

Although absent at tree level due to the GIM mechanism [123], FCNC top quark decays occur in the Standard Model at loop level. The branching ratios are, however, smaller than 10^{-12} [124–127], many orders of magnitude below the dominant decay mode into bW. There are several Standard Model extensions [128] such as quark-singlet models (QS), two-Higgs doublet (2HDM) and flavour-conserving two-Higgs doublet (FC 2HDM) models, the minimal supersymmetric (MSSM) model, SUSY with R-parity violation models, the Topcolour-assisted Technicolour model (TC2) [129] as well as models with warped extra dimensions (RS) [130], that significantly enhance the FCNC decay branching ratios, up to 10^{-4}. Here a model independent approach to top quark FCNC decays is performed using an effective Lagrangian [131–133]. Even if the LHC does not measure the top quark FCNC branching ratios, it can test some of these models or constrain their parameter space, and improve significantly the current experimental limits on the FCNC branching ratios. FCNC top-quark decays have been searched for in the past. The best current direct search limits are 3.2% for $t \to q\gamma$ [134] and 0.34% for $t \to Zq$ ($q = u, c$) [135].

Top quark pair production, in which one of the top quarks decays through the dominant Standard Model channel ($t \to bW$) and the other through a FCNC channel ($t \to q\gamma$, $t \to qZ$), is considered as signal. Several Standard Model processes are background to the present analysis: $t\bar{t}$ production (in which both top quarks decay via Wb), W+jets and Z+jets production, diboson production (WW, WZ and ZZ) and multi-jet production.

The sensitivity is evaluated selecting events as in [136] for the $t \to qZ$ channel and [137] for the $t \to q\gamma$ channel. For the $t \to q\gamma$ channel, the dominant backgrounds are $t\bar{t}$, Z+jets and W+jets events. For the $t \to qZ$ channel, the background is mainly composed of $t\bar{t}$, Z+jets and WZ events.

In the absence of FCNC decays, limits on production cross-sections are estimated and converted to limits on branching ratios using the SM $t\bar{t}$ cross-section. The HL-LHC expected limits at 95% CL for the $t \to q\gamma$ and the $t \to qZ$ channels, are in the range between 10^{-5} and 10^{-4}. Figure 8.8 shows the expected sensitivity in the absence of signal, for the $t \to q\gamma$ and $t \to qZ$ channels. Further improvements could come from the use of more sophisticated analysis discriminants.

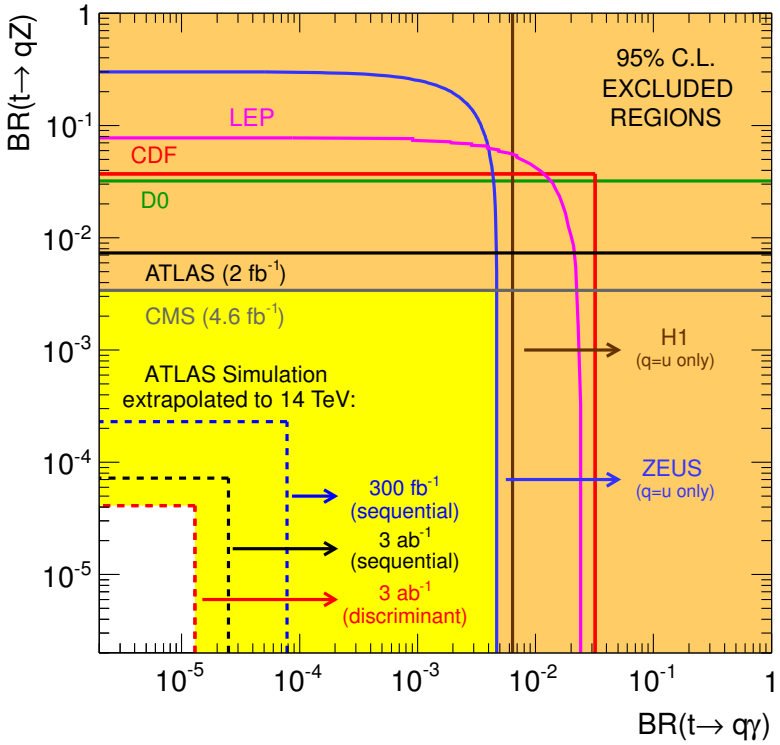


Figure 8.8: The present 95% CL observed limits on the $BR(t \to q\gamma)$ vs. $BR(t \to qZ)$ plane are shown as full lines for the LEP, ZEUS, H1, D0, CDF, ATLAS and CMS collaborations. The expected sensitivity at ATLAS is also represented by the dashed lines. For an integrated luminosity of $L = 3000\,\text{fb}^{-1}$ the limits range from 1.3×10^{-5} to 2.5×10^{-5} (4.1×10^{-5} to 7.2×10^{-5}) for the $t \to q\gamma$ ($t \to qZ$) decay. Limits at $L = 300\,\text{fb}^{-1}$ are also shown.

9. Installation and commissioning

9.1 Requirements

The LHC project is scheduling a long shutdown (LS3) in 2022 and 2023 to allow detectors and the accelerator to prepare for the conditions that will be met during the HL-LHC data taking period, which starts in 2024 with the aim of delivering an integrated luminosity of 2500 fb^{-1}over ten years.

The ATLAS cavern (UX15) will be accessible for about 27 months starting in December 2021. During this period, ATLAS will be opened to allow a series of modifications and improvements to the detector.

The mechanical operations required by the Phase-II upgrade are mainly driven by the insertion of a new Inner Detector (ITK).

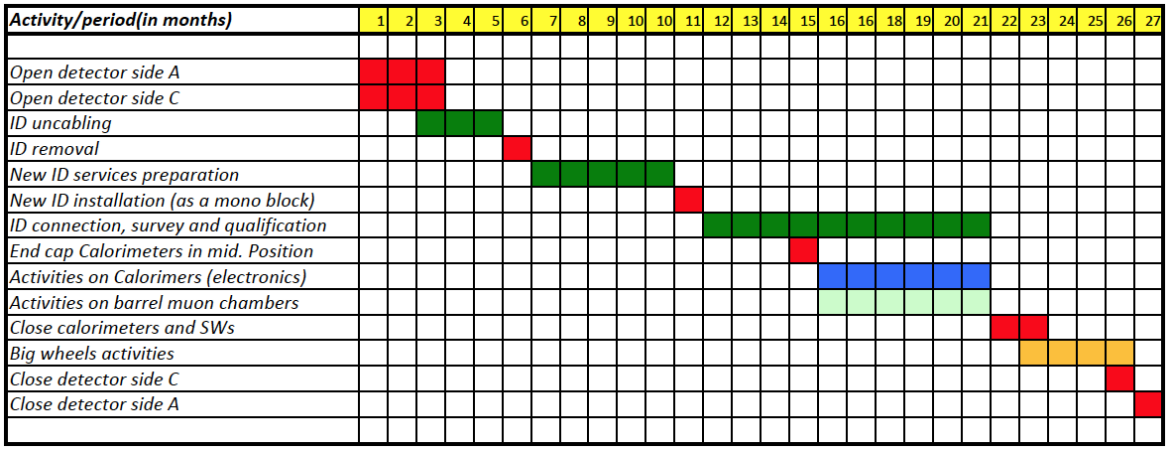


Figure 9.1: Items assigned to the various activities over the 27 months of access, scenario without considering the LAr end-caps cryostats opening

The old Inner Detector will be removed from the pit and brought to the surface for storage in a dedicated buffer zone at CERN for an initial radioactivity cool-down period until it can be decommissioned and/or disposed of. Before lifting it to the surface, all services must be disconnected. Because of the time criticality of this operation, the removal of the ID will get high priority at the beginning of the shutdown period over all other activities. After the removal of the old detector, the floor will be given to the refurbishing of the necessary services (pipes, fibres and cables). In all these operations the ALARA principle will be followed; the average radiation level is expected to be 10 μSv/h once the old ID is lifted to the surface.

To allow enough time for the refurbishing work, the newly built Inner Detector will be lowered into position approximately six months after the beginning of the shutdown. Only once the new ID is in place and re-cabled can other activities be started in parallel, such as the work for the muon barrel and calorimeters electronics upgrade. This work will be possible once the end-cap calorimeters are positioned back inside the toroid in its open position. During this period, the co-activity of all system-crews inside the Toroid area will need particular attention. The time between the installation of the ID and the repositioning inside the Toroid of the end-cap Calorimeters is

about 3 months, to allow the most intense cabling work and services installation to happen without disturbance and in optimal access conditions.

In this scenario it is assumed that there is no intervention necessary on the LAr end-cap cryostats (option 0 in chapter 3). The only other task requiring prime time, relates to the access to the muon big wheels, which are not fully accessible when the toroid end-caps are in the open position. It will be possible to schedule this task at the end of the shutdown period, when the end-cap toroids are in their running position. In this configuration there is no access to the ID end-plate region or to the calorimeters electronics. After a few months of work on the big wheels, the overall closing procedure will start.

9.2 Operation layouts and configurations

Figure 9.1 shows a schematic view of the time assigned to the various prime time activities. It is clear that all operations are very concentrated in time and that the schedule will be very challenging. To arrive at the main configuration, which allows the extraction of the old ID: the forward shielding and both muon Small Wheels will be moved to the surface; all beam pipes will be open, vented and moved to the surface; and the end-cap Toroids will be moved into their garage position at the side. The end-cap calorimeters will be moved by about 11.5 m longitudinally, just in front of the TAS Collimator, with no possibility to access the front-end electronics (see Figure 9.2). A dedicated floor, which will allow work and operations to be performed under good conditions in front of the barrel calorimeter front face, will be installed (mini vans) and dedicated scaffoldings will be erected as required. This scenario does not consider intervention required to install the new TAS, which will be necessary to accommodate a larger beam aperture.

All these opening operations are not new and will be exercised on many occasions over the coming years, by LS3 ATLAS technical coordination will have acquired considerable experience of these operations. The most critical and risky operation is the removal of the small wheels which will need to be brought to the surface. In the LS1 shutdown (ATLAS Phase-0 in 2013) this operation will be needed and therefore exercised and the specific technique should already be defined and qualified for the Phase-II needs.

The plan is to arrive to this main working configuration after 11 weeks of activity, working partially in 2 shifts. During all this period, for safety reasons, all other activities in UX15 will be very limited and strictly controlled.

9.3 LAr end-cap activities, additional scenario

As described in chapter 3.2, depending on the progress made in the understanding of the ageing of the HEC electronics and the degradation of the performance of the FCAL Calorimeter at high luminosity, several additional scenarios will need to be considered. Figure 9.3 refers to the extreme case where the HEC cold electronics needs to be exchanged and a new cold FCAL inserted. The end-cap calorimeters assembly will need to be positioned at the end of the barrel toroid, still on the main rails. A dedicated clean room will need to be constructed all around, preventing most of the other activities from being performed in parallel. The work on both calorimeters can be staggered by 3 months, for an individual length of operation of typically 9 months. The entire operation is very delicate and requires many complex manipulations such as the cryostat opening, the extraction

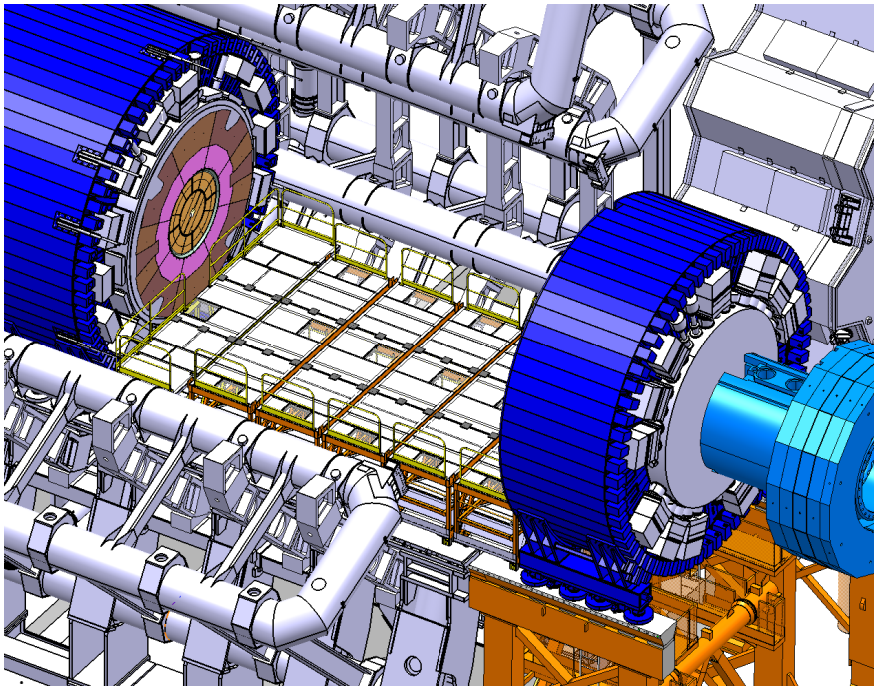


Figure 9.2: Detector configuration during LS3 shutdown, with both small wheels removed and access make for the new ID installation (large opening configuration)

of the two HEC wheels, and the final sealing of the cryostat flanges. All these activities will require a detailed ALARA analysis, since in this configuration the radiation levels that the intervention teams will be exposed to can reach values of 50-60 μSv/h. Particular attention will need to be given to avoid the contamination of the working area. All operations are done on the existing trucks at about 11 m above the cavern floor. As shown in Figure 9.3, this particular operation will add typically 9 months to the entire shutdown period, bringing the total time required for installation of the all the upgrades to 35 months. Before starting this operation enough time should be given to the ID for the cables refurbishment in the best and most confortable conditions (full access on both sides, large opening scenario). Intermediate scenarios, which do not require the opening of the HEC will be less time consuming and typically will be kept in to 30-32 months.

9.4 Radioprotection requirements

One of the new problems ATLAS will have to consider seriously in LS3 is the detector activation, after more then 10 years of beam exposure. In particular, all material near to the beam pipe (radius < 1.5m) will be confronted with such a problem. The beam pipe itself, the TAS collimator, the various shielding elements and the ID, will all be highly activated and any work nearby will need to be analyzed and optimized according to the ALARA principle.

CERN rules on radioprotection impose a maximal dose exposure of 6 mSv/year and 2 mSv/month for professional workers in a controlled area. This can be translated in a simple and pragmatic way to a maximal daily exposure of 60 μSv. Most of the detector areas will be confronted with such

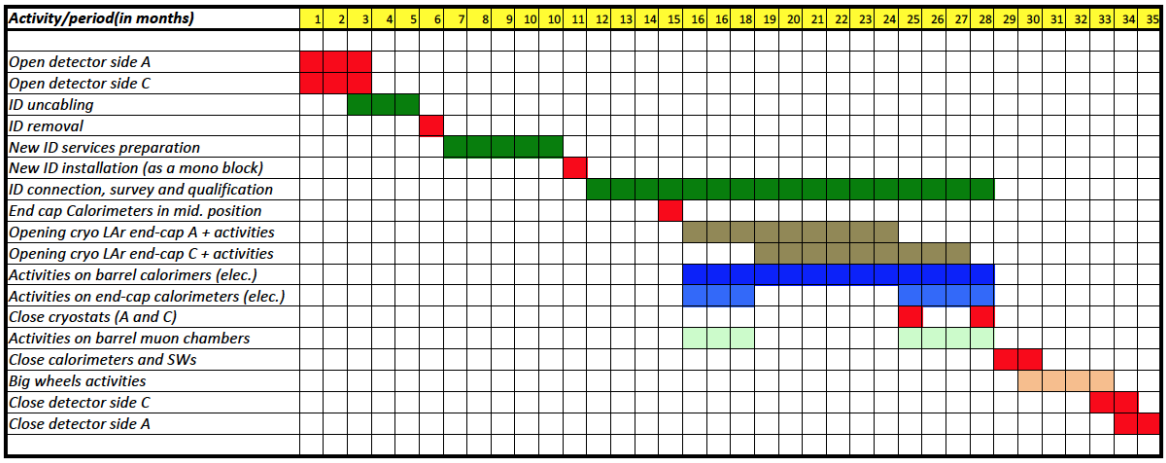


Figure 9.3: Items assigned to the various activities in case of the opening of the end-cap LAr cryostats

a challenge. For planning work, the ATLAS internal guideline for personal dose will be not to exceed 20 μSV/day/person.

During LS3, there will be 4 different situations to consider:

a) During the opening of the experiment and un-cabling of the inner detector. The opening procedure of ATLAS already takes into account the ALARA principle. The engineering procedures will be optimized to minimize exposure to the people involved in this activity. The intervention crew will be trained specifically for this work. More complex is the un-cabling procedure of the ID, where people will be working in front of the barrel for several weeks or even months, facing the ID. The beam pipes will be removed, but typically at 40 cm from the detector front face one can expect a radiation dose of 12-25 μSv/h depending on the location. This means that workers should limit their presence to a minimum number of hours per day and rotation of personnel during the 3 months allocated to this, prior to ID removal, is mandatory.

b) During ID removal and refurbishing of the ID services. During this stage, the ID activation will not be present anymore, no beam pipes and only a few services will remain. Access will be given to the inner bore of the Calorimeter (ID volume) to put in place new services and supports. It is expected that in this region and, similarly, at the front face of the Calorimeter, there will be dose rates of 5-10 μSv/h. However, this stage is much longer and therefore solutions again must be invented to minimise the dose to workers, using where possible individual protection, manipulators and rotation of personnel.

c) During opening of the LAr end-cap cryostats, in particular during the extraction and manipulation of the FCAL. This operation will require mechanical interventions on welds and structures and therefore one additional problem is the contamination of the entire working area. Here the dose might vary from 20 to 100 μSv/h depending on the location. Rotation of personnel, individual protection and special manipulators and procedures will be necessary and need to be part of the engineering design from the beginning.

d) Activities on the Big Wheels inner radius and just in front of the TAS and the JTT shielding. Here the dose is estimated to vary from 20 to 200 μSv/h, and any activity must be engineered such to protect people and minimize dose in an effective way.

For all 4 cases, a full ALARA procedure will be established and the work packages will need to be validated and approved by the CERN ALARA committee. A nuclear zone perimeter will be established at the beginning of the experiment by CERN/RP. All material exiting this perimeter must be, a priori, considered as radioactive and go through a special RP procedure which includes storage in a dedicated buffer zone and traceability for the entire life-cycle.

10. Resources

The collaboration decision and strategy is to express the costs associated with the activities described in this Letter of Intent in terms of CORE values.

The basic idea of CORE value is to focus on direct production costs (external costs) and exclude basic infrastructure and personnel costs. CORE costs thus includes items such as: components (but in general it does not include spares, except in the case of TDAQ), production (including industrial staff but not institute staff), outsourced parts of assembly, outsourced parts of installation, test and commissioning. CORE costs do NOT include items such as infrastructure (i.e. production area, R&D, design, early stage prototyping, institute staff, physicists, taxes and contingency).

The strategy is to rely on the concept of deliverables. Institutes and their Funding Agencies commit to provide given in-kind items, assemblies and detector parts as deliverable objects, after having agreed on the CORE value. What are defined as deliverables reflect the competences of the institutes providing them and will not be subject to central accounting and book-keeping.

Item	CORE cost (MCHF)	Possible additions	2015	2016	2017	2018	2019	2020	2021	2022
New Inner detector	131.500	26.000	2.400	5.600	35.660	32.460	29.160	15.360	10.860	0.000
LAr Calorimeter upgrades	32.124	15.096	0.547	3.170	1.015	2.003	4.517	14.379	6.494	0.000
Tile Calorimeter upgrades	7.483	2.517	0.000	0.000	0.000	1.122	1.629	4.070	0.602	0.060
Muon spectrometer upgrades	19.632	0.500	0.100	0.275	0.675	3.791	5.041	6.750	2.800	0.200
Trigger and DAQ upgrades	23.315	0.900	0.000	0.075	0.315	1.565	2.085	9.805	4.350	5.120
Common Fund	16.280	0.000	0.000	0.100	0.400	0.600	2.850	4.100	4.880	3.350
Total (MCHF)	230.334	45.013	3.047	9.220	38.065	41.541	45.282	54.464	29.986	8.730

Table 10.18: CORE Cost table

Table 10.18 shows the cost of the projects defined in this Letter of Intent. In the second column, each project CORE cost is expressed in MCHF units. In the third column some possible additional costs are anticipated. These are the costs of items which may still need to be part of the overall procurement. The final decision on these additional costs can only be taken at TDR stage, when the corresponding projects are better understood and defined. In the right-hand columns, the expenditure profile of the CORE cost is presented. It peaks during the years 2018-2020, when all major components will be procured. The total CORE cost amounts to 230.3 MCHF to be spent between 2015 and 2022. The cost of the additional options amounts to 45.0 MCHF.

About 16.28 MCHF out of the total CORE cost is of a common nature and will be handled centrally through a common fund, as was done for the original construction. The current estimate of 16.28 MCHF cover items related to infrastructure and shielding upgrades, installation work in the ATLAS pit and computing resources. At this stage in the project the value of the common funds have been kept to a minimum. As the projects evolve it is expected that several activities will emerge as common across the various systems, such as elements of the readout systems: the readout drivers (RODs) and the optical links, carbon fibre support structures, detector cooling systems, power supplies and integrated control systems. For these items, a common approach across the collaboration may be the most efficient way of funding them. It is expected that when the TDRs are

prepared, items that will benefit from a common approach will become evident and the fraction of the total cost identified as a common fund will be re-evaluated. This approach will give institutions that are not participating directly in a specific upgrade project a mechanism by which to contribute to the cost of the upgrade.

A more detailed cost table, with a break down item by item can be found through the following WWW-address: https://edms.cern.ch/document/1258343/1.

For each of the 5 projects and the common fund it is foreseen that a MoU document will be produced which will describe the duties of each institute involved in the construction, with a proper corresponding acknowledgement of the associated CORE value. Each MoU will then be signed by each Funding Agency involved and will define the time profile for each delivery. Before submission of the MoUs, ATLAS management will organize a proper scrutiny of the costs and of the related project milestones. Where needed by the relevant Funding Agencies to allow commitment of early resources, an interim MoU can also be established, to be superseded by the full MoU once that is adopted.

The cost figures quoted in this Letter of Intent are in 2012 Swiss Francs. No possible contingency figures related to exchange rate, inflation predictions or genuine uncertainties in the cost estimations are taken into account at this stage.

11. Conclusions

The High Luminosity LHC will provide the opportunity to probe the Higgs and EWSB sectors with unprecedented precision and significantly extend the reach of searches for new physics well in to the multi-TeV region, probing models such as SUSY, extra dimensions and compositeness. To be able to fully explore this new landscape it is required that the ATLAS experiment is upgraded to meet the experimental challenges at the HL-LHC: increased data rates, occupancy and radiation fluence; while preserving or improving the measurement of all physics objects: e, μ, τ, jets, γ, heavy flavour jets and E_T^{miss}.

Several physics studies using a parameterisation of the expected upgraded detector performance have been made. Rare Higgs decays, such as $ttH; H \to \gamma\gamma$ and $H \to \mu\mu$ can be measured, allowing the determination of the bosonic and fermionic coupling scale factors at the level of a few per cent. The first observation of Higgs self-coupling is accessible in the $HH \to b\bar{b}\gamma\gamma$ and $HH \to b\bar{b}\tau\tau$ channels. In weak boson scattering, the Standard Model process for a 126 GeV Higgs boson can be well measured and discovery of TeV scale di-boson resonances would become possible. Studies of possible searches for evidence of SUSY show that there are significant improvements in the range of parameters over which a discovery signal could be observed. Searches for new physics in $t\bar{t}$ and di-lepton resonances show that limits can be set at mass scales in the multi-TeV region. The upgrade programme described in this document presents the plans of the ATLAS collaboration to meet this challenge.

A concise list of the required upgrades is provided here:

- The radiation fluence accumulated by the end of 2022 and the high occupancy resulting from the HL-LHC luminosity makes installation of a new all silicon tracker mandatory. The new layout consists of pixel barrels and disks in the inner region, and strip barrels and disks in the outer region. The sensors are designed to cope with the high fluence, while the upgraded readout can handle the high occupancy and provide input to a level-1 track trigger.

- To preserve and improve the event selection capabilities, a new split level trigger architecture is proposed that can adapt to different experimental conditions and to new physics studies. The split level architecture will have a 500 kHz rate and ~ 6 μs latency capable Level-0 trigger based on inputs from: the Muon RPC/TGC trigger, possibly enhanced with MDT information, and the Calorimeter trigger. The Level-1 trigger will reduce the rate to 200 kHz within a latency of ~ 14 μs. This reduction will be principally accomplished by introducing a track trigger that will identify high p_T tracks, complementing information from the calorimeters and the muon system.

- The readout of the LAr and Tile calorimeters and the Muon system will be replaced, taking advantage of modern technologies to upgrade the on-detector electronics, and using large bandwidth optical links to move the data off-detector for processing, which will allow the use of FPGAs leading to more flexible and robust trigger and data acquisition systems.

- The developments in computing and software must address the experimental issues of larger data sizes and increasing event complexity and be able to adapt to changes in technology. The new simulation and reconstruction software will need to be optimised to run efficiently

in terms of CPU, memory and I/O usage. This will run within a new framework that can adapt to new computing technologies, and can work within new distributed computing paradigms. The issues of data preservation and long-term access will need to be built into the new framework.

The plans for the installation have been developed based on the experience accumulated over many years and that will profit from the additional operations required for the Phase-I upgrade. The number of steps for the upgrade can currently be accommodated within the programmed duration of the shutdown. However, if the ongoing studies on the degradation of the HEC and FCAL show a strong need for further actions, this would then require an extra nine months. The impact on the physics performance will have to be carefully weighed against the risks associated with this major additional intervention. The current understanding of the costs, and their breakdown by year and project, has also been presented.

The requirements to operate ATLAS for an additional decade and under conditions of unprecedented proton-proton collision rates have been outlined. Technical solutions have been proposed to meet the many resultant experimental challenges, giving access to a rich and unique physics programme at the energy frontier of particle physics.

References

[1] L. Rossi, O. Brüning et al. High-Luminosity Large Hadron Collider; A description for the European Strategy Preparatory Group. *CERN-ATS-2012-236*, 2012.

[2] The ATLAS Collaboration. Letter of Intent for the Phase-I Upgrade for the ATLAS Experiment. *CERN-LHCC-2011-012*, 2011.

[3] ATLAS Collaboration. FTK: a hardware track finder for the ATLAS trigger Technical Proposal, 2010. ATL-D-ES-0036; https://edms.cern.ch/document/1064724/1.0.

[4] G. Deputch et al. Development of 3D Vertically Integrated Pattern Recognition Associative Memory (VIPRAM). *FERMILAB-TM-2493-CMS-E-PPD-TD*, 2011.

[5] M. Garcia-Sciveres et al. System concepts for doublet tracking layers, Proceeding from Workshop on Intelligent Trackers, (WIT 2010). *JINST.*, 5:C10001, 2010.

[6] A. John et al. A Self Seeded First Level Track Trigger for ATLAS, Proceeding from Workshop on Intelligent Trackers, (WIT 2012), to appear in JINST .

[7] N.J. Buchanan, L. Chen, D.M. Gingrich, S. Liu, H. Chen, et al. Radiation qualification of the front-end electronics for the readout of the ATLAS liquid argon calorimeters. *JINST*, 3:P10005, 2008.

[8] N.J. Buchanan et al. The ATLAS Liquid Argon Calorimeter Electronics. *JINST*, 3:P02010, 2008.

[9] N Dressnandt, F M Newcomer, S Rescia, and E Vernon. LAPAS: A SiGe Front End Prototype for the Upgraded ATLAS LAr Calorimeter. *ATL-LARG-PROC-2009-017*, 2009.

[10] J. Ban, H. Brettel, A. Cheplakov, W. Cwienk, J. Fent, et al. Radiation hardness tests of GaAs amplifiers operated in liquid argon in the ATLAS calorimeter. *Nucl.Instrum.Meth.*, A594:389–394, 2008.

[11] The ATLAS Collaboration. Liquid Argon Calorimeter Technical Design Report. *CERN/LHCC/96-041*, 1996.

[12] D.M. Gingrich, G. Lachat, J. Pinfold, J. Soukoup, D. Axen, et al. Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter. *JINST*, 2:P05005, 2007.

[13] J. Ban, H. Brettel, W.D. Cwienk, J. Fent, L. Kurchaninov, et al. Cold electronics for the liquid argon hadronic end-cap calorimeter of ATLAS. *Nucl.Instrum.Meth.*, A556:158–168, 2006.

[14] S. Menke, Proton and Neutron Irradiation Tests of Readout Electronics of the ATLAS Hadronic Endcap Calorimeter, Proceedings of the IEEE/NSS 2012 in Anaheim, CA, arXiv:1211.4060 [physics.ins-det].

[15] R. Langstaff. Tooling specification for a HEC wheel electronics replacement cradle, 2010. ATU-SYS-ES-0011; https://edms.cern.ch/document/1175049/1.

[16] A. Artamonov, D. Bailey, G. Belanger, M. Cadabeschi, T.Y. Chen, et al. The ATLAS forward calorimeters. *JINST*, 3:P02010, 2008.

[17] J.P. Rutherfoord. Signal degradation due to charge buildup in noble liquid ionization calorimeters. *Nucl.Instrum.Meth.*, A482:156–178, 2002.

[18] B. Toggerson, A. Newcomer, J. Rutherfoord, and R.B. Walker. Onset of space charge effects in liquid argon ionization chambers. *Nucl.Instrum.Meth.*, A608:238–245, 2009.

[19] A. Glatte et al. Liquid argon calorimeter performance at high rates. *Nucl.Instrum.Meth.*, A669:47–65, 2012.

[20] M. Aleksa et al. Construction, assembly and tests of the ATLAS electromagnetic end-cap calorimeters. *JINST*, 3:P06002, 2008.

[21] J.Rutherfoord and R.B.Walker, Luminosity limits for liquid argon calorimetry, *J. Phys.*: Conf. Ser. 404 012016, 2012.

[22] M. Raymond. Time schedule for the replacement of the FCal, without opening of the cold cover, 2010. ATU-SYS-SC-0002, https://edms.cern.ch/document/1096694/1.

[23] Joshua Turner (on behalf of the ATLAS Liquid Argon Calorimeter Group), Upgrade Plans for ATLAS Forward Calorimetry for the HL-LHC, *Phys. Proc.* 37 301, 2012; John Rutherfoord (on behalf of the ATLAS Liquid Argon Calorimeter Group), Upgrade plans for the ATLAS Forward Calorimeter at the HL-LHC, *J. Phys.*: Conf. Ser. 404 012015, 2012.

[24] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[25] A. Marchioro. K. Kloukinas P. Moreira. The GBT, a Proposed Architecture for Multi-Gb/s Data Transmission in High Energy Physics, TWEPP-07: Topical Workshop on Electronics for Particle Physics, pages 332–336, 2007.

[26] M. Aleksa et al. Rate Effects in High-Resolution Drift Chambers. *Nucl.Instrum.Meth.*, A446:435, 2000.

[27] S. Horvat et al. Operation of the ATLAS Muon Drift-Tube Chambers at High Background Rates and in Magnetic Fields. *IEEE trans. on Nucl. Science Instr.*, 53, 2:562, 2006.

[28] Y. Arai et al. ATLAS Muon Drift Tube Electronics. *JINST*, 3:P09001, 2008.

[29] Y. Arai, AMT3 Users Manual, https://edms.cern.ch/file/897562/1/AMT3manual_034.pdf http://atlas.kek.jp/tdc/amt3/AMT3manual_034.pdf.

[30] J. Christiansen et al., Manual and User Application Notes for the HPTDC (2004), http://tdc.web.cern.ch/tdc/hptdc/docs/hptdc_manual_ver2.2.pdf http:/tdc.web.cern.ch/tdc/hptdc/hptdc.htm.

[31] P. Moreira et al. The GBT-SerDes ASIC prototype. *JINST*, 5:C11022, 2010.

[32] G. Aielli et al. Studies on fast triggering and high precision tracking with Resistive Plate Chambers. *Submitted to Nucl.Instrum.Meth., arXiv:1210.6696 [physics.ins-det]*, 2012.

[33] L. Pontecorvo. Study of the characteristics of single and double gap resistive plate chambers for a muon trigger at LHC. *Nucl. Instr. Meth.*, A344:105–109, 1994.

[34] J. Dubbert et al. Upgrade of the ATLAS Muon Trigger for the SLHC. *JINST*, 5:C12016, 2010.

[35] O. Sasaki et al. Design studies of the ATLAS muon Level-1 trigger based on the MDT detector for the LHC upgrade. *JINST*, 5:C12021, 2010.

[36] M Capeans, G Darbo, K Einsweiler, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. *CERN-LHCC-2010-013. ATLAS-TDR-019*, 2010.

[37] M. Keil. Operational Experience with the ATLAS Pixel Detector at the LHC. The 21st International Workshop on Vertex Detectors (Vertex 2012) to be published in *Proceedings of Science*.

[38] Y. Yacoob. Operational Experience with the ATLAS Strip Detector at the LHC. The 21st International Workshop on Vertex Detectors (Vertex 2012) to be published in *Proceedings of Science*.

[39] I. Dawson. Radiation background simulation and verification at the LHC: Examples from the ATLAS experiment and its upgrades. The 21st International Workshop on Vertex Detectors (Vertex 2012) to be published in *Proceedings of Science*.

[40] S. Gibson. Radiation damage to currently running LHC silicon detectors. The 21st International Workshop on Vertex Detectors (Vertex 2012) to be published in *Proceedings of Science*.

[41] ATLAS Collaboration. Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC. *New J.Phys.*, 13:053033, 2011.

[42] A. Clark, M. Elsing, and P. Wells. Performance Specifications of the Tracker Phase-II Upgrade. *ATL-UPGRADE-PUB-2012-003*, 2012.

[43] A. Clark, M. Elsing, N. Hessey, P. Mättig, N. Styles, P. Wells, S. Burdin, T. Cornelissen, T. Todorov, P. Vankov, I. Watson, and S. Wenig. Final Report: ATLAS Phase-II Tracker Upgrade Layout Task Force. *ATL-UPGRADE-PUB-2012-004*, 2012.

[44] I. Dawson, P. Majewski, and L. Nicolas. Neutron Moderator Studies for the ATLAS Inner Tracker Upgrade, 2009. ATU-TC-ER-0002; https://edms.cern.ch/document/982957/1.

[45] P. S. Miyagawa and I. Dawson. Radiation background studies for the Phase-II inner tracker upgrade. *ATL-UPGRADE-PUB-2013-012*, 2013.

[46] S Burdin et al. Tracking Performance of the Proposed Inner Tracker Layout for the ATLAS Phase-II Upgrade. *ATL-UPGRADE-PUB-2013-001*, 2013.

[47] S Burdin, M Garcia-Sciveres, N Hartman, and I Watson. Study of Conical Layout for Upgrade Pixel Detector. *ATL-UPGRADE-PUB-2013-003*, 2013.

[48] A. Ludwig, B. Heinemann, and M. Garcia-Sciveres. Studies of Inner Detector Layouts with 5 Pixel layers for the Phase-II Upgrade. *ATL-UPGRADE-PUB-2013-013*, 2013.

[49] P Delebecque, S Elles, N Geffroy, N Massol, T Rambure, and T Todorov. Alpine pixel detector layout. *ATL-UPGRADE-PUB-2013-009*, 2013.

[50] J. Weingarten et al. Planar Pixel Sensors for ATLAS Upgrade: Beam Tests results. *JINST*, 7:P10028, 2012.

[51] Christian Gallrapp (on behalf of ATLAS Planar Pixel Sensor R&D Project). Planar Pixel Sensors for the ATLAS tracker upgrade at HL-LHC. *arXiv:1206.3442 [physics.ins-det]*.

[52] C. Da Via et al. 3D silicon sensors: Design, Large area production and QA for pixel detector upgrades at the Large Hadron Collider. *Nucl. Instr. Meth.*, A694:321–330, 2012.

[53] M.Mikuz. Diamond sensors in HEP. 36th International Conference on High Energy Physics, ICHEP 2012, Submitted to *Proceedings of Science*.

[54] A. Lounis M. Benoit and N. Dinu. Simulation of guard ring influence on the performance of ATLAS pixel detectors for inner layer replacement. *JINST*, 4:P03025, 2009.

[55] Ivan Peric. A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology. *Nucl.Instrum.Meth.*, A582:876–885, 2007.

[56] M. Garcia-Sciveres, D. Arutinov, M. Barbero, R. Beccherle, S. Dube, et al. The FE-I4 pixel readout integrated circuit. *Nucl.Instrum.Meth.*, A636:S155–S159, 2011.

[57] M. Garcia-Sciveres et al. Towards Third Generation Pixel Readout Chips. International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (Pixel-2012), 2-7 September 2012, Inawashiro, Japan to appear in Nucl. Instr. Meth. A.

[58] S. Godiot et al. 3D electronics for hybrid pixel detectors. TWEPP-09: Topical Workshop on Electronics for Particle Physics, Paris, France, 21-25 sep 2009. p.220.

[59] L. Gonella et al. Experience with 3D integration techniques in the framework of ATLAS pixel upgrade for high luminosity LHC. International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (Pixel-2012), 2-7 September 2012, Inawashiro, Japan to appear in Nucl. Instr. Meth. A.

[60] R. Bates et al. Planar pixel module development for HL-LHC ATLAS pixel system. International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (Pixel-2012), 2-7 September 2012, Inawashiro, Japan to be published in Nucl. Instr, Meth.

[61] J. Pater et al. Pixel disk design and prototypes. *ATL-UPGRADE-PUB-2013-004*, 2013.

[62] P. Denes, R. Fly, and M. Garcia-Sciveres. A capacitor charge pump DC-DC converter for physics intrumentation. *IEEE Trans.Nuc.Sci.*, 56:1507, 2009.

[63] L. Gonella et al. A serial powering scheme for the ATLAS pixel detector at sLHC. *JINST*, 5:C12002, 2010.

[64] T. Stockmanns et al. Serial powering of pixel modules. *Nucl.Instrum.Meth.*, A511:S174, 2003.

[65] C. Bault et al. ATLAS upgrade tracker integration: Inner Tracker Support Structures, 2012. ATU-SYS-ER-0016; https://edms.cern.ch/document/1262222/1.

[66] C. Bault et al. ATLAS upgrade tracker integration: Inner Tracker Integration, Installation and Access, 2012. ATU-SYS-ER-0015; https://edms.cern.ch/document/1262221/1.

[67] NP Hessey. Building a Stereo-angle into strip-sensors for the ATLAS-Upgrade Inner-Tracker Endcaps. *ATL-UPGRADE-PUB-2013-002*, Feb 2013.

[68] Y. Unno, A.A. Affolder, P.P. Allport, R. Bates, C. Betancourt, et al. Development of n-on-p silicon sensors for very high radiation environments. *Nucl.Instrum.Meth.*, A636:S24–S30, 2011.

[69] L. Gonella, F. Faccio, S. Gerardin, et al. Total Ionizing Dose effects in 130-nm commerical CMOS technologies for HEP experiments. *Nucl.Instrum.Meth.*, A582:750–754, 2007.

[70] T. Affolder et al. System Electronics for the ATLAS Upgraded Strip Detector. *ATL-UPGRADE-PUB-2013-011*, 2013.

[71] Luis Amaral et al. The versatile link, a common project for super-LHC. *JINST*, 4:P12003, 2009.

[72] T. Affolder, I. Gregor and ITk-SC Module Working Group. Electrical Design and Performance of Single- and Double-Sided Silicon Strip Modules for the ATLAS Phase-II Upgrade. *ATL-UPGRADE-PUB-2012-005*, 2012.

[73] S. Diez on behalf of the ATLAS Upgrade Collaboration. System Implications of the Different Powering Distributions for the ATLAS Upgrade Strips Tracker. *Physics Procedia*, 37:960–969, 2012.

[74] A. Affolder, B. Allongue, G. Blanchot, F. Faccio, C. Fuentes, et al. DC-DC converters with reduced mass for trackers at the HL-LHC. *JINST*, 6:C11035, 2011.

[75] G. Beck et al. Phase 2 LOI Backup Document: Thermo-mechanical Local Support - Barrel Stave. *ATL-UPGRADE-PUB-2013-010*, Feb 2013.

[76] T. Jones et al. Phase 2 LoI Backup Document Silicon Strip Tracker R&D Status. *ATL-UPGRADE-PUB-2013-008*, Feb 2013.

[77] A. Clark, G. Barbier, F. Cadoux, M. Endo, et al. Development of a Silicon Microstrip Super Module Prototype for the High Luminosity LHC. *Nucl.Instrum.Meth.*, A699:S97–S101, 2013.

[78] S. Gonzalez-Sevilla, G. Barbier, F. Anghinolfi, F. Cadoux, et al. A silicon strip module for the ATLAS Inner Detector upgrade in the Super LHC collider. *Nucl.Instrum.Meth.*, A636:S97–S103, 2011.

[79] G. Barbier et al. Design and Assembly of Double-Sided Silicon Strip Module Prototypes for the ATLAS Upgrade Strip Tracker. *ATL-UPGRADE-PUB-2011-002*, 2011.

[80] S. Gonzalez-Sevilla, G. Barbier, F. Cadoux, A. Clark, et al. Electrical Performance of a Silicon Strip Super-Module Prototype for the High Luminosity LHC. *Nucl.Instrum.Meth.*, A699:S102–S106, 2013.

[81] B.T. Huffman et al. ITK optical links backup document. *ATL-UPGRADE-PUB-2013-005*, 2012.

[82] C. Bault et al. ATLAS upgrade tracker integration: Inner Tracker Common Items, 2012. ATU-SYS-ER-0014; https://edms.cern.ch/document/1262220/1.

[83] J. Godlewski et al. ATLAS upgrade tracker integration: Inner Tracker Cooling, 2012. ATU-SYS-ER-0017; https://edms.cern.ch/document/1262223/1.

[84] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.

[85] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.

[86] ATLAS Collaboration. Physics at a High-Luminosity LHC with ATLAS. *ATL-PHYS-PUB-2012-001*, 2012.

[87] ATLAS Collaboration. Physics at a High-Luminosity LHC with ATLAS - Update. *ATL-PHYS-PUB-2012-004*, 2012.

[88] ATLAS Collaboration. Studies of Vector Boson Scattering with an Upgraded ATLAS Detector at a High-Luminosity LHC. *ATL-PHYS-PUB-2012-005*, 2012.

[89] ATLAS Collaboration. Studies of the ATLAS potential for Higgs self-coupling measurements at a High Luminosity LHC. *ATL-PHYS-PUB-2013-001*, 2013.

[90] ATLAS Collaboration. Searches for Supersymmetry at the high luminosity LHC with the ATLAS Detector. *ATL-PHYS-PUB-2013-002*, 2013.

[91] ATLAS Collaboration. Studies of sensitivity to new dilepton and ditop resonances with an upgraded atlas detector at a high-luminosity lhc. (ATL-PHYS-PUB-2013-003), 2013.

[92] ATLAS Collaboration. Performance assumptions for an upgraded ATLAS detector at a High-Luminosity LHC. *ATL-PHYS-PUB-2013-004*, 2013.

[93] ATLAS Collaboration. Search for the Standard Model Higgs boson in the H to $\tau^+\tau^-$ decay mode in $\sqrt{s} = 7$ TeV pp collisions with ATLAS. *JHEP*, 1209:070, 2012.

[94] ATLAS Collaboration. Search for Neutral MSSM Higgs bosons in $\sqrt{s}$ =7 TeV pp collisions at ATLAS. *ATLAS-CONF-2012-094*, 2012.

[95] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. *CERN-2012-002, arXiv:1201.3084 [hep-ph]*, 2012.

[96] LHC Higgs Cross Section Working Group, A. David *et al.* LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle. *arXiv:1209.0040 [hep-ph], LHCHXSWG-2012-001*, 2012.

[97] HPAIR package: http://people.web.psi.ch/spira/proglist.html.

[98] Matthew J. Dolan, Christoph Englert, and Michael Spannowsky. Higgs self-coupling measurements at the LHC. *JHEP*, 1210:112, 2012.

[99] U. Baur, T. Plehn, and David L. Rainwater. Probing the Higgs selfcoupling at hadron colliders using rare decays. *Phys.Rev.*, D69:053004, 2004.

[100] Thomas Appelquist and Claude W. Bernard. Strongly Interacting Higgs Bosons. *Phys.Rev.*, D22:200, 1980.

[101] A. Dobado, M.J. Herrero, J.R. Pelaez, and E. Ruiz Morales. CERN LHC sensitivity to the resonance spectrum of a minimal strongly interacting electroweak symmetry breaking sector. *Phys. Rev.*, D62:055011, 2000. (and references therein).

[102] T. Sjöstrand et al. High-energy physics event generation with PYTHIA 6.1. *Comp. Phys. Comm.*, 135:238, 2001.

[103] A. Alboteanu, W. Kilian, and J. Reuter. Resonances and unitarity in weak boson scattering at the LHC. *arXiv:0806.4145*, 2008.

[104] Wolfgang Kilian, Thorsten Ohl, and Jurgen Reuter. WHIZARD: Simulating Multi-Particle Processes at LHC and ILC. *Eur.Phys.J.*, C71:1742, 2011.

[105] J. Alwell et al. MadGraph/MadEvent v4: The Web Generation. *JHEP*, 09:028, 2007.

[106] ATLAS Collaboration. Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV proton-proton collision data. *Accepted by Phys.Rev.*, 2012. arXiv:1208.0949 [hep-ex].

[107] ATLAS Collaboration. Search for direct top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 7$ TeV pp collisions using 4.7 fb^{-1} of ATLAS data. *ATLAS-CONF-2012-073*, 2012.

[108] ATLAS Collaboration. Search for direct production of charginos and neutralinos in events with three leptons and missing transverse momentum in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS detector. *ATLAS-CONF-2012-077*, 2012.

[109] ATLAS Collaboration. Search for direct slepton and gaugino production in final states with two leptons and missing transverse momentum with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV. *ATLAS-CONF-2012-076*, 2012.

[110] W. Beenakker, R. Hopker, M. Spira, and P.M. Zerwas. Squark and gluino production at hadron colliders. *Nucl.Phys.*, B492:51–103, 1997.

[111] Howard Baer, Vernon Barger, Andre Lessa, and Xerxes Tata. Discovery potential for SUSY at a high luminosity upgrade of LHC14. *Phys.Rev.*, D86:117701, 2012.

[112] Riccardo Barbieri and G.F. Giudice. Upper bounds on supersymmetric particle masses. *Nucl.Phys.*, B306:63, 1988.

[113] B. de Carlos and J.A. Casas. One loop analysis of the electroweak breaking in supersymmetric models and the fine tuning problem. *Phys.Lett.*, B309:320–328, 1993.

[114] C.G. Lester and D.J. Summers. Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders. *Phys.Lett.*, B463:99–103, 1999.

[115] A.J. Barr, C.G. Lester, and P. Stephens. m_{T2}: the truth behind the glamour. *J.Phys.G*, G29:2343–2363, 2003.

[116] F. Gianotti, M.L. Mangano, T. Virdee, S. Abdullin, G. Azuelos, et al. Physics potential and experimental challenges of the LHC luminosity upgrade. *Eur.Phys.J.*, C39:293–333, 2005.

[117] Kaustubh Agashe, Alexander Belyaev, Tadas Krupovnickas, Gilad Perez, and Joseph Virzi. LHC Signals from Warped Extra Dimensions. *Phys.Rev.*, D77:015003, 2008.

[118] Ben Lillie, Lisa Randall, and Lian-Tao Wang. The Bulk RS KK-gluon at the LHC. *JHEP*, 0709:074, 2007.

[119] Ben Lillie, Jing Shu, and Timothy M.P. Tait. Kaluza-Klein Gluons as a Diagnostic of Warped Models. *Phys.Rev.*, D76:115016, 2007.

[120] Robert M. Harris, Christopher T. Hill, and Stephen J. Parke. Cross Section for Topcolor Z'_t decaying to $t\bar{t}$: Version 2.6. *arXiv:hep-ph/9911288*, 1999.

[121] ATLAS Collaboration. Search for dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev. Lett.*, 107:272002, 2011. (and references therein).

[122] ATLAS Collaboration. A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C72:2083, 2012. (and references therein).

[123] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev.*, D2:1285–1292, 1970.

[124] J. L. Diaz-Cruz, R. Martinez, M. A. Perez, and A. Rosado. Flavour changing radiative decay of the t quark. *Phys. Rev.*, D41:891–894, 1990.

[125] G. Eilam, J. L. Hewett, and A. Soni. Rare decays of the top quark in the standard and two Higgs doublet models. *Phys. Rev.*, D44:1473–1484, 1991. Erratum-ibid.D59:039901,1999.

[126] B. Mele, S. Petrarca, and A. Soddu. A new evaluation of the $t \to cH$ decay width in the standard model. *Phys. Lett.*, B435:401–406, 1998.

[127] J. A. Aguilar-Saavedra and B. M. Nobre. Rare top decays $t \to c\gamma, t \to cg$ and CKM unitarity. *Phys. Lett.*, B553:251–260, 2003.

[128] J. A. Aguilar-Saavedra. Top flavour-changing neutral interactions: Theoretical expectations and experimental detection. *Acta Phys. Polon.*, B35:2695–2710, 2004.

[129] G. Lu, F. Yin, X. Wang, and L. Wan. The rare top quark decays $t \to cV$ in the topcolor- assisted technicolor model. *Phys. Rev.*, D68:015002, 2003.

[130] G. Perez K. Agashe and A. Soni. Collider Signals of Top Quark Flavor Violation from a Warped Extra Dimension. *Phys. Rev. D*, 75:015002, 2007.

[131] C. Caso et al. Review of particle physics, 1998-1999. review of particle properties. *Eur. Phys. J. C*, 3(1-4):1–794, 1998.

[132] W. Hollik, Jose I. Illana, S. Rigolin, C. Schappacher, and D. Stockinger. Top dipole form-factors and loop induced CP violation in supersymmetry. *Nucl.Phys.*, B551:3–40, 1999.

[133] M. Beneke et al. Top quark physics, in: Proceedings of the "1999 CERN Workshop on SM physics (and more) at the LHC". *arXiv:hep-ph/0003033*, 2000.

[134] CDF Collaboration, F. Abe et al. Search for flavor-changing neutral current decays of the top quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 80:2525–2530, 1998.

[135] CMS Collaboration. Search for flavor changing neutral currents in top quark decays in pp collisions at 7 TeV. *Phys.Lett.*, B718:1252–1272, 2013.

[136] ATLAS Collaboration. A search for flavour changing neutral currents in top-quark decays in pp collision data collected with the ATLAS detector at $\sqrt{s} = 7$ TeV. *JHEP*, 1209:139, 2012.

[137] ATLAS Collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. *arXiv:0901.0512 [hep-ex]*, 2009.