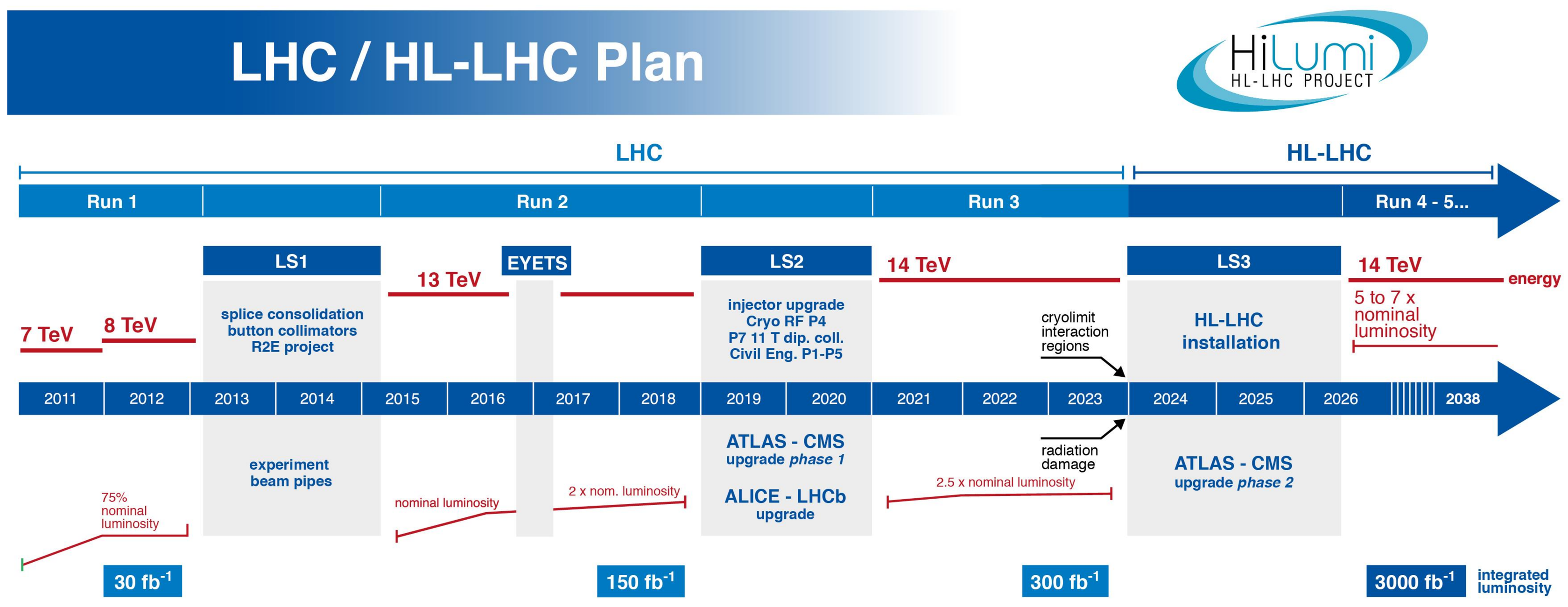
ATLAS Trigger and Data Acquisition Upgrades for the High Luminosity LHC

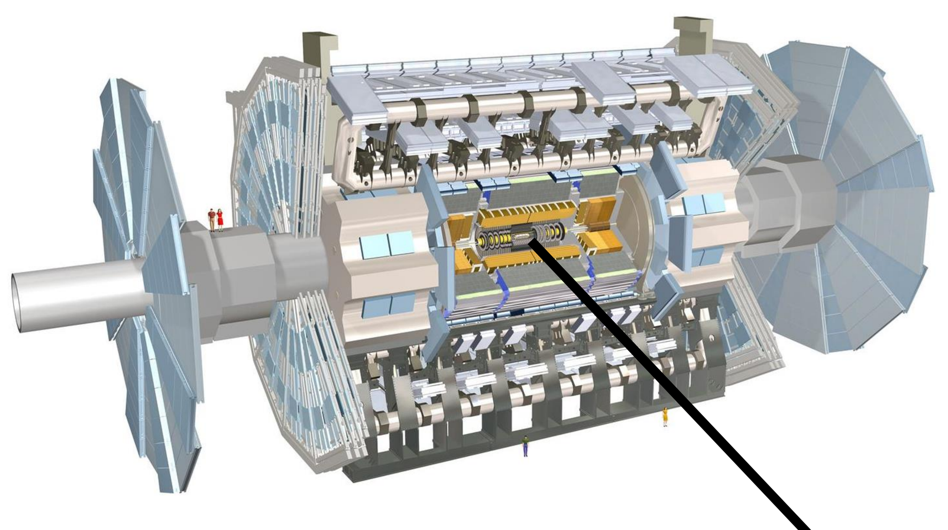
Authors: Eukeni Pozo Astigarraga on behalf of the ATLAS collaboration

High-Luminosity LHC

The Trigger and Data Acquisition system (TDAQ) of the ATLAS experiment [1] will be upgraded [2] for the High-Luminosity LHC (HL-LHC).

The HL-LHC is expected to start operations in the middle of 2026, and to reach ultimately a peak instantaneous luminosity of $L = 7.5 \times 10^{34}$ $cm^{-2}s^{-1}$, corresponding to approximately 200 inelastic proton-proton collisions per bunch crossing (pileup). In this configuration, more than ten times the integrated luminosity of the LHC Runs 1-3 combined will be delivered (up to 4000 fb^{-1}).





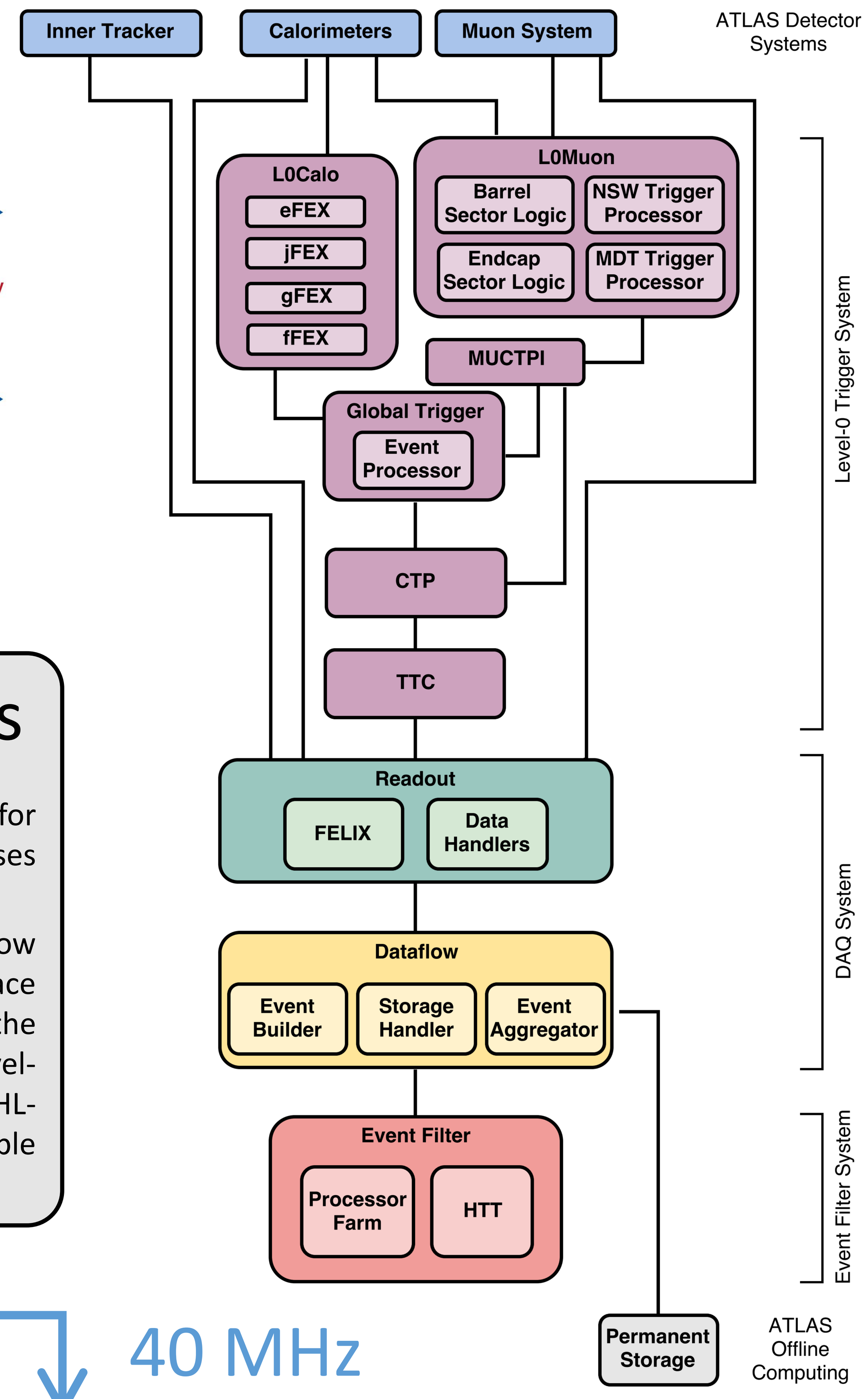
Physics Motivation

The LHC and ATLAS detector upgrades are designed to enable a broad physics program, including the study of the properties of the Higgs boson, precision Standard Model measurements, searches beyond the Standard Model and Flavour and Heavy-ion physics. The initial goals are described in [3] and [4].

Run 3 system limitations

The new conditions of the HL-LHC indicate a need for a factor of ten increase in the trigger rates which exposes fundamental limitations of the Run 3 TDAQ system.

On one side, the Run 3 detector readout, the dataflow components (network, storage, etc.) and the interface between both systems is not designed to cope with the updated throughput. On the other side, the Run 3 Level-1 Trigger system cannot be operated under the new HL-LHC rate and latency constraints without unacceptable loss of efficiency.



6 TB/s (Raw)

40 MHz

Readout

The Readout system receives data from the ATLAS detector Front-end Electronics (FE) at the L0-trigger rate (1 MHz) and performs basics processing before sending them to the Dataflow system. It forwards the TTC information to the FE.

FELIX

The Front-End LInk eXchange (FELIX) system provides a common interface to the custom detector FE via dedicated links. At 1 MHz L0-trigger rate and with an estimated event size of 6 MB, the FELIX system needs to sustain a 6 TB/s readout throughput.

In addition, FELIX is responsible for relaying the Trigger Timing and Control (TTC) information to the detector FE.

The FELIX system will be built on top of 300 commodity servers, hosting custom FPGA cards (connected via PCIe) where the more than 15,000 links coming from the FE are connected. Small subset of the FELIX readout will be commissioned already in the Phase-I system.

Readout Network

Data Handler

The Data Handler receives event fragments from the FELIX system and performs detector-specific formatting and monitoring tasks.

The FELIX server to Data Handler mapping is sliced on a per-subdetector basis. A high-throughput readout network interconnects the servers on a given slice, providing the possibility of a per-subdetector event aggregation.

The Data Handler system will be implemented on commodity servers and its size is estimated to more than 500 servers. Each of them connects to both the Readout and Dataflow networks.

L0-Accept

Level-0 Trigger System

The Level-0 Trigger System uses Calorimeter and Muon system information at 40 MHz to perform an initial event selection and to identify features to be examined at the subsequent trigger level. The maximum average L0-trigger rate is estimated to 1 MHz. The system makes extensive use of FPGAs and ATCA technology.

L0 Calo

Calorimeter data streams with coarse granularity are sent to the Feature Extractors (FEXs) for processing. The output of the L0 Calo subsystem is a set of trigger objects (electron, photon, tau lepton, etc.) that are sent to the Global Trigger.

L0 Muon

The L0 Muon trigger selects muons candidates and sends them to the Global Trigger via the MUCTPI component. The decision is based on the data from the upgraded muon spectrometer and the calorimeter input. Different detector technologies provide different angular coverage.

Global Trigger

The Global Trigger complements the L0 Calo trigger objects with additional high-granularity energy data coming directly from the upgraded calorimeter. It implements a different set of complex algorithms and sends the results to the CTP.

CTP

The Central Trigger Processor makes the final L0-Accept decision, aligning and combining several trigger inputs from different sources, including the Global Trigger and the MUCTPI. The L0-Accept signal is transmitted by the Trigger Timing and Control (TTC) system to the subdetector front-end electronics via FELIX.

6 TB/s (Processed)

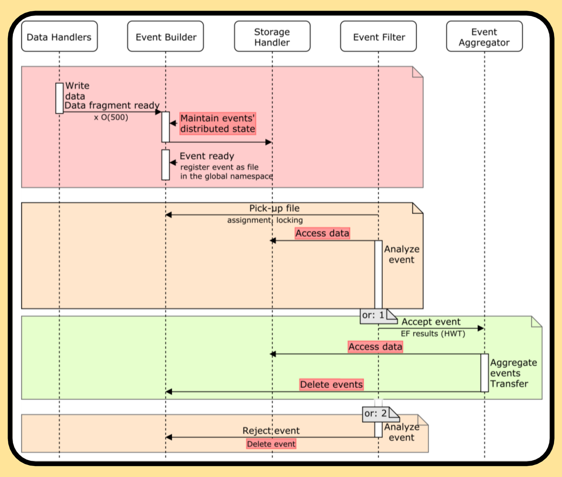
1 MHz

Dataflow

The Dataflow system buffers, transports, aggregates and compresses event data, exposing a simple interface to Readout and Event Filter systems. The main functional components are:

Event builder

The Event Builder is the logical interface between the Dataflow and Readout systems and the Dataflow and Event Filter systems. It is implemented as a software interface to the Storage Handler where the Data Handlers write the event fragments. Whether the storage is physically or logically built in the Storage Handler is a design choice to be taken.



Storage Handler

The Storage Handler decouples the Readout from the Event Filter (EF) system. It buffers data received from the Readout and serves them to the EF upon request.

In total, an aggregated I/O throughput of 7.8 TB/s is estimated. Assuming an SSD-based system and projecting current disk performance into 2025, the total system size can be estimated to 1800 SSDs, offering a total system capacity of 36 PB.

Event Aggregator

The Even Aggregator receives the selected events from the Event Filter and groups and compresses them before sending them to the Tier-0 permanent storage. It is implemented on top of the same hardware platform as the Storage Handler and provides a buffer area capable of storing up to 48 hours of accepted events.

DAQ Network

The network interconnects the Readout, Dataflow and Event Filter components and is the backbone of the Data Acquisition System.

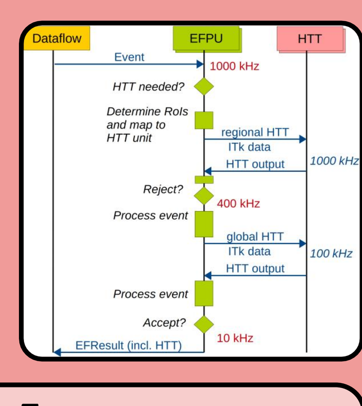
Based on modern commercial interconnection technologies, the network architecture has been designed in order to avoid traffic bottlenecks and to maximize the throughput of the system.

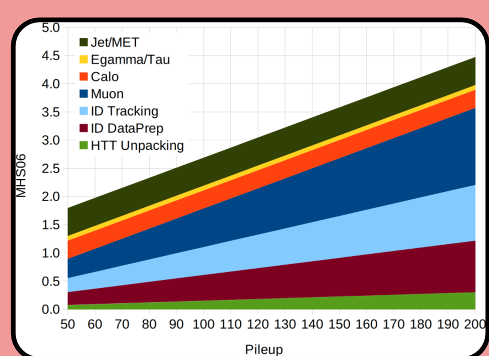
Network redundancy is considered in the design in order to minimize system downtimes.

Event Filter

The Event Filter (EF) system takes as input the detector data from events accepted by the preceding hardware trigger at 1 MHz. It must select the most useful events according to the trigger menu at a rate of 10 kHz and reject the rest.

The EF Processing Units (EFPUs) preform an initial selection combining information from the Level-0 hardware trigger with the regional hardware tracking in order to reduce the event rate down to around 400 kHz. This is followed by the full reconstruction of events and the execution of more complex selection algorithms assisted by the global hardware tracking.





Processor Farm

The most cost-effective computing platform for the Event Filter software is the commodity PC server. The market trends show an increased core-count and the incorporation of co-processors (GPGPUs or FPGAs).

Current estimates of the evolution of event processing times indicate the need for 4.5 million HEP-SPEC06 (MHS06). The farm infrastructure must fit in the available fixed rack space.

Hardware Track Trigger

The Hardware Track Trigger (HTT) provides tracks to the Event Filter reducing substantially the processing requirements. There are two levels of processing:

The regional HTT (rHTT) finds track candidates by finding hits in the inner tracker that match precomputed patterns stored in Associative Memory (AM) ASICs. The matched hit combinations are then processed in FPGAs to extract the tracking parameters.

The global HTT (gHTT) extrapolates the tracks found in the AM step to all the remaining inner tracker layers performing a full track fit and achieving the best possible track parameter resolution.

60 GB/s

10 kHz

Tier-0 Permanent Storage

Selected physics events are submitted to the Worldwide LHC Computing Grid (WLCG) for further analysis. The WLCG is a multi-tier computing infrastructure distributed around the world. The Tier-0 facilities are located at CERN and provide permanent storage capabilities to the LHC experiments.

For the HL-LHC, the total storage needs for the LHC experiments are estimated to 500 PB/year. Several R&D programs are ongoing to study how to better use the available technologies and opportunistic computing resources (Grid, HPC, Cloud, volunteer, etc.).

References

[1] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider 2008 JINST 3 S08003 1-407.
[2] ATLAS Collaboration, Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. CERN, Geneva, 2017, CERN-LHCC-2017-020.
[3] ATLAS Collaboration, Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment, Tech. Rep. CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, Dec, 2012. https://cds.cern.ch/record/1502664.
[4] ATLAS Collaboration, ATLAS Phase-II Upgrade Scoping Document, Tech. Rep. CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep, 2015. https://cds.cern.ch/record/2055248.

29 October - 1 November 2018

Amsterdam, the Netherlands