Technical Specification

ATLAS Level-1 Calorimeter Trigger Upgrade

Topology Processor (L1Topo)

K. Bierwagen, J. Damp, C. Kahra, M. Palka, U. Schäfer, R. Simoniello

Draft

Version: 1.0

27 October 2017

17:30

Table of Contents

1 Introduction

This document describes the specifications for the upgrade of the Level-1 topology processor module (L1Topo) of the ATLAS Level-1 Calorimeter Trigger Processor (L1Calo) [1.1] . An L1Topo processor has initially been introduced into the ATLAS trigger in Phase-0 for Run-2 to improve trigger performance by correlating trigger objects (electromagnetic clusters, jets, muons) and global quantities.

The new L1Topo will be installed in L1Calo during the long shutdown LS2, as part of the Phase-1 upgrade, and it will operate during Run 3 (and early-on in run 4). It is built to be forward compatible to a split Level-0/Level-1 Phase-2 system and may remain in the system after the Phase-2 upgrade in LS3, dependent on the eventual trigger architecture in Phase-2.

The ATLAS Phase-1 Level-1 Trigger system comprises eFEX [1.3] , jFEX [1.4] , and gFEX [1.5] subsystems as calorimeter data sources for L1Topo. They are providing trigger object data, “TOBs”, to L1Topo via optical fibre bundles. Another source of trigger objects is the ATLAS muon trigger subsystem.

L1Topo is a set of three (dual-width) ATCA [1.6] [1.7] modules, operated in a single ATCA “shelf” (crate), compliant with ATLAS and L1Calo standards. Real-time data are received via optical fibres exclusively. L1Topo runs a large number of concurrent and independent algorithms on the input data, to derive a number of trigger bits, typically one result bit and one overflow bit per algorithm. The result bits are forwarded to the Central Trigger Processor (CTP), which correlates these bits with further trigger and machine data to generate Level-1 Trigger and associated data words, to be transmitted back to the detector. Outputs to the CTP are available via electrical and optical data paths.

The allocation of the three L1Topo modules (with a total of 6 L1Topo processor FPGAs) to trigger algorithm types is described in [1.8]

The non-real-time data paths of L1Topo are basically identical to the eFEX and jFEX built for Phase-1: data are sent into the readout and the 2nd level Trigger via L1Calo RODs over the backplane of the ATCA shelf. The data formatting will be compatible with the already chosen formats. Control and global timing are accomplished via the backplane as well. To that end, the L1Topo module communicates with two Hub/ROD [1.9] [1.10] modules located in dedicated slots of the L1Topo shelf.

The Phase-1 Level-1 trigger system and the role of L1Topo within the Level-1 Calorimeter trigger system is described elsewhere in detail. Material on current Phase-0 L1Topo construction and performance is available as well. References are given in section 7.

2 Functionality

Figure 1 shows a block diagram of L1Topo. The various aspects of L1Topo functionality are described in detail below. While the data paths are implemented unalterably on the L1Topo PCB, most of the functionality described here is implemented in programmable firmware only, and is not directly affected by hardware details. An overview of module connectivity is given in Table 1. Implementation details of L1Topo are given in section 3.

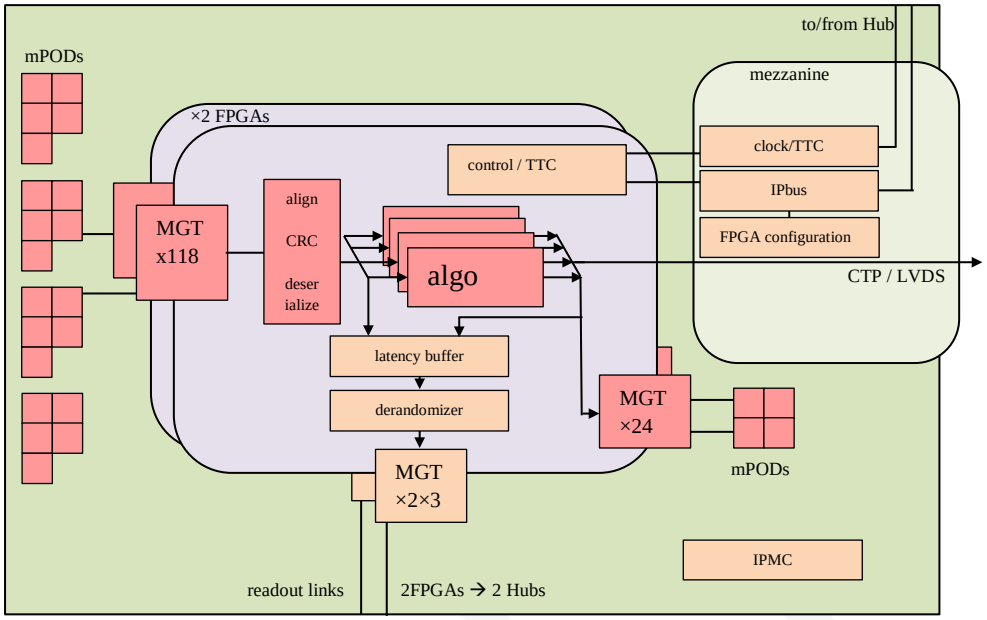

Figure 1. A block diagram of the L1Topo module

Number of input channels	2x118
Input data rate	6.4Gb/s, 11.2Gb/s and 12.8Gb/s
Inter FPGA parallel connectivity	64Gb/s
Maximum bandwidth to DROD	Up to 6 x 6.4 Gb/s
Maximum bandwidth to RROD	Up to 6 x 6.4 Gb/s
Real time output (CTP+Spare)	Up to 48 x (6.4 Gb/s (CTP),12.8 Gb/s (Spare))

Table 1 Overview of module interconnect

2.1 Real-Time Data Path

ATCA Backplane Zone-3 of L1Topo is used for real-time data transmission. The input data enter L1Topo optically from the back. The fibres are fed via four blind-mate backplane connectors (MTO-CPI) that carry 72 (or 48) fibres each. The optical signals are converted to electrical signals in 12-fibre receivers. For reason of design density, MiniPOD [1.11] receivers are used. The electrical high speed signals are routed into two FPGAs, where they are de-serialized in MGT receivers; the parallel data are presented to the FPGA fabric. The two FPGAs operate on their input data independently and in parallel. High bandwidth, low latency parallel data paths allow for real-time communication between the two processors. The signal results are transmitted towards the CTP on both optical fibres and electrical cables. The electrical signals are routed via an Extension Mezzanine.

2.1.1 Input Data

L1Topo will receive the topological output data of the sliding window processors from L1Calo and data from the L1Muon system. The data format transmitted into L1Topo comprises Trigger Object data for jets, clusters and muons, as well as energy sums. The data will consist of a description of the position of an object (jet, e/m cluster, tau and muons) along with some qualifying information, like the energy sum within the object.

2.1.2 Input Data Rates

So as to be compatible to the conflicting bitrate requirements of FEXes and Muon data sources (MUCTPI), the module will be built so as to support input data rates of either 6.4, 11.2 or 12.8 Gb/s on a given input channel. Since MGT input channels are organized in quads, with all four channels sharing clock generation, it is assumed that a given quad will be operated on one of the three bitrates only. For the relatively small number of channels that are also used for high speed output, the input bitrate might need to be chosen for compatibility with the output rate. That might create constraints for physical location of certain object types on the FPGA / on the fibre bundles.

2.1.3 Algorithms

Due to the large amount of logic resources in the chosen FPGAs, a significant number of algorithms is expected to be run on the real-time data in parallel.

The trigger menu will most probably be similar to the Run-2 menu and it is likely that many algorithms will be identical or very similar to the ones already introduced for Run-2.

A compact summary of all available Run-2 algorithms can be found in [1.13] .

In Run-2 there are some algorithms that are instantiated multiple times with different configurations to gather information about different cuts and being able to switch easily. Some of these algorithms might be dropped in Run-3. On the other side, several new and more complex algorithms can be added.

The procedure of organizing and configuring the algorithms is planned to be similar to Run-2. This means that as soon as most requested algorithms and the related parameters are known, a reasonable distribution of all algorithms over the available FPGAs has to be found. Due to the increased number of FPGAs and algorithms and the changes in the number and formats of the input TOBs it is not expected that the current distribution of algorithms can be reused.

For finding a new configuration, the main goal is to distribute the number of instantiated algorithms, as well as the allocated resources, approximately evenly across all FPGAs. This is a non-trivial task where no strict prescription exist. As soon as most desired algorithms are known, several configurations have to be tried so as to find the best one via trial and error. This configuration is then specified in the trigger menu along with the configurable parameters of the algorithms. In the end, the final version of the algorithm firmware top modules is generated automatically from the configuration in the trigger menu to ensure consistency.

Some changes are expected for topological triggers using the standard jets or missing energy, as it is possible that the corresponding algorithms benefit from the new globally built TOBs and quantities from the gFEX system.

In addition to the topological algorithms, the functionality of the CMX modules in the legacy Run-2 system (simple thresholding and multiplicity triggers) will be formed by algorithms in the Topological Processor. These “non-topological” triggers are based on algorithms with a single type of input TOBs and are expected to be simple and fast.

As in the current system, the algorithms will be flexible so that triggers with different thresholds but the same quantity can reuse the same algorithms, though in a separate instantiation.

2.1.4 Data Sharing

Topology data are processed in two separate FPGAs per module. There is no data duplication implemented at hardware level. The two processors can communicate via a parallel bus to get access to data that cannot be received directly via the multi-gigabit links. Though according to the device data sheets higher data rates should be possible, a maximum bit rate of 640 Mb/s per differential pair is anticipated for the inter-FPGA link, which is a convenient multiple of the bunch clock frequency. That will limit parallel connectivity to about 64 Gb/s of aggregate bandwidth (see section 3).

2.1.5 Output

The real-time output data of L1Topo to the CTP consist of individual bits indicating whether a specific algorithm passed or not, plus an overflow bit. The resulting trigger data are expected to exhibit a rather small volume. They will be transmitted to the CTP optically or electrically. A single fibre-optical ribbon connection per module that carries 48 fibres, running through the front panel of the module, is provided for this purpose. A mezzanine board will be required to interface L1Topo to the CTPCORE module electrically via 32 LVDS signals at low latency.

2.2 Error Handling

Input data are protected by several error detection schemes. The MGT hardware blocks can detect link errors and code errors. Additional protection is achieved by cyclic redundancy check characters included in the real-time data. Errors of all types will be monitored and the error counter will be incremented for any bunch clock cycle where there is at least one error in any input channel. Detailed information of the specific error will be stored in expert registers. Detection of an error will enforce zeroing the real-time data for the affected events.

2.3 Latency

The ATLAS Level-1 Trigger is a severely latency constrained system. The overall latency envelope is tightly controlled.

A breakdown of the estimated latency of the real-time path of the L1Topo system is given in the ATLAS TDAQ System Phase-1 Upgrade Technical Design Report [1.1] . Latency figures are kept up to date at [1.2] .

	ns	BCs	Sub Total	Total
Optical Input available from CMX, eFex, jFex & Muctpi				68,9
L1Topo Input Deserialisers	50	2,0		
Synchronize to local clock - 320-> 40 MHz	25	1,0		
Algorithmic Processing	125	5,0		
			8,0	
Electrical Output to CTP (multiplexed) (if used)	25	1,0		
Electrical Cable to CTP (if used) (2m)	10	0,4		
			1,4	
L1Topo electrical input available at CTP				78,3
Output Multiplexers 40-320 MHz (if used)	25	1,0		
Output Serialisers for optics (if used)	50	2,0		
Fibres to CTP (if used) (2m)	10	0,4		
			3,4	
L1Topo Optical inputs to CTP available			11,4	80,3

Figure 2. Latency estimate for L1Topo

Figure 2 shows the current latency estimates for L1Topo, as extracted from this document at the time of writing. Exact infrastructure latencies can only be measured once cabling and infrastructure firmware are in place and implemented.

Algorithm latency is assumed to be an upper limit. This might in turn limit the complexity of any future trigger algorithms. With an algorithmic latency of 3 LHC ticks on the Phase-0 Topology Processor, the headroom of another two ticks in phase 1 seems adequate. It should be noted that the latency critical path is assumed to be the muon subsystem. Therefore, any algorithms requiring muon data are assumed to have their results routed to the CTP at lowest possible latency, on electrical cables. In the current trigger menu, the most latency constrained trigger algorithm is the delayed muon trigger which commands a latency reduction of one bunch tick, so as to get late muon signals back into the bunch tick they are assumed to originate from. Therefore, for this specific trigger that part of the muon algorithms had to be (successfully) squeezed to a total algorithmic latency of two ticks.

Since calorimeter data are expected to arrive far earlier than muons, it is assumed that only the very last stages of algorithmic processing pertaining to muons are under severe latency restrictions. Any calorimeter date to be correlated with muons can be pre-processed under significantly relaxed constraints.

This scheme will maximize the usable latency budget, albeit at the cost of a more complex timing-in procedure.

2.4 Readout Data Path

Upon receipt of an L1Accept all L1Topo real-time output data will be captured and sent to the DAQ. Input data capture can be made dependent on possible occurrence of reception errors, or be fixed, software programmable. The number of slices worth of data per bunch tick is programmable. Data are pipelined and de-randomized on the processors and then serialized onto the backplane links to the ROD/Hub modules. Region-of-Interest data (RoI) can be captured separately and made available to the higher level triggers via the RoI builder, if required. In case the ROD will assert busy signal the readout will only finish any existing operations to form packets intended to be sent to the ROD. At the same time it will prevent sending any existing packets to the ROD and it will reject all incoming data when it exceed the capability to save new events. This will happen in situation when there will be not enough space in buffers and FIFOs.

Further details are found in section 3, the further processing of readout data in the RODs is described in [1.10] .

2.5 TTC and Clock

Timing signals are received in the L1Topo shelf via the Hub [1.9] module. There, the clock is recovered and commands are decoded, before being re-encoded using a local protocol. This use of a local protocol allows the TTC interface of the shelf to be upgraded to future timing distribution schemes without any modification of the L1Topo modules.

The L1Topo module receives the clock and TTC commands from the Hub module via the ATCA backplane. It receives the clock on one signal pair and the commands on a second (see section 3.11 for details). An auxiliary crystal clock will be available as well.

2.6 Module Control and Configuration

An IPbus interface is provided for high-level, functional control of L1Topo. This allows, for example, algorithmic parameters to be set, modes of operation to be controlled and spy memories to be read.

IPbus is a protocol that runs over Ethernet to provide register-level access to hardware. Here, it is run over a 1000BASE-T Ethernet link, which occupies one channel of the ATCA Base Interface. On L1Topo there is a local IPbus interface in every FPGA. These interfaces contain those registers that pertain to that device. A control FPGA, residing on a mezzanine, implements the interface between the topology processors and the shelf backplane, routing IPbus packets to and from the other devices as required. The control FPGA also contains those registers which control or describe the state of the module as a whole. For those devices such as MiniPODs, which have an I^2C control interface, an IPbus-I2C bridge is provided.

The processor FPGAs are configured upon power-up from flash based storage. The configuration data are clocked into the FPGAs via a parallel bus. Controller and flash memory are located on the mezzanine. For debug purposes the processors can be configured and accessed (Ibert, Chipscope ILA) via their JTAG interface.

2.7 Commissioning and Diagnostic Facilities

To aid in module and system commissioning, and help diagnose errors, L1Topo can be placed in Playback Mode via an IPbus command. In this mode, real-time input data to L1Topo are ignored and, instead, data are supplied from internal scrolling memories. These data are fed into the real-time path at the input to the algorithm logic, where they replace the input data from the FEXes and muons.

In spy mode, also selectable via an IPbus command, the scrolling memories can be filled with data received from the real-time inputs of L1Topo. The data captured can be read out via IPbus.

On the real-time output of L1Topo towards the CTP, the same playback/spy scheme is employed. By enabling the input and output play/spy scheme accordingly, it is possible to either test the interfaces with up/downstream modules by capturing input data and streaming output data from the memories. Alternatively, playback data injected into the input stage of L1Topo will allow to exercise the algorithms, with algorithm results captured in the output spy memories for subsequent readout and analysis.

In addition to the above facility, numerous flags describing the status of L1Topo can be read via the IPbus control. Access points are also provided for signal monitoring, boundary scanning and the use of proprietary FPGA tools such as ChipScope and IBERT.

2.8 Environmental Monitoring

L1Topo monitors the voltage and current of all critical power rails on the board. It also monitors the temperatures of all the FPGAs, of the MiniPOD receivers and transmitters, and of other areas of dense logic. Where possible, this is done using sensors embedded in the relevant devices themselves. Where this is not possible, discrete sensors are used.

A small set of voltage and temperature data are collected by the L1Topo IPMC, via an I^2C bus and are made available to ATLAS DCS via the shelf manager. Supplementary environment data are available to the control FPGA. These data can be accessed via IPbus.

FPGAs are protected against over temperature by internal monitoring and shutdown. This provides the lowest possible reaction time. Also, if any board temperature exceeds a programmable threshold set for a specific device monitored via IPMB, the IPMC powers down the board payload (that is, everything not on the management power supply). The thresholds at which this function is activated should be set above the levels at which the DCS will power down the module. Thus, this staged mechanism should activate only if the DCS fails. This might happen, for example, if there is a sudden, rapid rise in temperature to which the DCS cannot respond in time.

2.9 ATCA form factor

L1Topo is an ATCA module, conforming to the PICMG® 3.0 Revision 3.0 specifications. Backplane Zone-3 details are not part of the PICMG specification. Here the module design follows the L1Calo standards [1.15] .

The modules are dual width, they occupy two adjacent slots of an ATCA shelf each.

3 Implementation details

3.1 Modular Design

L1Topo consists of an ATCA sized main board, equipped with mezzanines. The mainboard mainly carries the real-time processing circuitry: Two processor FPGAs, connected up with 12 MiniPOD devices (10 × RX, 2 × TX) each. The FPGA/MiniPOD circuitry is two exact copies placed on the same PCB.

Module control via IPbus, and breakout of the electrical links to the CTP are implemented on the “Extension Mezzanine”. FPGA configuration memories are also located on the mezzanine, along with the TTC/clock reception and conditioning (jitter reduction). The mezzanine runs along the lower part of the front panel to allow for front panel connectivity and controls.

Further front panel connectivity and indicators are located on a separate, small front panel mezzanine in the upper part of the module.

Environmental monitoring and low level control is implemented on an IPMC controller module (LAPP IPMC).

Primary power supply is via standard PIM / converter brick(s). Viable solutions can be copied from either Phase-0 L1Topo or jFEX. Secondary power supplies are located on mezzanines.

3.2 Input Data Reception

L1Topo receives data from the L1Calo processors and the Muons via optical fibres. The bitrate is specified to 6.4, 11.2 and 12.8Gb/s, so as to be compliant with all data sources. The data are required to be 8b/10b coded data streams. Each fibre carries a net data volume of 224 (or 256 respectively) bits of data per bunch tick.

The input fibres to L1Topo are organised into 4 ribbons of 72 fibres each. They are routed to L1Topo via the rear of the ATCA shelf, where a rear transition module (RTM, [1.16]) provides mechanical support. Optical connections between the fibres and L1Topo are made by four 72-way Multi-fibre Push-On/Pull-Off (MPO) connectors, mounted in Zone-3 of the ATCA backplane. These connectors allow L1Topo to be inserted into, and extracted from, the shelf without the need to handle individual ribbon connections.

On the L1Topo side of the MPO connectors, 20 optical ribbons (each comprising 12 fibres) carry the signals to 20 MiniPOD receivers. These perform optical to electrical conversion. They are mounted on board, around the Processor FPGAs, to minimise the length of the multi-Gb/s PCB tracks required to transmit their output.

3.3 Processor FPGA

There are two Processor FPGAs on each L1Topo module. The functionality they implement can be grouped into real-time, readout and slow-control functions. Both FPGAs on an L1Topo module have the same wiring. Differences in functionality between Processor FPGAs on the same and different modules are due to different algorithms being run and are implemented via different firmware versions only.

Every Processor FPGA performs the following real-time functions:

- It receives, from MiniPOD optical receivers, up to 118 inputs of serial data at 6.4, 11.2 or 12.8 Gb/s per MGT link.
- It detects any data integrity issues with help of the MGT built-in error checks and with help of CRC checksums embedded in the user data.
  - Any errors are registered and counted,
  - Error counts can be read and reset via module control.
  - Any erroneous real-time data are zeroed.
- It allows for fine grain data alignment to word (bunch tick) boundaries.
- It allows for coarse grain data alignment in terms of full bunch ticks, up to 32 ticks.
- It runs topological algorithms on the conditioned real-time input data.
- It is able to share real-time data with the other on-board FPGA via parallel links
- It forwards the trigger results (typically a trigger bit with accompanying overflow bit) to the CTP.
- The CTP is fed with trigger results bits directly from each FPGA
  - Electrically (LVDS) via the extension mezzanine
  - Optically via MiniPOD

On the readout path, each Processor FPGA performs the following functions.

- The Processor FPGA records the input data and the output generated on the real-time path in scrolling memories, for a programmable duration of up to 10μs.
- On receipt of an L1A, it writes data from the scrolling memories to the FIFOs, for a programmable time frame. This is only done for those data enabled for readout by the control parameters.
- The Processor FPGA transmits data from the readout FIFOs to the ROD module (DROD), via a 6.4 Gb/s MGT backplane link.
- There is a further, equivalent data path from the Processor FPGAs to the RROD for purpose of RoI readout into the 2nd level trigger.

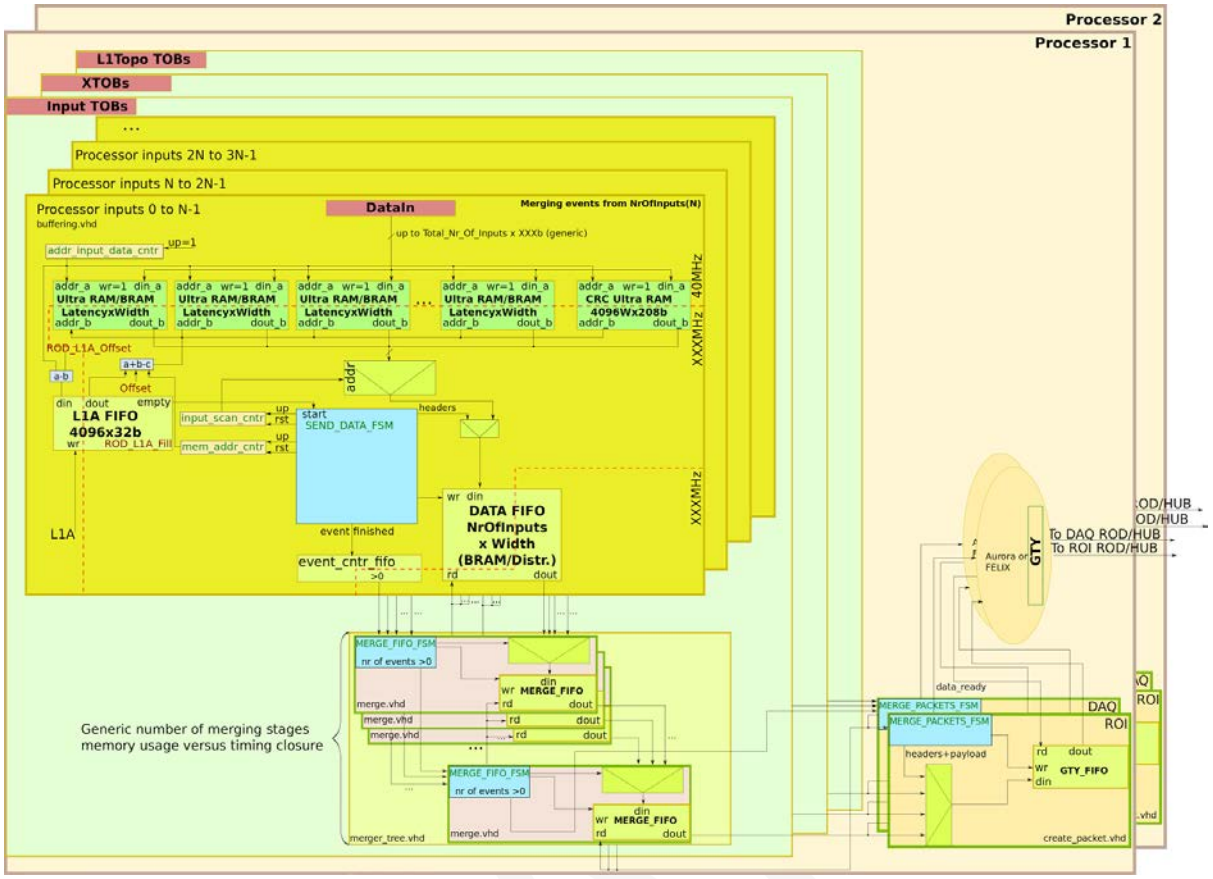

Figure 3. The block diagram of the L1Topo-I readout data flow.

Readout (Figure 3) will receive data from three different sources: input TOBs from FEXs and from muon subsystem, results from algorithms intermediate stages (XTOBs) and L1Topo TOBs containing RoI information. The data for each individual path (in TOBs/XTOB/TOB, in 13) will be first stored in the ring buffers (Ultra RAM/BRAM). In the next stage, after receiving L1A, a preselected number of these memories will be independently read out and merged (Processor inputs 0 to N-1, Processor inputs N to 2N-1 ...). In next steps the data will be again merged from a given number of data streams to a smaller number, when finally it will be stored in one FIFO. The number of merge stages, frequencies and memory types will be selected with the help of VHDL generics. This will allow to adjust the readout parameters to meet timing constraints where at the same time assuring required latency and reasonable memory usage. In each data path one will be able to: disable channels, send pre-scaled data, send data only on the CRC or input alignment error and select number of slices required to be sent.

For module control and monitoring, each Processor FPGA contains a local IPbus interface, which provides access to registers and RAM space within the FPGAs.

The Processor FPGA footprint on L1Topo is compatible to several FPGA types from the Xilinx UltraScale and UltraScale+ families. They are all 2577 ball devices. XCVU9P-2FLGA2577E is envisaged for L1Topo.

Of the 120 high speed links available in the XCVU9P, two are reserved for control purposes (TTC data and module control).

Regarding general-purpose I/O, of the total of 448 pins available, five banks of 24 pairs each are used for inter-FPGA data sharing. The pair count includes one pair of forwarded clock per bank. Each one-to-one bank interconnect is meant to be operated in one direction only. Receive and transmit lanes are not to be mixed within one bank. Inter-bank pin swapping is not allowed during PCB routing work.

3.3.1 Resource Estimate

Based on the configuration of the current topological processor, the following resource estimation is possible. The numbers are based on a synthesis of the firmware of the current module Topo00_U2 adapted to the envisaged XCVU9P FPGA. To be more confident about the estimate, the same estimation has been done for the current processor FPGA of the Virtex-7 family.
A comparison of the resulting resource estimation yields similar resource usage for both old and new system in terms of absolute numbers. Since the XCVU9P provides approximately three times as many resources as the current Virtex-7, the relative numbers drop by a factor of three.
The relative numbers of the estimated usage for the most important resources are shown in 0.

Resource	Estimated Usage
LUTs	19 %
Flip Flops	4 %
DSP Slices	7 %
BRAMs	3 %

Table 2 Resource estimate

3.4 Clocking

There are two types of clock sources on L1Topo: on-board crystal clocks and the LHC TTC clock, received from the ATCA backplane. These clock sources are fed via the clocking circuitry to the two processor FPGAs. The 40.079MHz TTC “clean” clock has potentially too much jitter to drive multi-Gb/s links directly. A PLL chip is therefore used to clean up the jitter on this clock. From the input of 40.079 MHz, the PLL chip can generate clocks of frequency $n \times 40.079$ MHz within a certain range. This flexibility allows the multi-Gb/s links on the L1Topo to be driven at a range of different rates. The Si5345 has been tested and verified on the jFEX prototype and will be used on L1Topo. The clock (re)generation circuitry is located on the extension mezzanine, the individual clock trees for MGT reference clocks and global clocks are actively fanned out on the main board.

The four MGT reference clock trees are operated at CML signal level, they are AC coupled into the FPGAs. The main clock tree supplies the real-time inputs running at 6.4 / 11.2 / 12.8Gbps. There are additional trees for real-time output (6.4/12.8Gb/s) and for backplane output towards the RODs (6.4Gb/s). Note: According to UltraScale+ documentation it should be possible to derive all required MGT internal PLL frequencies from a common 160.32 MHz reference clock. Therefore, the additional clocks will probably not all be in use on the production modules. A separate crystal clock will be available to the MGT quads carrying IPbus links.

3.5 High-Speed signals on the PCB

L1Topo is a very high-speed and very high-density ATCA module, which has many optical fibre links and some electrical backplane links running at a speed of up to 12.8Gb/s. In addition, the tight ATLAS L1Calo latency margin requires a large number of parallel links running at nominally 640Mb/s between FPGAs for data sharing on L1Topo.

Signal integrity is a challenge for the L1Topo design. It benefits, however from the detailed PCB simulations that have been done for the jFEX prototype, from which the phase-1 L1Topo is being derived. Cross talk is limited by maximising differential pair pitch, impedance is guaranteed (10%) by the PCB manufacturer.

3.6 FPGA configuration

The configuration of the two large processor FPGAs is controlled from the Extension Mezzanine. To this end all signal lines required for either master SPI mode or slave SelectMAP are routed to the mezzanine.

The baseline configuration option is SPI mode. Though dual SPI mode (ie. Byte wide configuration) is supported by this scheme, the mezzanine currently under construction will use a single SPI flash memory chip per processor FPGA. The flash devices can be written either via JTAG or via IPbus. The latter operation will require specific firmware and software to be written.

The configuration scheme will allow for both the current production firmware and a “golden” recovery image to be stored on the SPI flash devices. Whether that feature will actually be used is as yet undecided, since the processor FPGAs can always be configured through the mezzanine-based control FPGA, even with an erased or corrupted flash chip connected up to the processors. Direct JTAG configuration of the processor FPGAs is an additional option for debug purposes.

3.7 The Extension Mezzanine

The Extension Mezzanine module provides many of the (non-realtime) services described above. It carries mainly module control, clock/control, and configuration circuitry. It also provides initialization circuitry for the FPGAs and acts as an interface to environmental monitoring devices. The only real-time signals running via the mezzanine are the electrical outputs to the CTP.

The “intelligent” module controller is an FPGA from the XILINX Artix-7 family. This Control FPGA handles incoming IPbus requests and forwards the data and control packets to the processors on the mainboard via MGT (GTP) links. MGT links are also used to replicate incoming TTC data into the two processors (see below).

The IPbus communicates with its control PC(s) via an Ethernet Phy chip. The chip type chosen is VSC8221. It is an electrical Ethernet (1000BASE-T) to SGMII device. The SGMII link is connected to an MGT link of the control FPGA. The 1000BASE-T port is linked to the

Hub/ROD module-1 via the backplane. This link is AC-coupled with series capacitors. Magnetics (transformers) are not required due to the choice of Phy chip, which is specifically designed (voltage mode drivers, internal biasing) to support magnetics-free links.

The backplane clock arriving from the Hub modules is transmitted at the LHC bunch crossing frequency of 40.079 MHz and meant to be of high quality, low jitter. However, locally on the mezzanine this clock is run through a jitter cleaner / clock synthesizer chip (Si5345) where it is refreshed and multiplied to higher ratios of the bunch clock. The jitter cleaner delivers four multiples of the base frequency: x1 multiplication, just jitter cleaned for purpose of global clock into the FPGA fabric, a multiple suitable for 6.4/11.2/12.8 Gb/s real-time input reference, a multiple for the backplane readout links and a separate multiple for the real-time outputs.

The mainboard processors are fed from the jitter cleaner outputs via clock fan-out chips. The global (FPGA fabric) clocks are of LVDS level, the MGT reference clocks of CML. Separate crystal clocks are provided for local use on IPbus/Ethernet and optionally for TTC data inputs.

The TTC data links are received from the backplane, one AC-coupled MGT link from each Hub/ROD module. The data are routed into the control FPGA, where they are interpreted and forwarded to the processor FPGAs, again on AC-coupled MGT links. The TTC data links are synchronous to the LHC bunch clock and therefore require an LHC clock multiple for re-transmission to the processors on the mainboard.

While the processor FPGAs are accessible through their JTAG ports at any time, the configuration bit stream required at any power-up is meant to be provided by local storage. Default storage device is one large (quad) SPI flash memory per FPGA. The device chosen for the first version of the mezzanine is MT25QU01 or MT25QU02. Different configuration schemes can be made available with further versions of the mezzanine card, should the updates of the flash devices, required for any persistent processor firmware updates, be considered inconveniently slow. The update process can be triggered and controlled from either the control FPGA or via JTAG. The control FPGA itself will in any case be configured from a small SPI flash chip, which due to smaller capacity and rare updates, is assumed to be a rather painless update operation. For the control FPGA, in-situ (live) updates are possible due to the use of a Xilinx-provided fall-back / golden image scheme.

Environmental data (voltages, currents, temperatures) are collected on the mainboard by I2C based sensors, and routed to the mezzanine via the bidirectional I2C buses. Parameters in the respective devices are set in the same way. Data are originating from dedicated monitoring chips, or from monitor/control interfaces available in core functionality devices, e.g. MiniPODs. They are routed into the control FPGA with an optional breakout onto headers. The control FPGA allows for access to these data via IPbus. The status/control data exchanged that way are complementary to the IPMC data. The handling of serialized slow control data on FPGAs and the description of the required state machines in VHDL and the maintenance of such circuitry is not particularly efficient in terms of engineering effort. For this reason, an updated mezzanine with a complementary, small microcontroller for housekeeping functionality is envisaged. Alternatively, an embedded processor might be used on the FPGA.

The real-time signals forwarded to the CTP via the mezzanine are plain route-through only. They are run via the mezzanine so as to allow for re-grouping signals from the two processor

FPGAs into a single cable port, should that be required. This scheme is taken over from the Phase-0 Topology processor. At current the signal distribution is symmetric, same bandwidth from each of the processors. That's the baseline for the Phase-1 modules as well, unless specific requirements are presented. Pinout of the VHDCI connector will be unchanged wrt. Phase-0 L1Topo[1.12] .

3.8 The IPM Controller

For the purposes of monitoring and controlling the power, cooling and interconnections of a module, the ATCA specification defines a low-level hardware management service based on the Intelligent Platform Management Interface standard (IPMI). The Intelligent Platform Management (IPM) Controller is that portion of a module (in this case, L1Topo) that provides the local interface to the shelf manager via the IPMI bus. It is responsible for the following functions:

- interfacing to the shelf manager via dual, redundant Intelligent Platform Management Buses (IPMBs), it receives messages on all enabled IPMBs;
- negotiating the L1Topo power budget with the shelf manager and powering the payload hardware only once this is completed (see section 3.9);
- managing the operational state of L1Topo, handling activations and deactivations, hot-swap events and failure modes;
- implementing electronic keying, enabling only those backplane interconnects that are compatible with other modules in shelf, as directed by the shelf manager;
- providing to the Shelf Manager hardware information, such as the module serial number and the capabilities of each port on backplane;
- collecting, via an I^2C bus, data on voltages and temperatures from sensors on L1Topo, and optionally exchanging these data with the control FPGA;
- driving the ATCA-defined LEDs.

L1Topo uses the IPMC module produced by LAPP as the IPM Controller [1.17] .The form factor is DDR3 VLP Mini-DIMM.

3.9 Power Management

With regard to power, the hardware on the L1Topo is split into two domains: management hardware and payload hardware. The management hardware comprises the IPM Controller plus the primary DC-DC converters and any non-volatile storage that this requires. By default, on power up, only the management hardware of L1Topo is powered (drawing no more than 10 W), until the IPM Controller has negotiated power-up rights for the payload hardware with the shelf manager. This is in accordance with the ATCA specification. However, via a hardware switch it is also possible to place L1Topo in a mode where the Payload logic is powered without waiting for any negotiation with the shelf controller. This feature, which is in violation of the ATCA specification, is provided for diagnostic and commissioning purposes.

On power-up of the payload hardware, the sequence and timing with which the multiple power rails are turned on can be controlled by a programmable device.

Excluding the optional exception noted above, the L1Topo conforms to the full ATCA PICMG® specification (issue 3.0, revision 3.0), with regard to power and power management. This includes implementing hot swap functionality, although this is not expected to be used in the trigger system.

Power is supplied to L1Topo on dual, redundant -48V DC feeds. A standard power input module (eg. PIM400) and a step down convertor, both “quarter brick” sized, are employed for power conditioning and conversion down to 12V. Alternatively a combined PIM/converter device is considered. The 12V supply is stepped down further, by multiple (secondary) switch-mode regulators, to supply the multiplicity of voltages required by the payload hardware.

For the power supplies to the FPGA multi-Gigabit transceivers, the PCB design guidelines and noise requirements specified in the UltraScale+ Series FPGAs GTY Transceiver User Guide (UG578) will be observed. The secondary convertors are located on small mezzanine modules.

An initial power estimate has been derived from the jFEX figures, taking into account Xilinx XPE spreadsheets and (extrapolated) measurements. The FPGA internal voltages (0.95V) are expected not to exceed a maximum current of 45A per FPGA; MGT Voltages (1V and 1.2V) are expected to be loaded with less than 16 and 19 A respectively. That is a maximum of ~80W Watt per FPGA, plus board level supplies (1.8V, 2.5V, 3.3V) of about 50W total.

3.10 Front-panel Inputs and Outputs

The following signals are, or can be, sent or received via the L1Topo front panel.

- Electrical differential (LVDS) signals are sent to the CTP via an SCSI VHDCI style connector, located on the mezzanine. Wiring details copied from Phase-0 L1Topo.
- Fibre-optical output to CTP via MPO/MTP connectors. A total of 48 fibres can be sent out of the front panel, largest fraction assumed to be spares for possible use at Phase-2.
- Auxiliary clock in. MMCX connector. This input allows L1Topo to be driven by an external 40.079 MHz clock, in the absence of a suitable clock on the backplane.
- Clock out. MMCX connector

The following bi-directional control interfaces are available on the front panel. See section 3.13 for the use of these interfaces.

- JTAG Boundary Scan.
- 1G Ethernet socket (optional, not to be used in production environment).

3.11 Rear-panel Inputs and Outputs

3.11.1 ATCA Zone-1

This interface is configured according to the ATCA standard. The connections include

- dual, redundant -48V power supplies,
- hardware address (used to derive MAC/IP addresses for IPbus)
- IPMB ports A and B (to the Hub modules),
- shelf ground,
- logic ground.

Figure 4 shows the backplane connections between the L1Topo and the Hub module, which are located in zones 1 and 2 of the ATCA backplane. See the ATCA specification for further details.

3.11.2 ATCA Zone-2

3.11.2.1 Base Interface

The Base Interface comprises eight differential pairs. Four of these are connected to Hub slot one and are used for module control (IPbus), the other four are connected to Hub slot two and are used to interface to the IPMC.

3.11.2.2 Fabric Interface

The Fabric Interface comprises 16 differential signal pairs, eight of which are connected to Hub slot one, and eight of which are connected to Hub slot two. Those signal pairs connected to Hub slot one are used as follows:

- One signal pair is used to receive the TTC “clean” clock of 40.079 MHz.
- One signal pair is used to receive decoded TTC commands, plus near real-time signals such as ROD busy. This lane is connected into a multi-Gigabit receiver on the extension mezzanine. The exact protocol is defined by the Hub module developers and is implemented in firmware. The link speed does not exceed 10 Gb/s.
- Six signal pairs are used to transmit readout data via MGT links. The protocol is being jointly defined by L1Topo and ROD module developers. The link speed is 6.4 Gb/s. Two out of these six signal pairs are used as receivers in standard ATCA backplanes. They are operated in inverse direction on all L1Calo modules to increase the possible readout bandwidth. These two links are considered spares on L1Topo

The same connectivity is available into Hub slot 2. For details on backplane use see [1.14] .

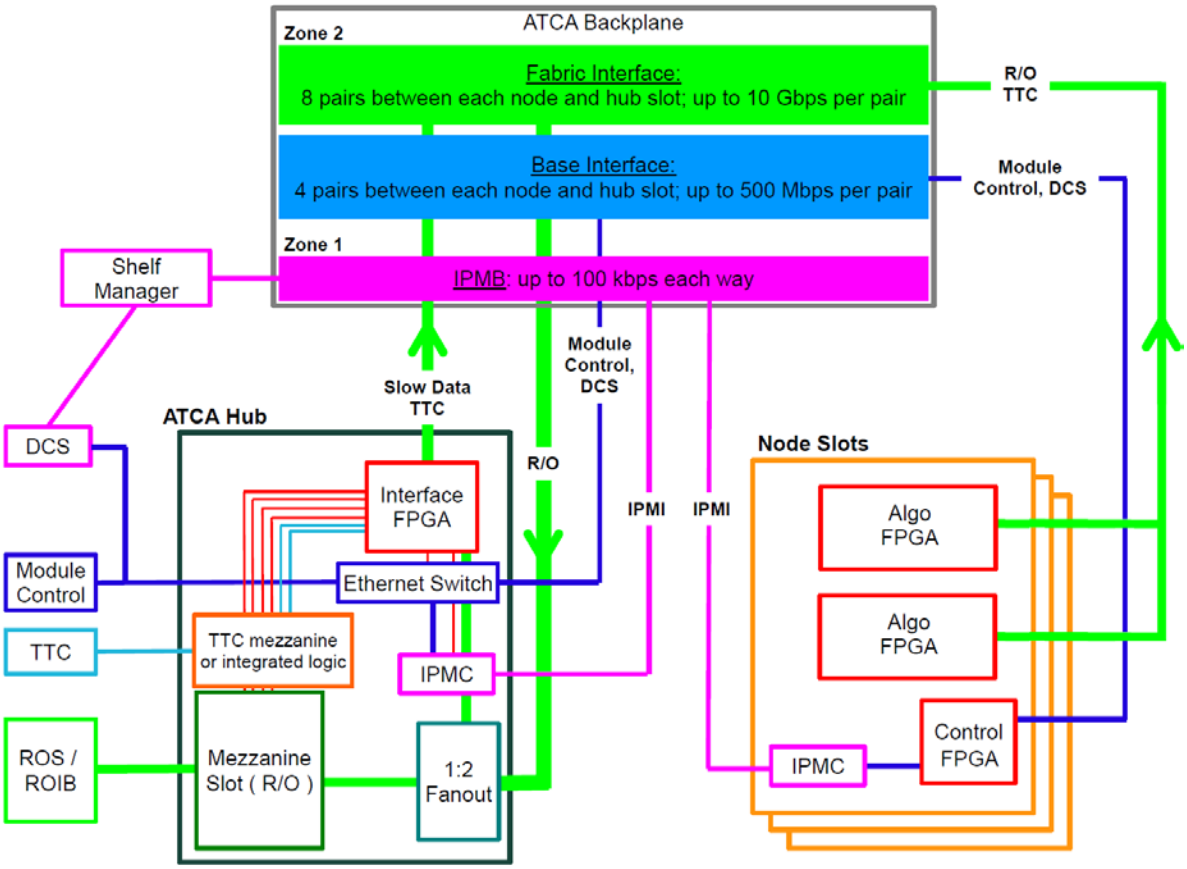

Figure 4. The ATCA backplane connections between the L1Topo and the Hub module.

3.11.3 ATCA Zone-3

ATCA Zone-3 houses four optical MPO connectors. That allows for up to 288 fibres, carrying data from the feature extractors and muons to L1Topo (see section 3.1). These fibres are supported in the L1Topo shelf by a (passive, mechanical) rear transition module (RTM,[1.16]). On the L1Topo side of the connectors, fibre ribbons carry the calorimeter data to MiniPOD receivers, mounted in board. The optical connections are made on the insertion of the L1Topo into the shelf, and broken on its extraction. Dependent on the requirements, real-time output can possibly be run on otherwise dark fibres (spares). However, it is anticipated that real-time optical output connection is rather made via the front panel.

3.12 LEDs

All LEDs defined in the ATCA specifications are located on the L1Topo front panel. In addition, further status LEDs are provided on either the front panel or the top side. These indicate functions like power, DONE signals, L1A receipt und further LEDs for diagnostic purposes for all FPGAs.

3.13 Instrument Access Points

3.13.1 Set-Up and Control Points

The following interfaces are provided for the set-up, control and monitoring of L1Topo. They are intended for commissioning and diagnostic use only. During normal operation it should not be necessary to access the L1Topo via these interfaces.

- The JTAG Boundary Scan port: via this port a JTAG/boundary scan test can be conducted to check board level connectivity, all FPGAs on the L1Topo can be configured, the configuration memory of the Configurator can be loaded and the FPGA diagnostic/evaluation tool ChipScope can be run, including for IBERT tests. Form factor is Xilinx standard 14-pin, 2mm pitch. This port is on the front panel.
- The 1G Ethernet port (optional): this port provides an auxiliary control interface to L1Topo, over which IPbus can be run, should there be a problem with, or in the absence of, an IPbus connection over the shelf backplane. It is on the front panel and located on the extension mezzanine. This is an optional front panel port.

3.13.2 Signal Test Points

Due to the sensitive nature of multi-Gb/s signals, no test points are provided on PCB tracks intended to carry multi-Gb/s data. If such signals need to be examined, this must be done via firmware. Test points are placed on a selection of those data and control tracks that are not operating at multi-Gb/s.

For each FPGA, a few spare, general-purpose IO pins are routed to 2.54mm headers. Furthermore, spare multi-Gb/s links are routed to MMCX sockets. With appropriate firmware these connections allow internal signals, or copies of data received, to be fed to an oscilloscope, for example, or driven from external hardware.

The exact number of test connections, and those signals on which a test point can be placed most usefully, are to be determined in the final stage of module layout.

3.13.3 Ground Points

At least six ground points are provided, in exposed areas on the top side of the module, to allow oscilloscope probes to be grounded.

3.14 Floor plan

Figure 5 shows a preliminary floor plan of the L1Topo module. This will be used as a guide for the layout process; the exact location of components may change.

The routing of c. 300 signals at multi-Gb/s presents a significant challenge for the design of the L1Topo PCB. In order to minimise track lengths and routing complexity for these signals, the Avago MiniPOD receivers are placed around the Processor FPGAs. However, this creates an additional constraint on the layout: the need to accommodate routing paths for the fibre-

optic ribbons carrying the data to these receivers. To connect the MPO connectors to the receivers the ribbons need to twist, curve and bypass large heat sinks on the FPGAs. The Extension Mezzanine is shown bottom left, the IPMC along the Zone-3 connectors.

In addition to those components shown in Figure 5, glue logic is placed on the underside of the module.

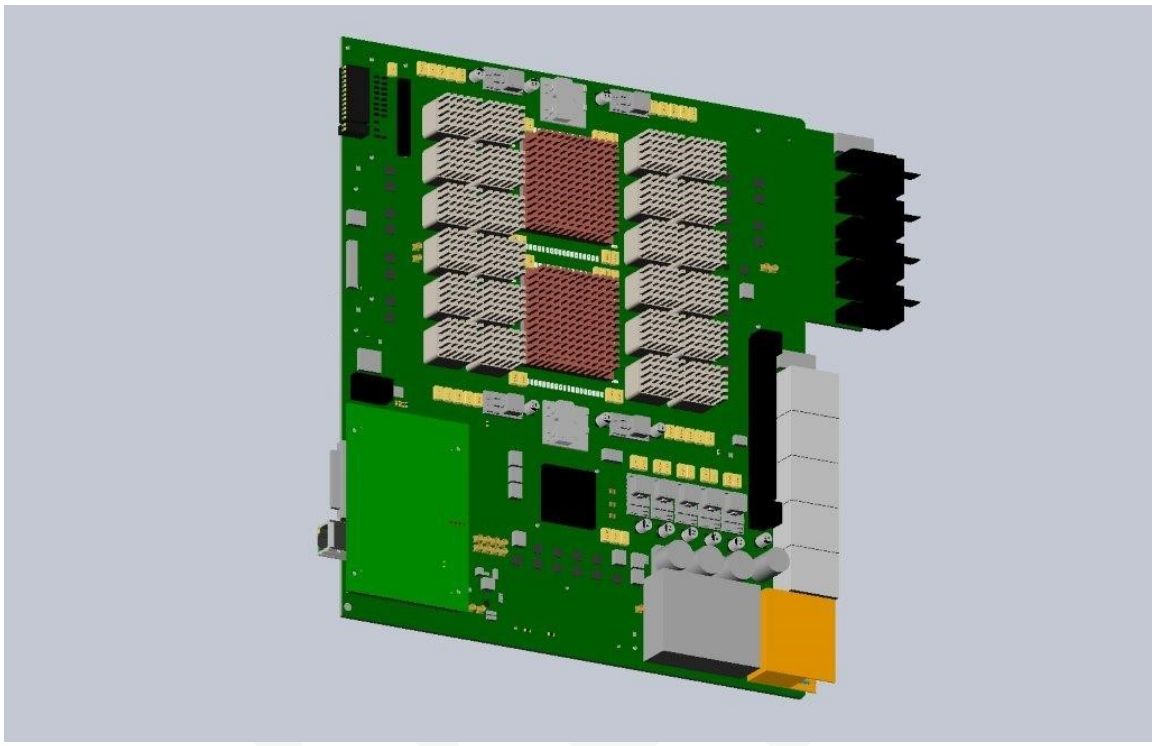

Figure 5. Sketch of L1Topo, showing a preliminary placement guide.

4 Front-Panel Layout

Figure 6. Preliminary front panel layout.

Figure 6: A drawing of the front panel will be added once the component placement is final and the panel design has been made. For the time being only the front panel elements are listed: ATCA specific LEDs, LEDs for FPGA DONE, IPbus activity, L1A, power status. Optional RJ45s for IPbus and IPMC, clock monitor, external clock, CTP electrical out, CTP optical out (MPO/MTP), JTAG

5 Summary: Interfaces

Some important details of interfaces to external systems as described above are summarized in this section.

5.1 Internal Interfaces

The Extension Mezzanine is connected to the L1Topo main board via a large 1mm pitch connector. The pinout will be documented as soon as the design is finalized.

5.2 External Interfaces

5.2.1 Electrical TTC interface (backplane input)

A clean clock of 40.079MHz is received as a differential electrical signal via the ATCA backplane. The signal is AC-coupled on the extension mezzanine and routed into an “any-in” differential receiver or directly into the jitter cleaner

The clock is accompanied by a TTC data signal, differential, AC coupled on the mezzanine, electrically compatible to Xilinx MGT. The data rate is assumed to be 3.2Gb/s, 8b/10b encoded. Format being defined by the Hub designers.

Data paths supported from both Hub slots 1 and 2.

5.2.2 Electrical DAQ interface (backplane output)

Readout data are sent to the DAQ/2nd level trigger via the ATCA backplane on 6 links to each Hub/ROD, LHC bunch clock synchronous, AC coupled on L1Topo, Xilinx MGT compatible, at 6.4Gb/s. Data paths are supported into both Hub slots 1 and 2. The data formats are being defined together with the Hub/ROD community. Readout/RoI paths are supported from both FPGAs to both Hub slots, i.e. a total of 4 times 2+1spare link.

5.2.3 IPbus interface (backplane I/O)

Module control links are standard Gigabit Ethernet via the backplane from/to Hub slot 1. The phy chip is located on the extension mezzanine. The envisaged phy chip (VSC8221) allows for magnetics-free, capacitive coupling, which will be the baseline.

5.2.4 DCS interfaces (backplane I/O)

The IPMC module is linked to the outside world via an I2C (IPMB) bus in ATCA Zone-1, and a standard Ethernet link to Hub slot 2 via the base interface.

5.2.5 Electrical CTP interface (front panel output)

The Central Trigger Processor is interfaced electrically via a VHDCI SCSI style connector. Pinout is unchanged with respect to the Phase-0 L1Topo module. Signals level is LVDS. All signal pairs can be driven from the two processor FPGAs. The allocation of pairs to individual FPGAs is implemented on the extension mezzanine. The interface is assumed to be data lines only, though parity and clock signals could be generated in FPGAs if required. The signal level is LVDS.

5.2.6 Optical CTP interface (front panel output)

The Central Trigger Processor is interfaced fibre-optically via an MTP/MPO connector on the front panel. Up to 48 total fibres can be driven from the two processor FPGAs through MiniPODs. The maximum bitrate is 14Gb/s, the CTP interface is assumed to run at 6.4 Gb/s, spare links 12.8Gb/s, synchronous to the LHC clock. Data encoding is 8b/10b.

5.2.7 Optical FEX/Muon interface (rear input)

The calorimeter FEXes (e/j/g-FEX) and the muon trigger are fibre-optically interfaced via the backplane, on 72-way MTP/MPO connectors. The mechanical interface to the RTM is Molex MTP-CPI. Four of these shrouds are available in ATCA Zone-3. The signals are routed through MiniPODs (up to 14 Gb/s) and received into FPGAs via MGT links. Encoding is 8b/10b. Data rate is specified for mixed operation 6.4/11.2/12.8Gb/s. Signal rates are not to be mixed in same quad.

6 Appendix : Data formats

The formats of the data received and generated by L1Topo are about to be finalised. Details are found in separate documents. Tables can be added once formats are final. This section gives a coarse overview only.

6.1 Real-Time Input Data

Real-time input from FEXes and Muon Trigger is 8b/10b-encoded at 6.4, 11.2 or 12.8 Gb/s. This yields a line capacity of 224 or 256 bits total per bunch crossing. The raw data are accompanied by a CRC check sum and by comma characters, required for line synchronization. Comma characters are sent upon link start-up and either in regular intervals, or in place of otherwise empty data fields, replacing 0x00 data bytes. Comma characters are in either case injected in fixed and unique positions within a full-BC data word only. For purpose of overall alignment and monitoring, bunch count information will be embedded into the data stream. Detailed formats are currently being finalized together with the L1Calo community.

6.2 Real-Time Output Data

The Real-time output of L1Topo into the CTP is composed of trigger information, accompanied by overflow information. On the electrical interface this information is sent without any further formatting, as an 80Mb/s stream. On the optical interface the raw data will be protected by a CRC check sum and aligned with help of embedded comma characters, plus overall alignment with embedded bunch count information.

6.3 Backplane data formats

Readout streams into DAQ and RoI systems are routed through the two Hub/ROD modules in the shelf. The formats on the data links are being defined together with the ROD community. It should be noted that it will not be possible to run all DAQ or RoI output in a channel bonded scheme, since it is actually two separate streams from distinct sources, the two processor FPGAs. Detailed data contents might partially be modelled on the existent Phase-0 L1Topo protocol, though eFEX protocol details (headers/trailers) will be employed where possible, to simplify ROD firmware production and maintenance.

The TTC data running on the backplane from the Hub modules to the L1Topo modules are re-coded on the Hub. The exact protocol is being defined by the Hub designer community.

7 Related Documents

[1.1] ATLAS TDAQ System Phase-I Upgrade Technical Design Report, CERN-LHCC-2013-018, http://cds.cern.ch/record/1602235/files/ATLAS-TDR-023.pdf

[1.2] Phase-I Latency Envelopes for the Level-1 Trigger, https://edms.cern.ch/document/1256858/

[1.3] L1Calo Phase-I eFEX Specification, https://edms.cern.ch/document/1419789

[1.4] L1Calo Phase-I jFEX Specification https://edms.cern.ch/document/1419792

[1.5] L1Calo Phase-I gFEX Specification https://edms.cern.ch/document/1425502

[1.6] ATCA Short Form Specification, http://www.powerbridge.de/download/know_how/ATCA_Short_spec.pdf

[1.7] PICMG 3.0 Revision 3.0 AdvancedTCA Base Specification, http://www.picmg.com/

[1.8] L1Topo requirements, https://indico.cern.ch/event/638444/contributions/2639285/attachments/1490897/

[1.9] L1Calo Phase-I Hub Specification, https://edms.cern.ch/document/1415974

[1.10] L1Calo Phase-I ROD specification, https://edms.cern.ch/document/1404559

[1.11] Foxconn 14Gb/s MiniPOD devices, http://www.fit-foxconn.com/Product/ProductDetail?topClassID=Electronic Module&&PN=AFBR-822VxyZ

[1.12] Phase-0 L1Topo module, http://esimioni.web.cern.ch/esimioni/TPF/TP_mainh.html

[1.13] Phase-0 L1Topo algorithms/firmware, https://gitlab.cern.ch/sartz/L1TopoFirmwareDocumentation/blob/master/L1TopoFirmware.pdf

[1.14] L1Calo usage of ATCA backplane, see https://edms.cern.ch/file/1492098

[1.15] L1Calo 8U front board form factor, see https://edms.cern.ch/file/1492098

[1.16] L1Calo 8U RTM form factor, see https://edms.cern.ch/file/1492098

[1.17] LAPP IPMC module, see http://lappwiki.in2p3.fr/twiki/bin/view/AtlasLapp/ATCA

8 Glossary

ATCA	Advanced Telecommunications Computing Architecture (industry standard).
BC	Bunch Crossing: the period of bunch crossings in the LHC and of the clock provided to ATLAS by the TTC, 24.95 ns.
DAQ	Data Acquisition.
DCS	Detector Control System: the ATLAS system that monitors and controls physical parameters of the sub-systems of the experiment, such as gas pressure, flow-rate, high voltage settings, low-voltage power supplies, temperatures, leakage currents, etc.
DROD	Data ROD (Data into DAQ) in Hub slot 1
ECAL	The electromagnetic calorimeters of ATLAS, considered as a single system.
eFEX	Electromagnetic Feature Extractor.
FEX	Feature Extractor, referring to either an eFEX, gFEX or jFEX module or subsystem.
FIFO	A first-in, first-out memory buffer.
FPGA	Field-Programmable Gate Array.
gFEX	Global feature extractor
HCAL	The hadronic calorimeters of ATLAS, considered as a single system.
IPbus	An IP-based protocol implementing register-level access over Ethernet for module control and monitoring.

IPMB	Intelligent Platform Management Bus: a standard protocol used in ATCA shelves to implement the lowest-level hardware management bus.
IPM Controller	Intelligent Platform Management Controller: in ATCA systems, that portion of a module (or other intelligent component of the system) that interfaces to the IPMB.
IPMI	Intelligent Platform Management Interface: a specification and mechanism for providing inventory management, monitoring, logging, and control for elements of a computer system. A component of, but not exclusive to, the ATCA standard.
jFEX	Jet Feature Extractor.
JTAG	A technique, defined by IEEE 1149.1, for transferring data to/from a device using a serial line that connects all relevant registers sequentially. JTAG stands for Joint Technology Assessment Group.
L0A	In Run 4, the Level-0 trigger accept signal.
L0Calo	In Run 4, the ATLAS Level-0 Calorimeter Trigger.
L1A	The Level-1 trigger accept signal.
L1Calo	The ATLAS Level-1 Calorimeter Trigger.
LHC	Large Hadron Collider.
MGT	As defined by Xilinx, this acronym stands for Multi-Gigabit Transceiver. However, it should be noted that it denotes a multi-gigabit transmitter–receiver pair.
MiniPOD	An embedded, 12-channel optical transmitter or receiver.
MicroPOD	An embedded, 12-channel optical transmitter or receiver, smaller compared to the MiniPOD.
MMXC	Sub-Miniature coaxial RF connector.
MPO	Multi-fibre Push-On/Pull-Off: a connector for mating two optical fibres.
PMA	Physical Media Attachment: a sub-layer of the physical layer of a network protocol.
ROD	Readout Driver.
RROD	ROI ROD (Data into 2^{nd} level trigger), in Hub slot 2
RoI	Region of Interest: a geographical region of the experiment, limited in η and ϕ, identified by the Level-1 trigger (during Run 3) as containing candidates for Level-2 trigger objects requiring further information.
RTM	Rear Transition Module, module in the back compartment of a shelf, allowing for connections to the front module. Here: passive fibre coupling mechanics.
Shelf	A crate of ATCA modules.
TOB	Trigger Object.
TTC	The LHC Timing, Trigger and Control system.

XTOB	Extended Trigger Object. A data packet passed to the readout path, contained more information than can be accommodated on the real-time path.

9 Document History

Version	Comments
0.7	Internal circulation without mezzanine section
0.8	Added mezzanine section
0.9	Draft for initial distribution
1.0	Draft circulated for review