High-Density Optical Receiver Mezzanine board for ATCA-SRS

Sorin Martoiu, Gabriel Stoicea, Michele Renda (IFIN-HH, Bucharest)

Table of Contents

Overview of the mezzanine board 1
Mezzanine Board Specifications 6
1. Processing Units 6
2. Communication Links 7
3. Clock Distribution 9
4. Power Distribution 12
5. Board Management 14
Schedule and Responsibilities 14
Appendix A. Application Example of the HORX Mezzanine card as NSW Trigger Processor 15
Appendix B. Compatibility with Phase-II 17

Overview of the mezzanine board

The High-Density Optical Receiver (HORX) mezzanine board provides up to 72 bidirectional optical links using 12 high-speed Avago MicroPod modules specified up to 12.5 Gbps, and high-density programmable logic processing power provided by two Xilinx Virtex-7 FPGA units. The two FPGAs are interconnected by low-latency LVDS lines (64) and high-speed GTH transceiver lines (8), which may offer the possibility to join the processing capability of the two FPGA units into a single powerful processing unit.

The two FPGAs on the mezzanine board are independently connected to the ATCA-SRS mezzanine interface consisting of high-speed GTH transceiver lanes (8), low-latency LVDS signals (50) and single-ended LVCMOS signals (16). On the carrier board the interface signals are connected to an FPGA, giving full flexibility in terms of the protocols used. The mezzanine interface also provides clock signals which are sourced by the carrier board, management interfaces (JTAG and IPMI) which are driven by the IPMC

controller resident on the ATCA carrier and power lines. More details of the ATCA-SRS carrier board are given in the following chapter.

The mezzanine board implements a flexible clock distribution scheme which allows for separate clock domains for the FPGA fabric and two distinct domains for the GTH transceivers, for each of the two FPGAs. Each of the 6 clock domains are controlled by independent PLL circuits with multiple outputs. The reference clocks for these PLL circuits can be either local oscillators, clock recovered from any input optical stream, or reference clock signals supplied by the carrier board.

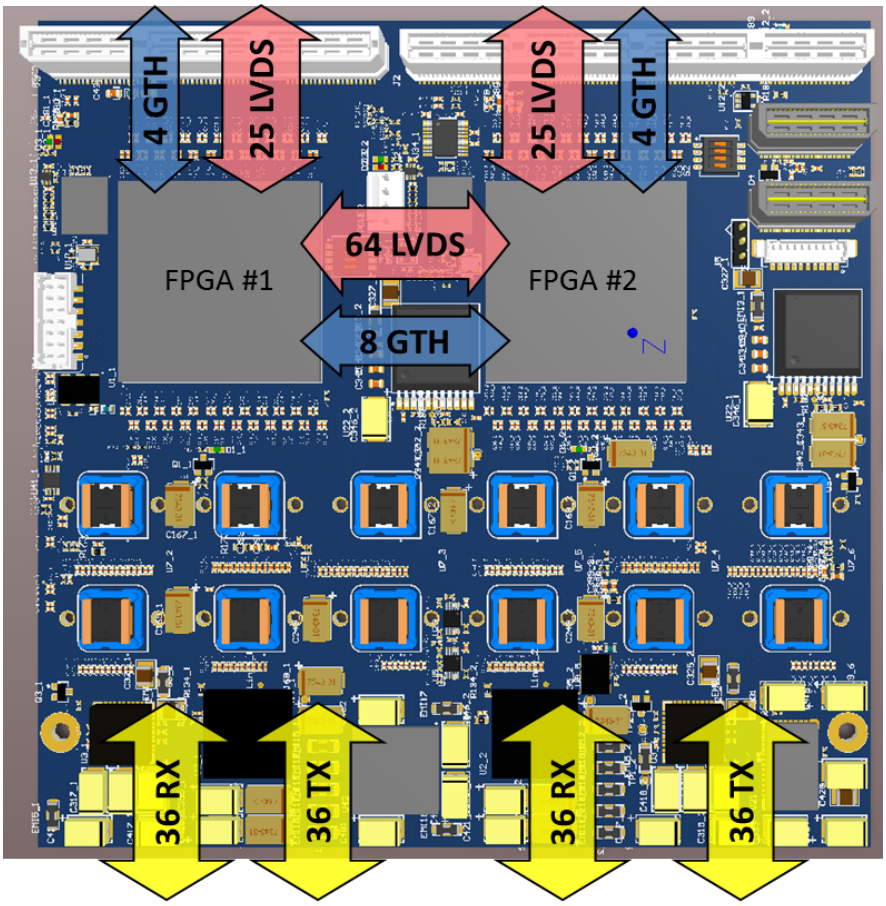

Figure 1. Ilustration of the High-Density Optical Receiver Mezzanine Board.

The system clock conditioners implement a zero-delay option designed to provide clock phase synchronization across the system. The clock source may come from the ATCA-SRS carrier board via the RTM port (discrete LVDS connections) or derived from an input serial stream supplied to an SFP port on RTM. In addition, the ATCA-SRS carrier board is compatible with the AXIe standard which allows skew controlled point-to-point propagation of reference clock over the AXIe-enabled backplane. In this case, clock synchronization can be achieved across the entire ATCA crate.

Integration of the mezzanine board with the ATCA-SRS Carrier board

The ATCA-SRS board was designed within the RD51 Collaboration as a versatile platform for data acquisition in medium or large scale experiments in high-energy physics or other scientific domains. The board is based on a modular concept which allows the integration of application specific front-end interfaces on mezzanine cards which connect to general purpose processing units based on FPGAs. The communication links inside the ATCA-SRS system are conceptually designed to provide scalability both in application space, allowing the user to integrate different detector types within the same acquisition system, and in space domain, giving the possibility to expand the system size from small to medium or large scale systems based on the same components.

Figure 3 shows how the mezzanine board is connected to the ATCA carrier board, as well as the connection on the carrier to the ATCA fabric and RTM module. Each mezzanine is connected to a dedicated FPGA on the carrier board.

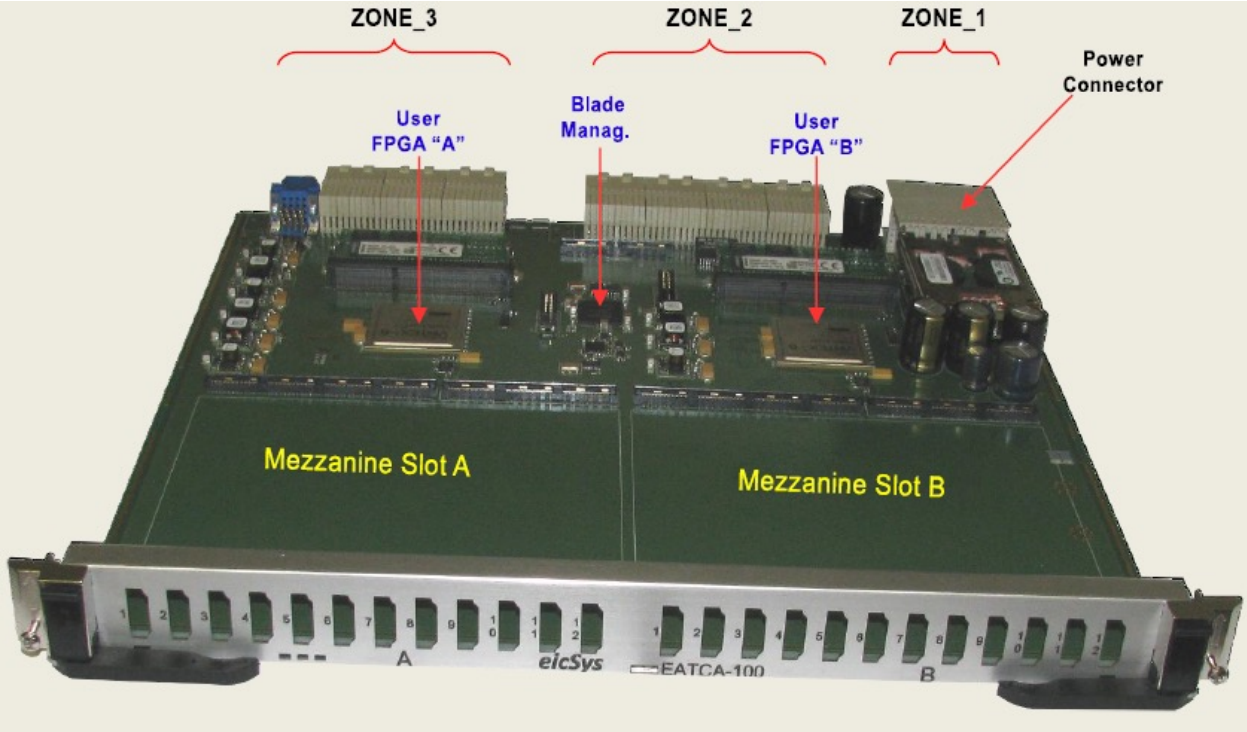

Figure 2. Photo of the current existing version of the ATCA-SRS carrier board

The two carrier FPGAs are linked to the ATCA backplane fabric in two modes. The hub channels of the fabric interface are used for high speed data exchange over the backplane using either proprietary protocols (i.e. like GBT), or an industry standard protocol like 10 GB Ethernet or PCIe. A high-speed multiplexer allows either a parallel connection of the two carrier FPGAs to the hub channel or a master-slave configuration where one only one FPGA is connected, while the other (slave) FPGA is connected to the master FPGA via an internal bus. The other fabric channels can be used for low-latency data exchange between any two node cards of the system using direct LVDS connections to the carrier FPGAs.

The two FPGAs are also linked together by a combination of serial links and LVDS signals.

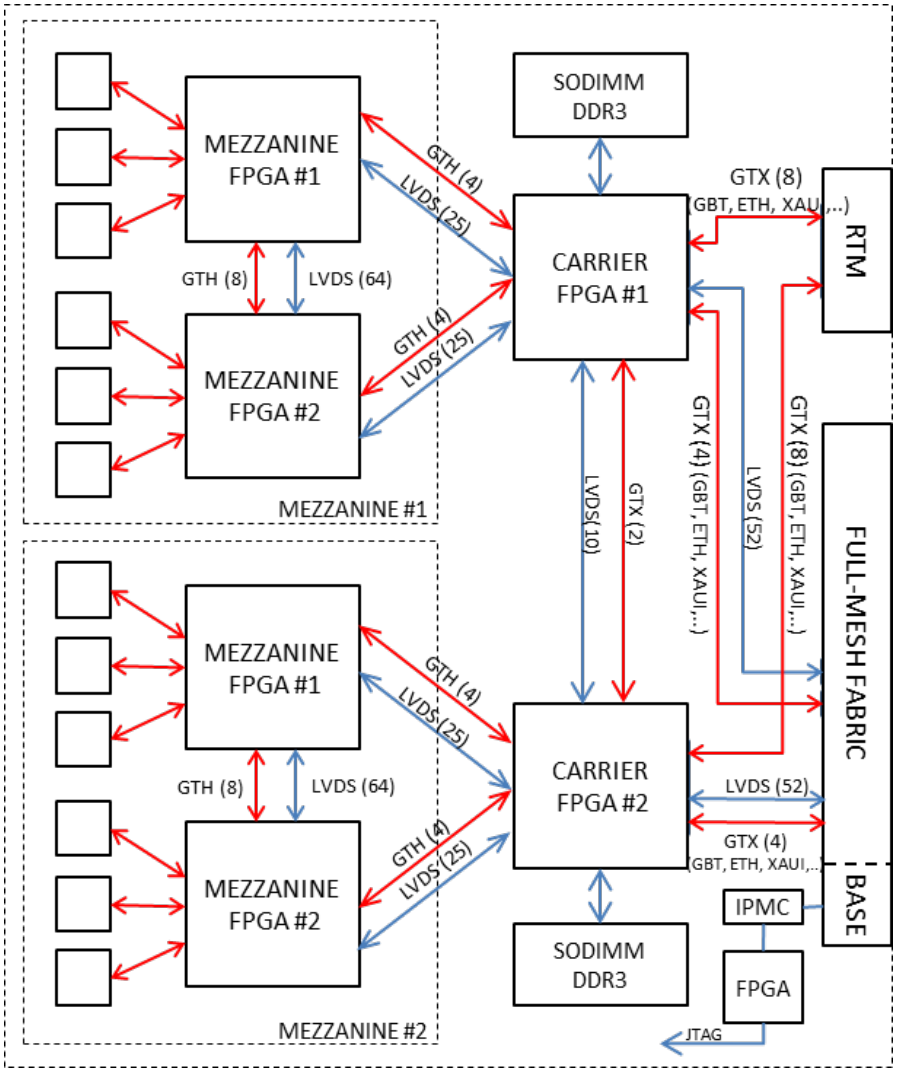

Figure 3. Simplified block diagram showing communication links of the mezzanine board and the ATCA carrier board.

The carrier board can receive reference clock from 3 sources: local oscillator, RTM (dedicated LVDS signal), or the AXIe enabled ATCA backplane. The board integrates a PLL circuit which distributes synthesized versions of the reference clock to the other components of the blade, including mezzanine boards. The PLL system can be used in 0-delay mode which allows for synchronous phase control over the entire system. In addition to the clock distribution PLL, each of the two carrier FPGA has a dedicated jitter-cleaner PLL which can be used to regenerate the clock recovered from an input serial stream. The upgraded version of the ATCA-SRS blade which will be released soon, will integrate separate clock domains for the system reference clock and the high-speed protocol reference clocks (see Table 1).

Table 1. Main specification of the current and future versions of the ATCA-SRS carrier board.

	ATCA-SRS blade Version 1	ATCA-SRS blade Version 2
Dimensions	288mm x 322mm compliant to PICMG3.x Specification	
FPGA	2 x XC6VLX130T/240T/385T	2 x Series 7 (Kintex-7 or Virtex-7)
Memory	2 x SODIMM DDR3 up to 4GB	2 x DDR3 or DDR4 SODIMM
Mezzanine Interface	50 x LVDS up to 1GHz DDR 8 x GTX 16 x LVCMOS Power 12V, 5V, 3.3V	Same Second power connector for higher current
Back-end Interface	ZONE 2: • Full-mesh connectivity • GTX (hub channel) • Clock and Trigger (AXIe compliant) ZONE 3: • 16 x GTX • CLK IN	ZONE 2: • Full-mesh connectivity • GTH (hub channels) • Clock and Trigger (AXIe compliant) ZONE 3: • 16 x GTH • CLK IN
Clock Distribution	AXIe compliant Local oscillator or RTM clock sources. PLL distribution of system and GTX clocks. System-wise phase synchronization using AXIe and 0-delay PLL.	same Separate PLL circuits for system and GTH clocks for greater flexibility.
Management	Modular IPMC Firmware upgrade through Ethernet or RTM	

On the ATCA-SRS carrier board the dedicated management part has been isolated from the two carrier FPGAs, in order to fully accommodate the management features available in ATCA standard and allow clear separation between User part and System part. The management part consists of lower profile FPGA and an IPMC extension slot, which can be used for additional module, implementing the communication with the ATCA shelf manager. The FPGA is used for peripheral management, JTAG chain configuration, or power management. The functionality of this part can be accessed either via IPMC module or dedicated USB connector on the front panel.

Mezzanine Board Specifications

1. Processing Units

The board integrates two independent FPGA processor units from the Xilinx Virtex-7 family with independent front-end optical links and independent back-end LVDS/GTH connection to motherboard FPGA.

Each FPGA unit includes its own configuration memory (BPI flash). Clock and power domains are also independent (more details given in sections 3 and 4).

For local debugging, the two FPGAs can be programmed and controlled via a local JTAG connection. In-system programming is possible via the JTAG connection present on the mezzanine interface. This connection can be controlled via the IPMI management system of the ATCA carrier.

1.1. FPGA Specifications

- Footprint: 1FFG1158
- Compatible Part Numbers: XC7VX415T, XC7VX485T, XC7VX550T, XC7VX690T

1.2. Configuration options

- Individual BPI PROM (MT28GU512)
- Local JTAG connector
- JTAG connection from carrier. This connection can be controlled via the IPMI management connection of the ATCA carrier board.

	Part Number	XC7VX415T	XC7VX485T	XC7VX550T	XC7VX690T
	EasyPath™ Cost Reduction Solutions[1]	XCE7VX415T	XCE7VX485T	XCE7VX550T	XCE7VX690T
Logic Resources	Slices	64,400	75,900	86,600	108,300
	Logic Cells	412,160	485,760	554,240	693,120
	CLB Flip-Flops	515,200	607,200	692,800	866,400
Memory Resources	Maximum Distributed RAM (Kb)	6,525	8,175	8,725	10,888
	Block RAM/FIFO w/ ECC (36 Kb each)	880	1,030	1,180	1,470
	Total Block RAM (Kb)	31,680	37,080	42,480	52,920
Clocking	CMTs (1 MMCM + 1 PLL)	12	14	20	20
I/O Resources	Maximum Single-Ended I/O	600	700	600	1,000
	Maximum Differential I/O Pairs	288	336	288	480
Integrated IP Resources	DSP Slices	2,160	2,800	2,880	3,600
	PCIe® Gen2[2]	—	4	—	—
	PCIe Gen3	2	—	2	3
	Analog Mixed Signal (AMS) / XADC	1	1	1	1
	Configuration AES / HMAC Blocks	1	1	1	1
	GTX Transceivers (12.5 Gb/s Max Rate)[3]	—	56	—	—
	GTH Transceivers (13.1 Gb/s Max Rate)[4]	48	—	80	80
	GTZ Transceivers (28.05 Gb/s Max Rate)	—	—	—	—
Speed Grades	Commercial	-1, -2	-1, -2	-1, -2	-1, -2
	Extended[5]	-2L, -3	-2L, -3	-2L, -3	-2L, -3
	Industrial	-1, -2	-1, -2	-1, -2	-1, -2

Figure 4. Overview of features of the pin compatible FPGA family which can be used with the HORX board

2. Communication Links

2.1. Front-end Optical Links

2.1.1. Optical Transceivers: Avago MicroPOD RX and TX modules

- Max Data Rate: 10.3 Gbps or 12.5 Gbps

2.1.2. FPGA Transceivers: GTH/GTX banks

- Max Data Rate GTX: 8.0, 10.3 or 12.5 Gbps (depending on FPGA speed grade)
- Max Data Rate GTH: 8.5, 11.3 or 13.1 Gbps (depending on FPGA speed grade)

2.1.3. Number of modules for each FPGA: 3 RX + 3 TX (36/36 fibers)

2.1.4. Number of modules in total: 6 Rx + 6 TX (72/72 fibers)

2.1.5. Management: Avago modules are controlled by the corresponding FPGA via I2C bus.

2.1.6. Front-end optical connectors:

- Type: 24-fiber MTP connector (12 TX + 12 RX)
- RX/TX only option: 12-fiber MTP (RX or TX)
- Total Number of connectors : 6

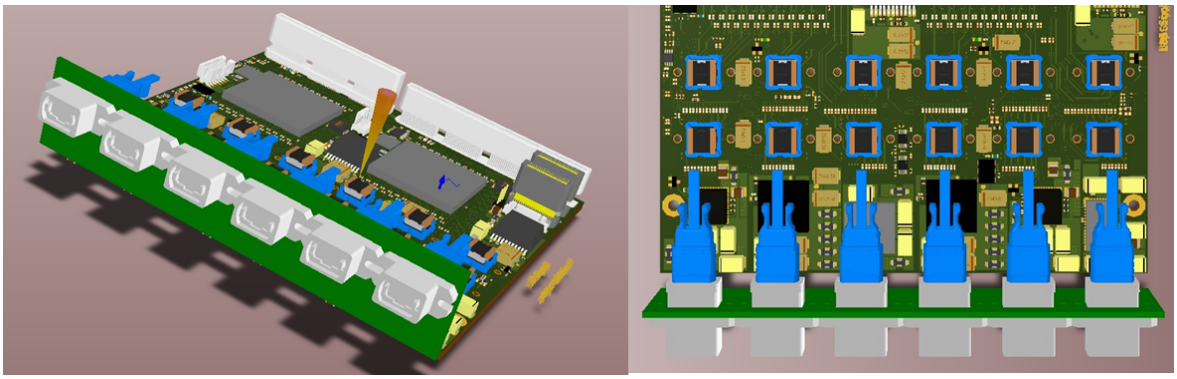

Figure 5. Illustration of the MTP connectors on the front-panel

2.2. Mezzanine connection

Both FPGAs are connected independently via the mezzanine connector to the ATCA-SRS carrier board. The connection includes high-speed GTH transceiver lanes (8), low-latency LVDS signals (50) and single-ended LVCMOS signals (16) which are directly connected to one FPGA unit on the carrier board.

From the 50 differential LVDS lines of the interface, 10 are connected to clock capable inputs on the carrier FPGA and 40 are connected to regular SelectIO pins.

The direction and functionality of the interface is not defined. Any standard or user defined protocol can be implemented.

The mezzanine interface also integrates reference clock resources for the mezzanine card and management interface.

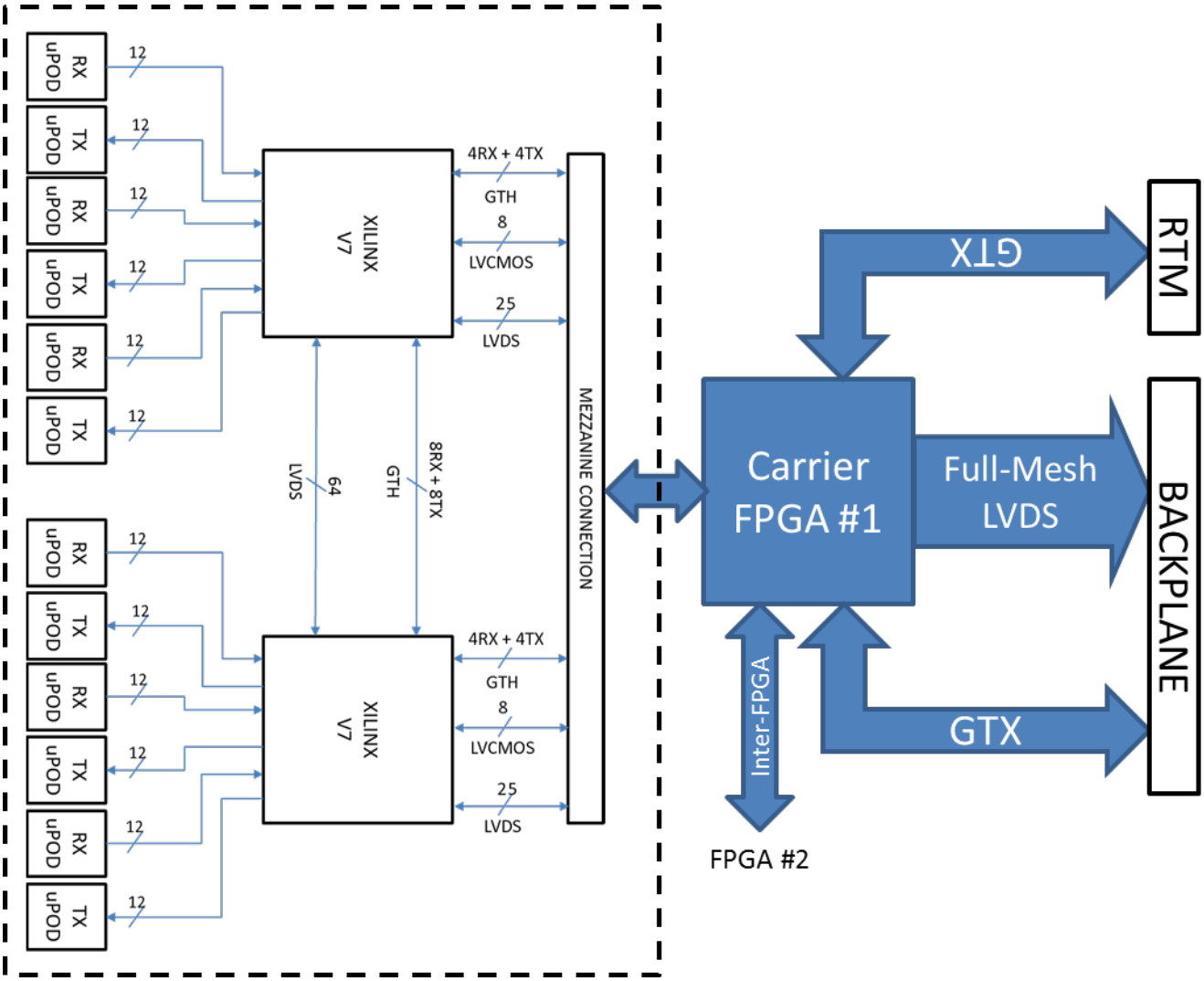

Figure 6. Diagram of the mezzanine communication links

2.2.1. LVDS Connections

- Number of lines: 25 / each FPGA
- Protocol: no protocol assumed. Each line can be used in any direction or bidirectional

2.2.2. LVCMOS Connections

- Number of lines: 8 / each FPGA
- Protocol: no protocol assumed. Each line can be used in any direction or bidirectional

2.2.3. GTH Connections

- Number of lines: 4RX+4TX / each FPGA
- Protocol: no protocol assumed.

2.2.4. Clock Signals

- System Clock (eg. 40 MHz LHC clock)
- GTH reference clock

2.2.5. Management signals

- IPMI bus from the ATCA carrier IPMC controller
- PRESENT and POWER_GOOD pins

2.2.6. Power

- 12V / 15A
- Other voltages (not used): 3.3V, 5V

2.3. Inter-FPGA connection

The inter-FPGA bus bridges the two FPGA together, allowing for high-bandwidth and low-latency data transfer over a combination of high-speed transceivers (12) and LVDS lines (64). Depending on the application, the bridge may offer the possibility to join together the processing power of the two FPGA units into a single powerful processor.

The direction or functionality of the LVDS bus is not constraint; the user may implement any protocol for the data exchange.

2.3.1.LVDS Connection

- Number of lines: 64
- Protocol: no protocol assumed. Each line can be used in any direction or bidirectional

2.3.2.GTH Connection

- Number of lines: 8RX+8TX
- Protocol: no protocol assumed.

3. Clock Distribution

3.1. Mezzanine Clock Sources

3.1.1.Mezzanine Interface

- System Clock. Generated on the carrier motherboard by a 0-delay PLL (LMK03200) from the following possible sources:
  - Local Oscillator on ATCA carrier board
  - RTM (discrete LVDS signals or derived from serial stream)
  - Backplane (AXIe standard)
- GTH Reference Clock.
  - On current carrier design, the GTH clock is generated by the same PLL circuit as the system clock. The PLL ratio may differ from the one of the system clock.
  - In upcoming upgrade of the motherboard, a separate PLL is foreseen for the GTH reference clock. The reference sources for this secondary PLL will be the same as for the system-clock PLL (local oscillator, RTM or AXIe clock).

3.1.2.Local oscillators

There are 3 local oscillators for each FPGA unit. One is used for deriving system clocks supplied to FPGA fabric and the other two are used to supply reference clocks to the GTH transceivers.

The frequency of the local oscillators is user defined.

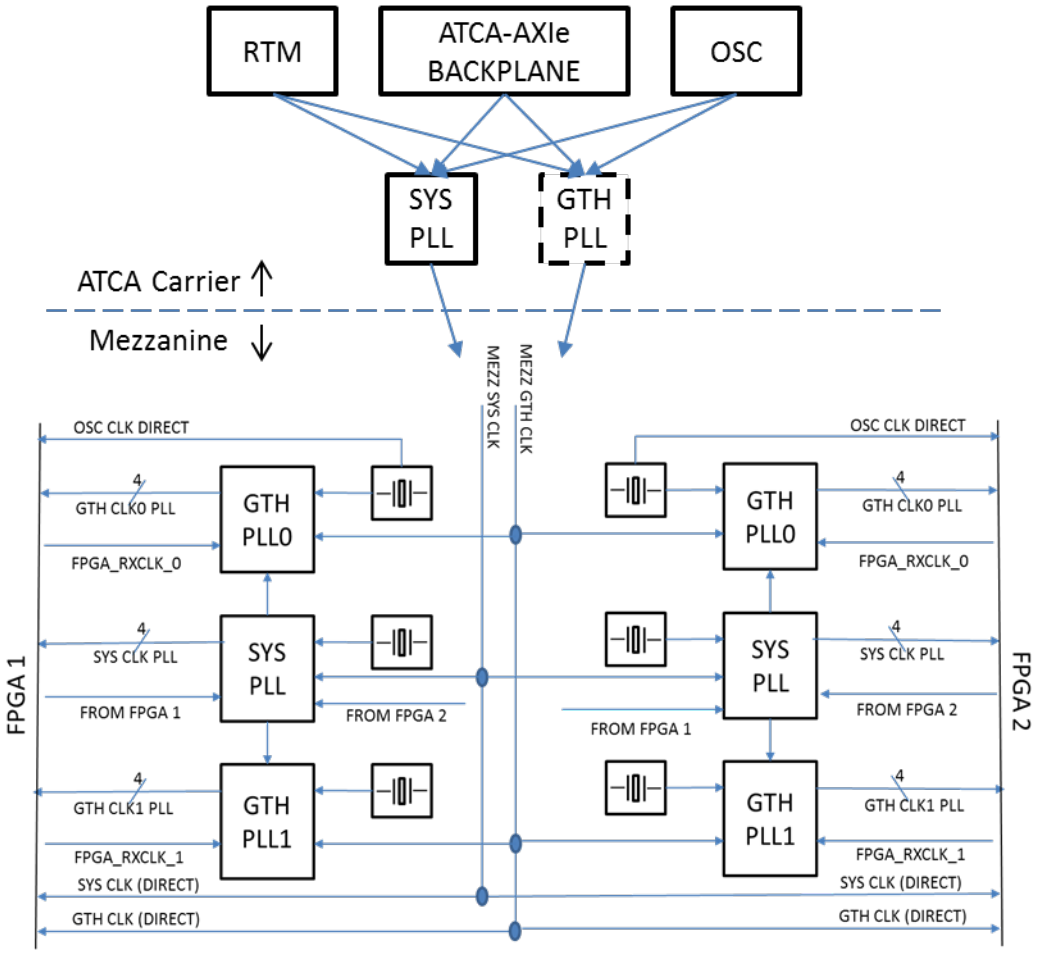

Figure 7. Diagram of the Clock Distribution scheme.

3.2. Clock conditioners

Each FPGA system has its own dedicated clock distribution tree, consisting of one 0-delay PLL (LMK03200 - Precision 0-Delay Clock Conditioner with Integrated VCO) for the system clock and 2 PLLs (LMK03806 - Ultra-low jitter clock generator) for the serializers clock references. The two clock sources (system and GTH reference) can be supplied from the carrier motherboard via the mezzanine interface. Optionally, the source can be switched to dedicated local crystal oscillators or clock outputs from the FPGA (i.e. recovered from incoming optical links).

The second GTH reference clock domain allows for greater flexibility, giving the possibility to implement two independent protocols which cannot share the same clock seed (eg. GBT derived from 40.08MHz LHC clock and 10GbE derived from 156.25MHz).

The system and GTH clocks from the carrier board and the local system oscillator are fed via fanout buffers directly to the FPGA, bypassing the PLL circuits. These may be used as start-up clocks or when further clock synthesis is not required.

The PLL circuits are managed by the FPGA units. Optionally, the management interface can be connected to the emulated IPMI extension in the FPGA firmware (see Management section 5)

3.2.1. System-Clock PLL

- Purpose: Jitter cleaner, multiple clock synthesizer and/or 0-delay buffer for the system clock
- Type: TI LMK03200 (LMK03200 - Precision 0-Delay Clock Conditioner with Integrated VCO)
- Sources:
  - Mezzanine System Clock
  - Local Oscillator 0
  - FPGA processor 0
  - FPGA processor 1
- Outputs:
  - FPGA (4 outputs)
  - GTH-Clock PLL 1 (1 output)
  - GTH-Clock PLL 2 (1 output)

3.2.2. GTH Clock PLL 1

- Purpose: Jitter cleaner, multiple clock synthesizer for GTH clock reference
- Type: TI LMK03806 (LMK03806 - Ultra-low jitter clock generator)
- Sources:
  - Mezzanine GTH Clock
  - Local Oscillator 1
  - FPGA recovered clock
  - System Clock PLL (derived from System Clock PLL, see section 3.2.1)
- Outputs:
  - FPGA GTH REFCLK (4 outputs)
  - FPGA Fabric (2 outputs)

3.2.3.GTH Clock PLL 2

- Purpose: Additional jitter cleaner, multiple clock synthesizer for GTH clock reference. This additional PLL increases the flexibility, allowing two different protocols to be implemented in the front-end and/or backend system.
- Type: TI LMK03806 (LMK03806 - Ultra-low jitter clock generator)
- Sources:
  - Mezzanine GTH Clock
  - Local Oscillator 1
  - FPGA recovered clock
  - System Clock PLL (derived from System Clock PLL, see section 3.2.1)
- Outputs:
  - FPGA GTH REFCLK (4 outputs)
  - FPGA Fabric (2 outputs)

GTH PLL 0	GTH PLL 1	FPGA Pin	GTH bank	Link Allocation
ClkOut0	-	MGTREFCLK0_115	MGTHQUAD_114	uPod 0
			MGTHQUAD_115	
-	ClkOut0	MGTREFCLK1_115		
			MGTHQUAD_116	
ClkOut2	-	MGTREFCLK0_118	MGTHQUAD_117	uPod 1
			MGTHQUAD_118	
-	ClkOut2	MGTREFCLK1_118	MGTHQUAD_119	
ClkOut4	-	MGTREFCLK0_215	MGTHQUAD_214	uPod 2
			MGTHQUAD_215	
-	ClkOut4	MGTREFCLK1_215		
			MGTHQUAD_216	
ClkOut6	-	MGTREFCLK0_218	MGTHQUAD_217	Carrier GTH link
			MGTHQUAD_218	Inter-FPGA bridge
-	ClkOut6	MGTREFCLK1_218	MGTHQUAD_219	

Table 2. GTH PLL allocation table

3.2.4. PLL Management

- All 3 PLLs are in direct control of the FPGA processing unit
- Optionally the PLLs can be controlled via the emulated IPMI extension in the FPGA firmware (see Management section 5.1.2)

3.2.5. Direct clock connections to FPGA processing unit (start-up clocks)

- System Clock from carrier
- GTH Clock from carrier
- Local Oscillator 0

4. Power Distribution

4.1. Power sources

4.1.1. 12V power originates from the carrier board (ATCA supply chain)

4.1.2. Optional 12V connector, for benchtop operation

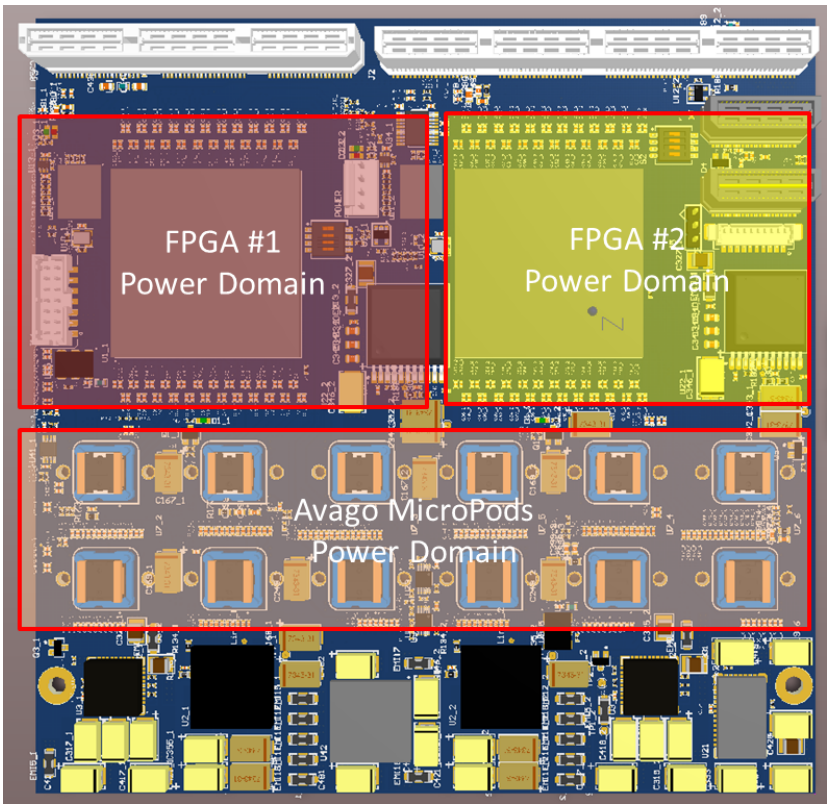

Figure 8. Separate Power Domains

4.2. FPGA power domains

The following power domains are duplicated for each FPGA unit:

- 1.0V@20A : core and BRAM
- 1.8V@1A : VAUX and IO
- 1.8V@1A : MGT VCC
- 1.0V@10A : MGT AVCC
- 1.2V@4A : MGT AVTT

4.3. Common power domains

- 3.3V@4A : optical modules
- 3.3V@4A : PLLs
- 2.5V@6A : optical modules

4.4. Total Power Rating:

4.4.1. Power Output: 120W

4.4.2. Power Input : 150W

4.5. Power Management

- All power circuits are controlled by an MMC microprocessor unit which is connected to the IPMI bus of the carrier board
- Separate sequencer circuits for the 2 FPGA power domains
- Phase control circuit for the DC/DC converters to avoid high current surge from the 12V input supply.

5. Board Management

5.1. IPMI Bus

The local IPMI bus is connected to the IPMI bus of the carrier board via the mezzanine connector. A local MMC controller will monitor the status and control local power converters or can reset the two FPGA independently.

5.1.1.MMC Microcontroller

- Type: Freescale Kinetis K10P64 Cortex-M0 MCU
- Inputs
  - Power monitors of the local DC/DC converters
  - Temperature sensors
- Outputs
  - Power enable for the two FPGA processing units (independent)
  - Reset signals for the two FPGA (independent)
- IPMI extension
  - I2C bus connected to FPGAs for representing their sensors on IPMI
- Auxiliary control
  - RS232 for debug/benchtop operation

5.1.2. FPGA Emulated IPMI extension

The MMC is connected to the two FPGAs over local I2C buses.

The MMC represents the sensors implemented on FPGA to the IPMI management infrastructure. This allows controlling FPGA parameters or other peripherals which are connected to the FPGAs (i.e. PLL settings, Xilinx XADC – power levels and die temperature).

The communication between MMC and FPGA is made over simple I2C master read/write transactions. No IPMI software is required on FPGAs, the firmware will implement only a simple memory emulator which is read by the MMC.

Schedule and Responsibilities

The HORX mezzanine board is currently being manufactured, test results are expected soon. The ATCA-SRS board is already available and it was tested extensively within the RD51 collaboration. The next version with of the board is now being designed and is expected in Q2/Q3 2015.

The design and realization of the HORX mezzanine board was externalized to a specialized company (Samway Electronics, Bucharest) on order of IFIN-HH, Bucharest. The final design is part of the contracted deliverables therefore all design sources are in possession of our institute. The ATCA-SRS Carrier Board is being manufactured under a licensing contract between Eicsys Gmbh, Germany as

licensee and the IP Holder Consortium within RD51 collaboration (CERN, IFIN-HH and Polytechnic University of Valencia), as licensor. The licensing contract is not exclusive.

Firmware and software modules for various functions of the system were already developed and tested within RD51 Collaboration, to which ATLAS institutes already have access to.

Apart for the two companies mentioned which will provide support for the two boards, the following team from our institute will be in charge of the design, specifications and firmware & software support:

- Sorin Martoiu, Research Engineer
- Michele Renda, Graduate Student
- Gabriel Stoicea, Physicist

Appendix A. Application Example of the HORX Mezzanine card as NSW Trigger Processor

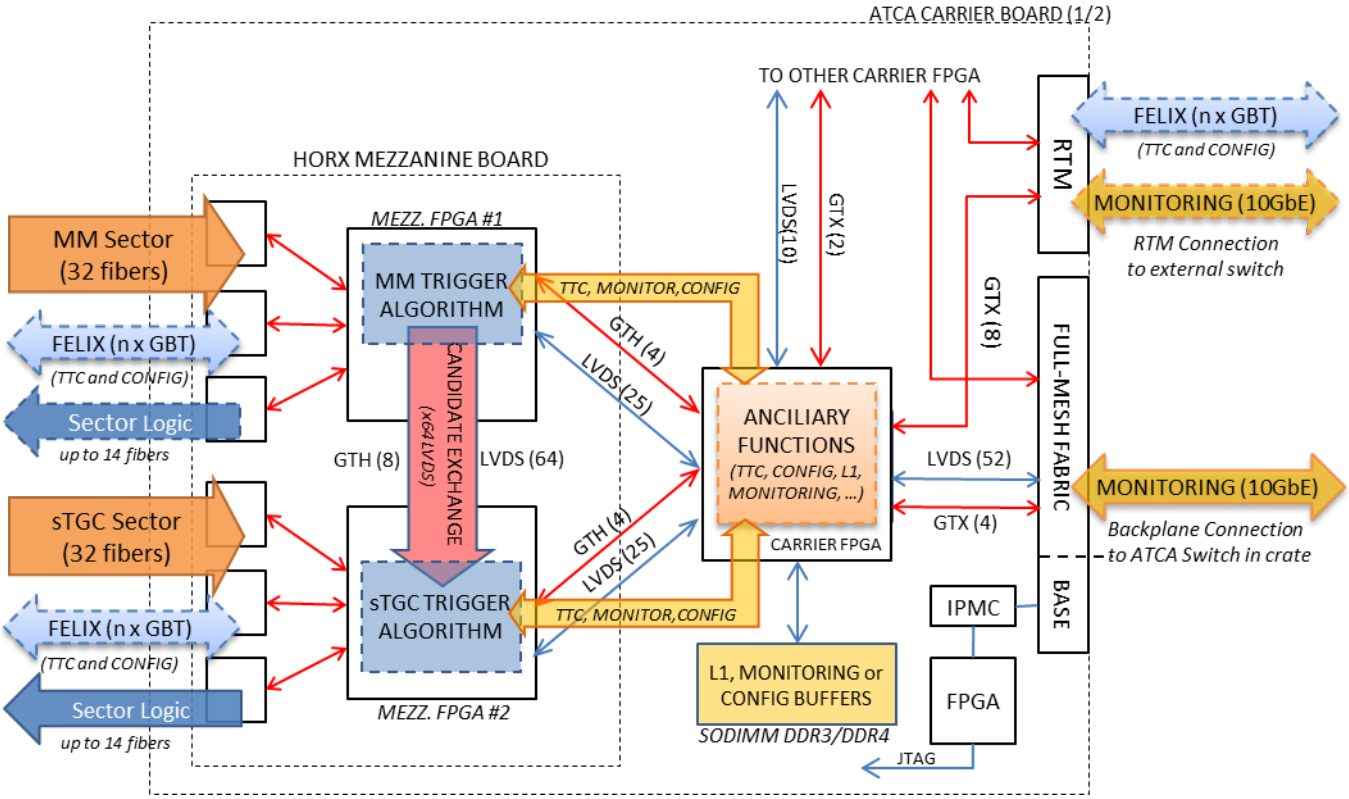

Figure 9. Example of using the HORX mezzanine as Trigger Processor Card for NSW

Figure 9 shows how the High-Density Optical Receiver mezzanine can be used for the NSW Trigger Processor application. The 12 front-end MicroPod modules can be mapped to the front-end fibers coming from a NSW sector, connecting each detector technology (Micromegas and sTGC) to one of the two mezzanine FPGAs, which will implement the corresponding MM and sTGC trigger algorithms. The same front-end interface can be used for the optical output connection to the Sector Logic, which

requires up to 14 fibers. The flexible clocking scheme implemented by the mezzanine board allows the use of different serial speeds on the TX and RX GTH transceivers.

The track vectors computed by the MM and sTGC processors have to be combined before being sent to the Sector Logic. The low-latency 64-bit LVDS bus can transport data from one processor to the other with minimum latency. Considering a conservative 640MHz DDR transfer rate, one can transfer up to 8 24-bits words between processors in less than 2.5 ns.

The FPGA on the ATCA carrier board can be used for the Ancillary Functions of the NSW Trigger Processor (TTC, configuration, Level-1 Data Buffering, monitoring, etc.). Each of the two FPGAs will correspond to one full sector of the NSW detector. In this way the mezzanine FPGAs are off-loaded of the ancillary functions, leaving more programmable resources for the trigger algorithms. The dedicated SODIMM memory module can be used for buffering Monitoring and Level-1 data, and/or for storing simulated input data which can be used to test the performance of the processing algorithm. The current version of the ATCA blade offers support for up to 4GByte DDR3 memory for each carrier FPGA. Next blade version might increase the memory limit and/or integrate DDR4 technology.

Optionally, low-latency data can be exchanged between sectors either via on-blade inter-FPGA LVDS bus or via the full-mesh LVDS backplane connectivity.

The 8 x GTX transceivers connected to the RTM interface, for each of the two carrier FPGAs, provide enough GBT connections for TTC, Level-1output data and configuration to/from FELIX. Optionally, some or all of the connections to FELIX may be realized on the front-end interface of the mezzanine boards, which have space for 4 bidirectional links.

Monitoring of the Processor operation can be done by a private network using, for example, 10Gb Ethernet protocol. This can be achieved either using RTM connection to an external device or by using the backplane hub interface with an ATCA Switch Blade. This Switch blade can be an off-the-shelf commercial device.

Firmware for 10Gb Ethernet communication, board management or DDR3 memory interface were designed and tested on hardware within the RD51 Collaboration.

Appendix B. Compatibility with Phase-II

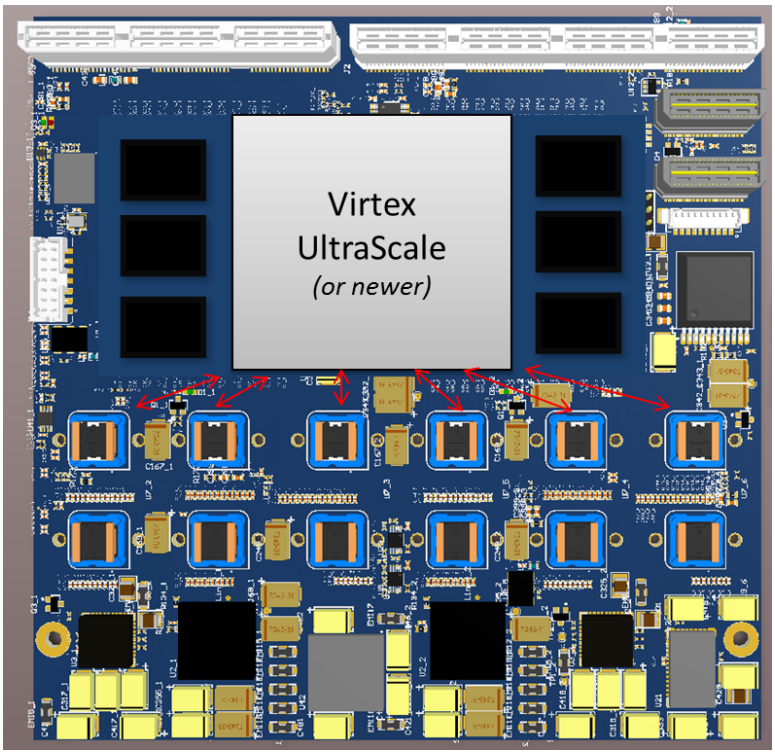

Figure 10. HORX upgrade example using Virtex UltraScale or other future series device. All front-end optical transceiver modules can be connected a single processing device.

In Phase II, with higher latency budget and taking advantage of newer FPGA technology, the NSW trigger processor may do a more refined calculation of muon pointing and momentum, or improve robustness and redundancy, by processing together the prompt signals coming from all detector layers, Micromegas and sTGC alike. The form factor of the HORX mezzanine board already allows that all uplink fibers from the Micromegas and sTGC sections of a NSW sector be connected to the same board. Future FPGA devices are capable of integrating more and more high-speed transceiver links, therefore it is feasible to integrate a single FPGA device on the mezzanine card which is able to aggregate data from both detector technologies. As an example, the Virtex UltraScale series is foreseen to integrate up to 120 high-speed transceivers with signaling speed of 16.3 and 30.5 Gb/s, as well as up to more than 4 million logic cells (http://www.xilinx.com/publications/prod_mktg/ultrascale-virtex-product-table.pdf). These devices will therefore be a valid candidate for integrating the foreseen NSW trigger processing algorithm for the Phase II operation.