

Ultrascale+ for the new ATLAS calorimeter trigger board dedicated to jet identification

B. Bauss, V. Bücher, R. Degele, H. Herr, C. Kahra, S. Rave, E. Rocco, J. Schäfer, U. Schäfer, J. Souza, Duc Bao Ta, S. Tapprogge, M. Weirich

Institut für Physik, Johannes Gutenberg Universität, Mainz Germany

A. Brogna

Detector Lab, PRISMA cluster of Excellence, Mainz Germany

Abstract—To cope with the expected increase in luminosity at the Large Hadron Collider in 2021, the ATLAS collaboration is planning a major detector upgrade to be installed during Long Shutdown 2. As a part of this, the Level-1 trigger based on calorimeter data, will be upgraded to exploit the fine granularity readout using a new system of Feature EXtractors (FEXs), which each reconstruct different physics objects for the trigger selection. The jet FEX (jFEX) has been conceived to identify small/large area jets, large area tau leptons, missing transverse energy and the total sum of the transverse energy. The use of the latest generation Xilinx field programmable gate array (FPGA), the Ultrascale+, was dictated by the physics requirements which include substantial processing power and large input bandwidth, up to $\sim 3\text{Tb/s}$, within a tight latency budget of less than 390 ns. The modular design of the jFEX board allowed for an optimal routing of a large number of high speed signals within the limited space of an ATCA board. To guarantee the signal integrity, the board design has been accompanied by simulation of the power, current and thermal distribution. The printed circuit board has a 24-layer stack-up and uses the MEGTRON6 material, commonly used for signal transmission above 10 Gb/s. The jFEX system, consisting of 6 boards, will be produced by the end of 2018 to allow the installation and commissioning of the full system in time for the LHC restart at the beginning of 2021.

I. INTRODUCTION

The ATLAS experiment [1], one of the multi-purpose detectors at the LHC, has planned a major detector upgrade for Run 3 [2] (from 2021 to 2026) to adapt data readout capabilities for increased beam intensity up to an instantaneous luminosity of $2.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$.

A significant part of this upgrade is related to the Level-1 Calorimeter Trigger (L1Calo), which uses data from the calorimeters to reconstruct different physics objects for the first trigger selection. In order to exploit the fine granularity readout with a trigger rate at up to 100 kHz, a new system of Feature EXtractors will be used for trigger selection. There are three dedicated module types: electron FEX (eFEX) [3], jet FEX (jFEX) and global FEX (gFEX)[4]. During the commissioning phase of the new system, the legacy system will run in parallel but will be retired once the new system is fully commissioned. Fig. 1 shows the Level-1 Calorimeter trigger system at the beginning of LHC Run 3.

The jFEX has been designed to identify jets, taus and energy related quantities. The use of the latest generation of the Xilinx Virtex Ultrascale+ FPGAs is dictated by the physics requirements which include a significant processing power and

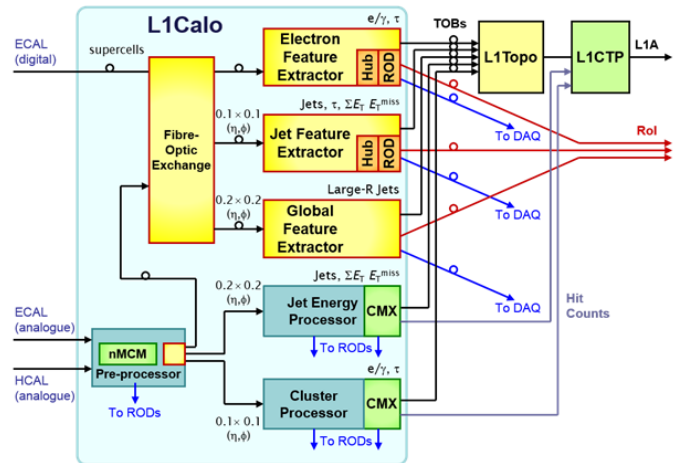


Fig. 1. Level-1 Calorimeter trigger system at the beginning of LHC Run 3. New components are shown in yellow. The new system will run in parallel with the current legacy system until the new system will be fully commissioned. [1]

an enormous input bandwidth, up to $\sim 3\text{Tb/s}$, within a latency budget of less than 390 ns.

II. BOARD REQUIREMENTS AND FEATURES

The jFEX board is meant to identify jet and large-area tau candidates and calculates observables indicative of the presence of hadronic jets or neutrinos. Its input data comes from the central and forward calorimeter with various granularities over the η range of ± 4.9 .

In order to fully exploit the calorimeter information, the jFEX system bandwidth has been maximized whilst maintaining the capability of large area jet clustering. Additionally, a large processing power is required on the new system due to the complexity of the algorithms that will be implemented.

Several algorithms for the jet identification will run on the jFEX, and the current baseline algorithm identifies the local maximum using a sliding window algorithm and for each local maximum summing the surrounding cells with a Gaussian weighting distribution.

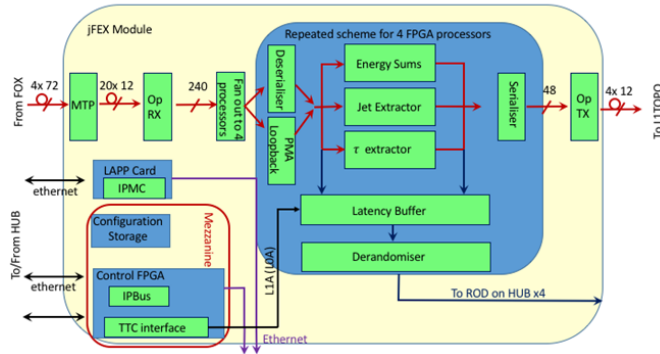


Fig. 2. jFEX block diagram. The sketch shows the functionality of one of the four FPGAs. In the bottom left the functionalities of the extension mezzanine are shown. Data and module monitoring are performed through the connection to the Hub and the ROD via the backplane.

The block diagram in Fig.2 shows the jFEX processing functionality. The digitised calorimeter data are fed into the board through the Fibre Optical Plant (FOX) and then received on the jFEX through the electro-opto devices, miniPODs [5], and sent afterwards to the processor FPGAs. The whole ring in ϕ is covered by each module, with overlapping regions between processors in η and ϕ duplicated on board using PMA loop-back.

The algorithm results are sent to the Level-1 Topological Processor board [6] (L1Topo) as Trigger Objects (TOBs) and also through miniPODs to the Topological Fibre Optical Plant (TopoFOX).

The shelf communication and readout of the jFEXs system is performed via the FEX system ATCA switch module (Hub) and the Readout Driver (ROD) mezzanine [2].

III. BOARD OVERVIEW

The jFEX is an ATCA [7] card equipped with 4 processor FPGAs, 20 miniPODs for data reception (5 per FPGA) and 4 miniPODs for data transmission (1 per FPGA). Each board receives data through a total of 240 optical inputs (60 per FPGA) and transmits the TOB through 48 optical outputs (12 per FPGA) to the L1Topo boards. For the data sharing between the processor FPGAs 58 Multi-Gigabit Transceivers (MGTs), 29 transmitters and 29 receivers, are used per FPGA pair.

The jFEX design is modular and consists of four identical blocks, each containing 1 FPGA and 6 miniPODs (5 receivers and 1 transmitter). This approach allows a symmetric design and consequently, an equalization of the length of the traces between processors and miniPOD and also for the FPGA inter-connections. The bus lengths are minimised, without any crossing, in order to improve the signal integrity.

The large input bandwidth and processing power required for the data were the main factor for the FPGA choice to be used on the board. Therefore, a large number of MGTs and logic resources, with a reasonable power consumption, were the most relevant features for the processor FPGAs.

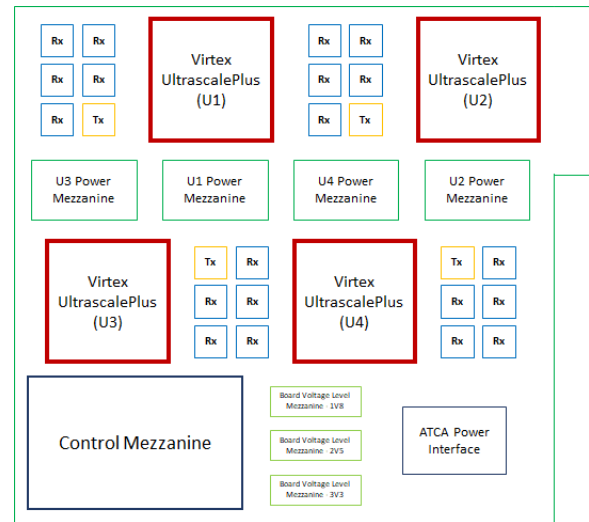


Fig. 3. The jFEX simplified block diagram. The board is divided in four identical blocks, each consisting of one FPGA and six miniPODs.

The Xilinx Virtex Ultrascale+ XCVU9P-2PFLGA2577E is the chosen device.

In order to gain flexibility for future board changes, the board control functionalities are located on a mezzanine card. This philosophy allows a smooth and reliable way of upgrading control functions and components without affecting the main board. Additionally, this approach guarantees smooth and reliable board operation and compatibility with the surrounding trigger system for the next twenty years of data taking.

The design strategy based on mezzanines is also used on the power distribution for the voltage rails with current consumption higher than 3 A, which use switching regulators. Fig. 3 shows an overview of jFEX hardware, with its mezzanines.

A. Control Mezzanine

Many of the (non-real-time) services required on the jFEX are provided by the Control Mezzanine card. It carries mainly module control, clock and configuration circuitry. It also provides initialization circuitry for the FPGAs and acts as an interface to environmental monitoring devices.

The "intelligent" module controller is an FPGA which handles incoming IPbus requests and forwards the data and control packets to the processors on the mainboard via MGT links.

A significant part of the clock (re)generation circuitry is located on the jFEX Control Mezzanine, the individual clock trees for MGT reference clocks (real-time data path, Timing, Trigger and Control (TTC) data and IPbus) and global clocks are actively fanned out on the main board.

The IPbus communicates with its control PC(s) via an Ethernet Phy chip. 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 via the backplane.

The TTC data links are received from the backplane at 6.4 Gb/s from each Hub/ROD module. The data are routed into the fanout chip where they are forwarded to the processor FPGAs.

The configuration of all four processor FPGAs is controlled from the mezzanine. To this end all signal lines required are routed to the mezzanine. The configuration mode used is Master SPI x4.

1) *Clock Distribution*: The high number of MGTs used on the jFEX require a complex scheme for distribution of all MGT reference clocks and the FPGA global clock, in order to minimize the number of signals routed on the PCB. Additionally, two constraints should be respected:

- Based on the the GTY user guide the reference clocks for a QUAD can also be sourced from up to two QUADs below or above;
- There should be no crossing of clocking signals between different FPGA Super Logic Regions (SLR).

The jFEX module receives the LHC clock on two differential pairs (Fabric Interface) via the ATCA backplane. Later the signal is distributed to one of the inputs of the jitter cleaner chip placed on the Control Mezzanine and to all processor FPGAs via several fanout chips.

From its inputs the PLL chip can generate clocks of frequency $n \times 40.079$ MHz, up to 712.5 MHz. This flexibility allows the multi-Gb/s links on the jFEX to be driven at a large range of different rates. Fig. 4 shows the clock distribution on the jFEX module.

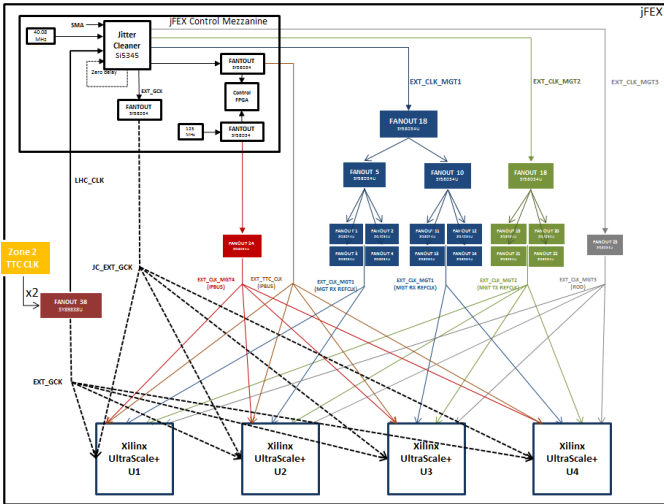


Fig. 4. jFEX clock scheme. The LHC clock is received via the ATCA backplane and routed to the jitter cleaner placed on the Control Mezzanine, to be distributed to the FPGA MGT reference clocks and fabric, via several fanout chips.

B. Power Mezzanine

The power supplies for the processor FPGAs are mainly located on mezzanines. This approach allowed a careful evaluation of ripple noise and stability of each design, before connecting the mezzanines to the jFEX and powering up the board.

The jFEX Power Mezzanines are based on the TDK-Lambda iJX series of non-isolated DC/DC converters. Three main factors were taken into account while making this selection:

- High current consumption requirement by the Virtex Ultrascale+;
- The MGTs on the Virtex Ultrascale+ are very sensitive to noise in its power rails ($V_{MGTAVCC}$, $V_{MGTAVTT}$ and $V_{MGTVCCAUX}$) requiring a maximum 10 mV_{pp} of noise on the FPGA power pins over the band from 10 kHz to 80 MHz;
- Out-of-the-box monitoring tool for output voltages, currents and temperatures (on-board PMBus monitoring).

For the $V_{MGTAVCC}$, the iJB (max. 60 A output current) is used, while $V_{MGTAVCC}$, $V_{MGTAVTT}$ and V_{CCIO} uses the iJA (max. 35 A output current). The iJA DC/DC converter is also used to supply the board level voltages 3.3 V and 2.5 V.

Two different mezzanines have been designed for the jFEX: the *FPGA Power Mezzanine* equipped with two iJA for $V_{MGTAVCC}$ & $V_{MGTAVTT}$ and one iJB for V_{CCINT} and the *Board Level Voltage Mezzanine* equipped with one iJB, for V_{CCIO} , 3.3 V and 2.5 V. The *Board Level Voltage Mezzanine* are directly connected to the jFEX.

Each processor FPGA have a single independent *FPGA Power Mezzanine*. The V_{CCIO} is the only voltage rail shared between all FPGAs, being supplied by a single *Board Level Voltage Mezzanine*. The 3.3 V and 2.5 V voltage rails are also supplied by separate *Board Level Voltage Mezzanine*. In total the jFEX board hosts four *FPGA Power Mezzanines* and three *Board Level Voltage Mezzanines*.

For safety, apart from the over-voltage/current protection mechanism intrinsic to the iJX DC/DC regulators, the jFEX Power Mezzanines also include an over-voltage protection circuit (crowbar).

IV. HARDWARE CHARACTERIZATION

The jFEX prototype was produced in November 2017 and extensively tested during the following 12 months. The test are described below and focus on the power consumption, ripple measurement on the MGTs power pins, thermal dissipation, MGT reference clock quality and validation of the optical inputs and parallel IOs. Fig. 5 shows a picture of the jFEX prototype.

A. Ripple Measurement

The MGTs on the Virtex Ultrascale+ are very sensitive to noise in its power rails ($V_{MGTAVCC}$ and $V_{MGTAVTT}$) requiring a maximum 10 mV_{pp} of noise on the FPGA power pins over the band from 10 kHz to 80 MHz. Therefore, a ripple measurement was performed on each MGT power rail. Table I shows the measured noise which is within the Xilinx specifications.

B. Power Consumption and Thermal Dissipation

The FPGAs are individually powered by mezzanines which were separately tested before connecting them to the main board.

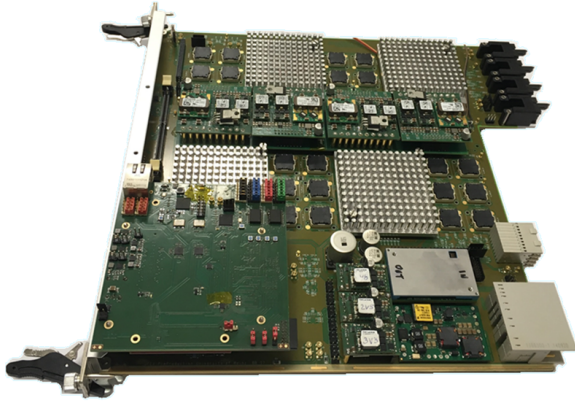


Fig. 5. Picture of the jFEX final prototype.

TABLE I

RIPPLE MEASUREMENT OF $V_{MGTAVTT}$ AND $V_{MGTAVCC}$ (PEAK TO PEAK) FOR ALL THE FOUR FPGAS. THE VALUE ARE $< 10mV$ SO WITHIN THE XILINX SPECIFICATIONS.

FPGA	$V_{MGTAVCC0.9V}$ (mV)	$V_{MGTAVTT01.2V}$ (mV)
U1	1.63	1.51
U2	2.99	2.90
U3	4.02	3.17
U4	4.07	2.79

In order to verify the reliability of the power design, a high current test was performed on the jFEX using IBERT cores with all MGTs enabled simultaneously. Table II shows the highest power consumption of all the four FPGAs when all MGTs are activated at 11.2 Gb/s.

TABLE II

DETAILS OF POWER CONSUMPTION OF THE FOUR FPGAS.

Supply	Voltage Rail	Current (A)
Board Level Voltages	1V8	0.47
	2V5	14.55
	3V3	7.42
FPGA processor U1	$V_{CCINT} 0.85V$	23.75
	$V_{MGTAVCC} 0.9V$	10.95
	$V_{MGTAVTT} 1.2V$	20.42
FPGA processor U2	$V_{CCINT} 0.85V$	20.41
	$V_{MGTAVCC} 0.9V$	10.75
	$V_{MGTAVTT} 1.2V$	20.43
FPGA processor U3	$V_{CCINT} 0.85V$	17.17
	$V_{MGTAVCC} 0.9V$	10.45
	$V_{MGTAVTT} 1.2V$	20.14
FPGA processor U4	$V_{CCINT} 0.85V$	18.82
	$V_{MGTAVCC} 0.9V$	10.73
	$V_{MGTAVTT} 1.2V$	20.21
Total(W)		265.86

The Vivado Design Suite provides the possibility to estimate the FPGA power consumption [12] and the current estimates are reported in Table III. The results obtained of the jFEX prototype are consistent with the estimations from the Vivado Design Suite, except for the $V_{CCINT} 0.85V$, which depends on the parameters used during the tests.

TABLE III

ULTRASCALE+ POWER CONSUMPTION ESTIMATED WITH VIVADO.

Vivado Current Estimation (A)	
$V_{CCINT} 0.85V$	13.22
$V_{MGTAVCC} 0.9V$	10.326
$V_{MGTAVTT} 1.2V$	22.111

The Xilinx Virtex Ultrascale+ FPGA allows a junction temperature up to 125 °C, however during all the tests the prototypes had the automatic FPGA shutdown set to 85°C. Moreover the temperature value reported for the “U1” in Table IV, comes from the missing air guides and from not closed jFEX front panel (single height instead of double, implying a not fully closed shelf), without considering the option of increasing further the crate fan speed.

TABLE IV

THE FPGA TEMPERATURES READ OUT FROM THE VIVADO SYSTEM MONITORING, FOR THE JFEX INSTALLED IN THE CRATE WITH THE FAN’S SPEED OF 10/15 WITHOUT THE JFEX FRONT PANEL.

FPGA	Temperature (°C)
U1	83.20
U2	69.30
U3	54.30
U4	62.30

Figs. 6 and 7 are pictures taken with the thermal camera when the prototype was switched-on: on the bench without any configuration FPGA on the processor FPGAs (Figs. 6); on the crate with all the MGTs enabled, during an IBERT test at 11.2 Gb/s (Figs. 7). The FPGAs temperatures were monitored via XADC and measured values are shown in Table IV.



Fig. 6. Thermal picture of the jFEX prototype switched-on on the bench.

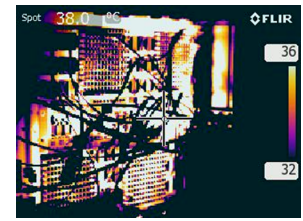


Fig. 7. Thermal picture of the jFEX final prototype in the shelf.

The conclusions drawn from this test are that the heat sink is adequate for the FPGA and the implementation of air guides will facilitate the cooling uniformity of the board.

C. Clock Phase Noise

The Xilinx Virtex Ultrascale+ specifications for the MGT reference clock [13], MGT RFLCLK, are defined in the frequency domain. Therefore to measure the jitter of the clock a spectrum analyser has been used. The probed clock signal is gathered through MMCX connectors on the board. Fig. 8 shows the measurement results, while table V compares it with

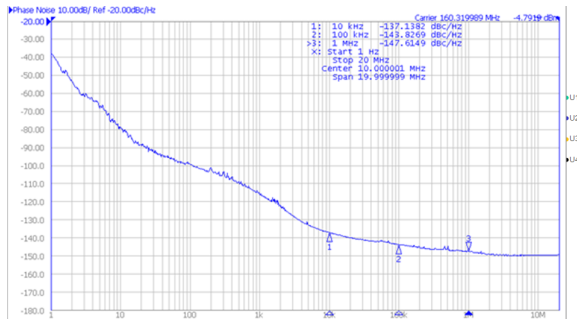


Fig. 8. Phase noise measurement with Spectrum Analyzer.

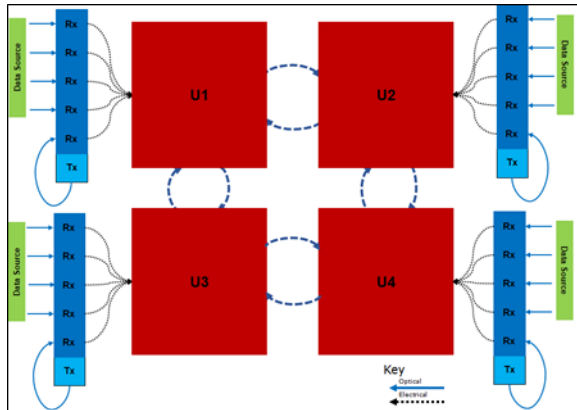


Fig. 9. Setup used for testing all 240 input fibres simultaneously.

the Xilinx specifications. The clock signal is within the Xilinx specs.

TABLE V
JITTER MEASUREMENT RESULTS FOR THREE FREQUENCIES COMPARED WITH XILINX SPECIFICATIONS.

Jitter measurement at three different frequencies		
Offset Frequency	Xilinx Specs	jFEX Meas.
10 kHz	-112 dBc/Hz	-137.14 dBc/Hz
100 kHz	-128 dBc/Hz	-143.83 dBc/Hz
1 MHz	-145 dBc/Hz	-147.615 dBc/Hz

D. Validation of Optical Inputs and Outputs

The jFEX prototype has 240 MGTs inputs, with 232 duplicated on board via PMA loopback. In order to validate the board design all the MGTs must run concurrently at 11.2 Gb/s (BER smaller than 10^{-15}). For that purpose an IBERT core was used with all MGTs enabled.

Fig. 9 shows the fibre routing between the miniPODs and the processor FPGAs in the test setup. The LHC clock was received via backplane at a frequency of 40.08 MHz.

The IBERT tests (using the PRBS31) at 11.2 Gb/s ran over 36 hours reaching a BER $< 10^{-15}$.

Fig. 10 and 11 shows the distribution of the open area of the eye diagrams for all 240 inputs and for the data duplication via PMA loopback, respectively. Some channels have a smaller

opening area and the explanation is related to the quality of the fibre connection to the miniPODs (some cleaning was necessary to remove some dust) and some fibre bundles were forced with a bending radius larger than allowed due to the height of the FPGA power mezzanine. Fig. 12 and 13 show the eye scans for the best and worst cases, respectively.

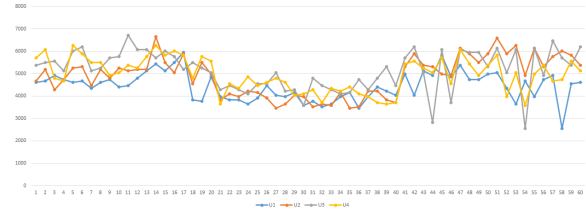


Fig. 10. Eye Diagram open areas for all the 240 MGT inputs measured at 11.2 Gb/s.

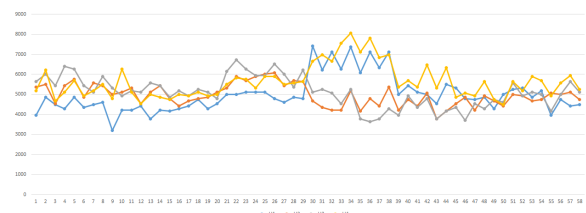


Fig. 11. Eye Diagram open areas for all the PMA loopback MGT inputs measured at 11.2 Gb/s.

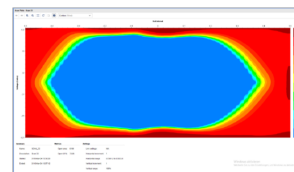


Fig. 12. Eye Diagram with open area around 8000 a.u. (MGT with the largest Eye Diagram open area).

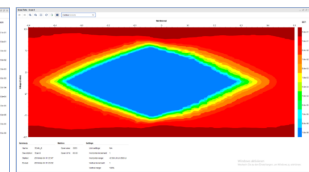


Fig. 13. Eye Diagram with open area around 3000 a.u. (MGT with the smallest Eye Diagram open area).

E. Validation of Parallel I/Os

Additionally to the data sharing using MGTs, the FPGAs have a total 186 parallel I/O lines connected, which can be used for that purpose, with a lower data rate. During the validation of the hardware those lines were tested at 1000 Mb/s (chosen as line rate for measurement purpose - eye scan), with one half of the lines used as transmitters, the other half as receivers.

Eye scans were performed with IDELAY blocks (count mode) on receiver side for each line with a scan resolution 2.1 - 12 ps (512 taps) and the data eyes open are between 84% and 95%. Fig. 14 shows the open area of the eye scans for all lines.

F. FPGA Configuration

While the processor FPGAs are accessible through their JTAG ports at any time, the configuration bitstream required

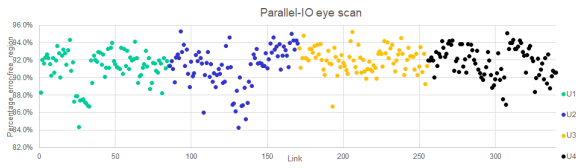


Fig. 14. Open area of parallel IOs eye scans.

at any power-up is meant to be provided by local storage, placed on the jFEX Control Mezzanine. The configuration mode used is Master SPI $\times 4$ with each FPGA having its own configuration flash memory. The configuration should be done in parallel.

Due to the limited number of devices available on the market, the first version jFEX Control Mezzanine is designed using the MT25QU01GBBB8ESF-0AAT, a 1 GB device. In order to use the Multiboot feature of the Virtex Ultrascale+ a larger device is required, therefore the next version jFEX Control Mezzanine is going to be equipped with MT25QL02GCBB8E12-0SIT, a 2 GB device.

1) *FPGA Multiboot*: The FPGA MultiBoot and fall-back features are used on the jFEX board to support updating bitstream images dynamically between LHC runs. The FPGA MultiBoot feature enables switching between images on the fly. In this method, the flash memory holds two images:

- Golden image: previously validated image, which includes only basic slow-control functions implemented;
- Multiboot or user image: working image to be used during normal operation. If an error occurs during loading of the MultiBoot image from the upper address space, the fall-back circuitry triggers the golden image to be loaded.

The FPGA MultiBoot feature allows the IPBus communication to the jFEX to be always maintained, in case the user image is corrupted. Additionally, it allows the image to be re-written via IPBus-SPI bridge.

In order to test the FPGA MultiBoot scheme on the jFEX prototype, a special MCS file containing two different bitstream files has been generated with the Vivado Design Suite and loaded on all memories populated on the Control Mezzanine. A reset/power-up was applied several times and the correct functioning of the four processor FPGAs was verified by checking the configuration status registers.

To validate the fall-back feature, the "user" part of the MCS file previously used, was intentionally corrupted (IDCODE was changed manually) and loaded on the memories. Again, a reset/power-up was applied several times and the FPGA and the correct functioning of the fall-back feature verified.

V. CONCLUSION

The jFEX prototype board was delivered in November 2017. The board has been extensively and successfully tested over the past 12 months.

A pre-production module, with no hardware changes with respect to the prototype, is currently being manufactured and is expected to be delivered before the end of 2018.

The final jFEX modules are expected to be installed in the ATLAS detector before the end of 2019.

ACKNOWLEDGMENT

This research has been supported by German Federal Ministry of Education and Research (BMBF) and by the Cluster of Excellence, Precision Physics, Fundamental Interactions and Structure of Matters (PRISMA) funded by the German Research Council (DFG).

The authors would like to thank all the collaborators of ATLAS L1Calo collaboration for the fruitful discussions and suggestions.

REFERENCES

- [1] ATLAS Collaboration, The ATLAS Detector, JINST 3 (2008) S08003
- [2] "Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System", CERN-LHCC-2013-018
- [3] "Design and test performance of the ATLAS Feature Extractor trigger boards for the Phase-I Upgrade", W. Qian *et al.*, under press in JINST Proceeding section for the TWEPP 2016 conference.
- [4] "The Prototype of Global Calorimetric Hardware Trigger for ATLAS at High Luminosity LHC", H. Chen *et al.* contribution to the IEEE NSS and MIC 2016 conference Strasbourg.
- [5] www.avagotech.com
- [6] "Upgrade of the ATLAS Level-1 Trigger with event topology information", E. Simioni *et al.* J.Phys.: Conf. Ser. (JPCS), Volume 664, 2015
- [7] <https://www.picmg.org/openstandards/advancedtca/>
- [8] www.xilinx.com
- [9] www.ge.com
- [10] <https://www.xilinx.com/products/boards-and-kits/1-58nuel.html>
- [11] "SFF-8431 Specification for Enhanced Small Form Factor Pluggable Module SFP+", <ftp://ftp.seagate.com/sff>
- [12] <https://www.xilinx.com/support/documentation/navigation/design-hubs/dh0008-vivado-power-estimation-and-optimization-hub.html>
- [13] https://www.xilinx.com/support/documentation/data_sheets/ds923-virtex-ultrascale-plus.pdf
- [14] https://www.xilinx.com/support/documentation/sw_manuals/xilinx2016_3/ug973-vivado-release-notes-install-license.pdf