

Overview of ATCA Boards Developments in ATLAS for Phase-1

S. Veneziano, INFN Rome
for the ATLAS Collaboration

(many thanks to the real experts working on Phase-I Upgrade sub-systems,
additional thanks to I.Brawn, R.Spiwoks, L.Levinson)

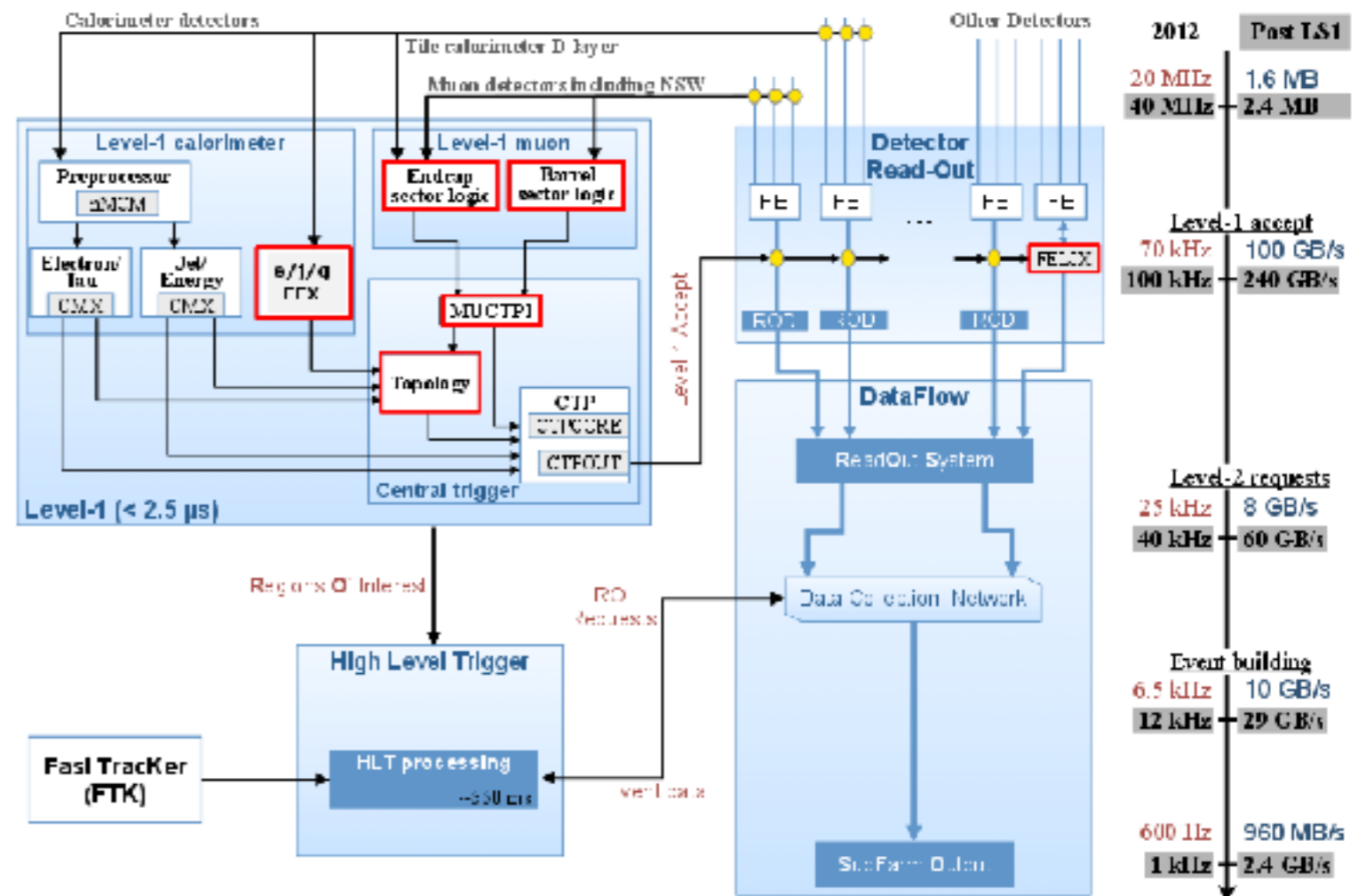
ATLAS Trigger and DAQ after Phase-I upgrade

In Phase-I, ATLAS TDAQ system will face event rates and pile-up levels much higher than the original design values. TDAQ will adapt to this new environment by:

- having a powerful L1Calo using increased granularity to achieve better isolation
- keeping low energy thresholds;
- the Muon Endcap Trigger will be upgraded and suppress fake rate using New Small Wheel detectors;
- New detectors will have high-bandwidth readout compatible with Phase-II (FELIX);

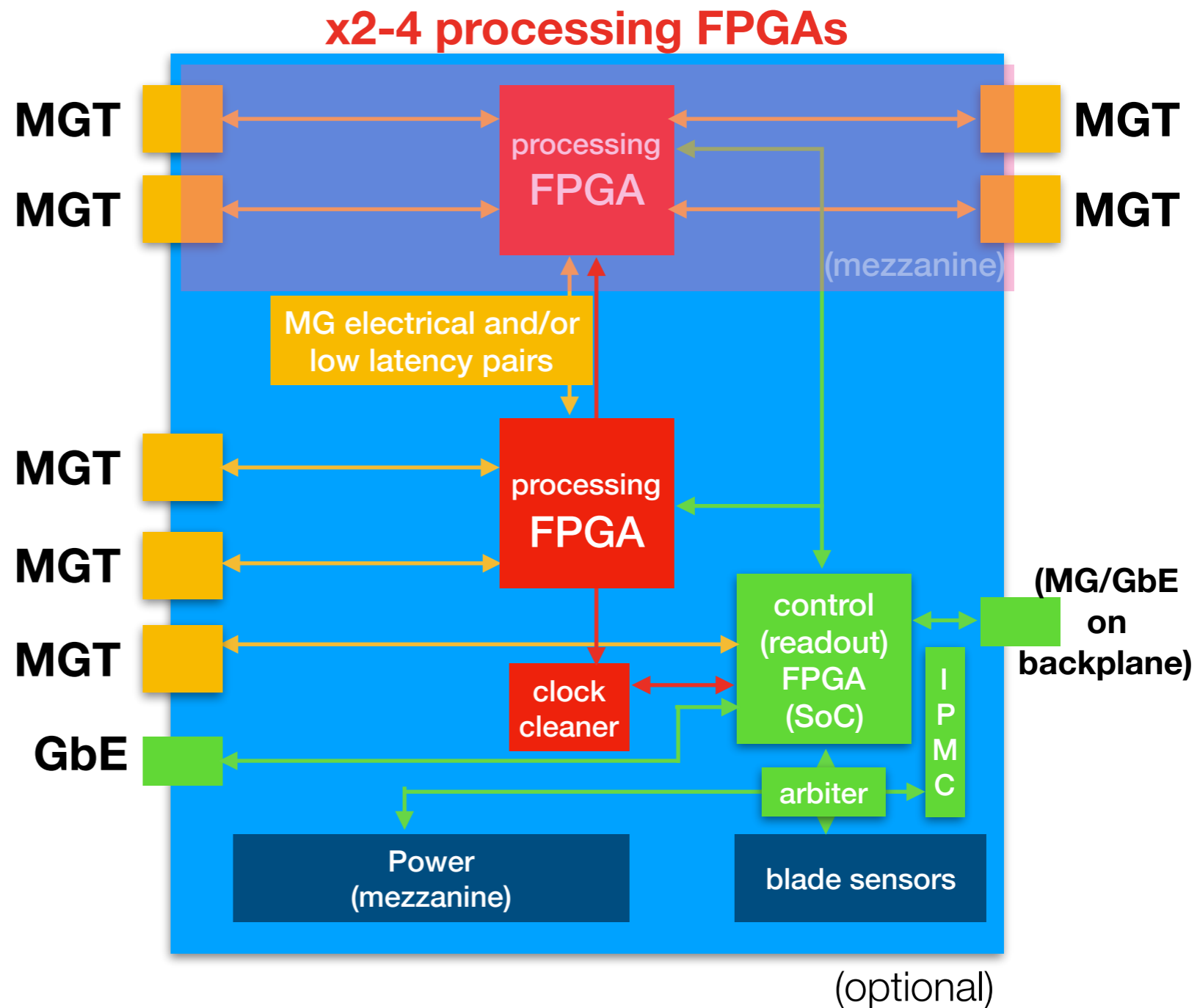
It will be achieved with:

- an upgraded L1Trigger: a real-time, low latency path using:
 - Multi-Gbps (6.4-12.8 Gbps) optical IOs
 - Algorithms implemented in large FPGAs
 - ATCA (VME) boards hosting multiple interconnected FPGAs using Multi-Gbps links.
- an upgraded Dataflow:
 - Custom boards hosted on commodity PCs.



Phase-I ATCA board model

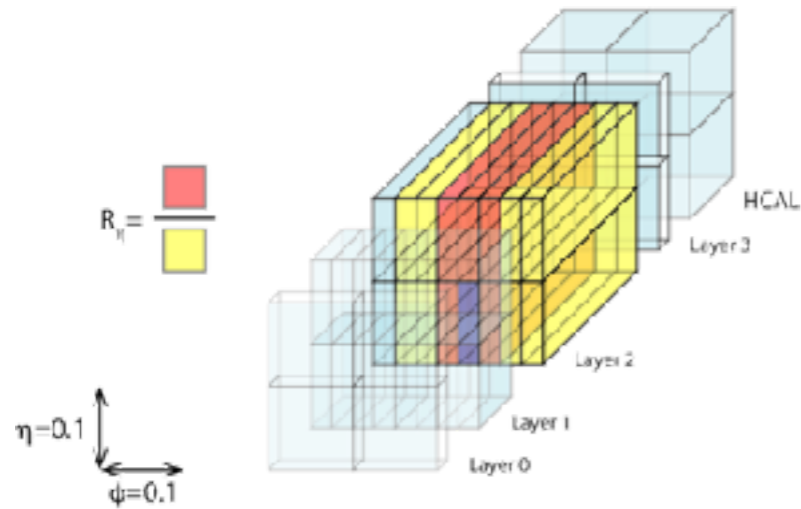
- Multi-Gbps (6.4-12.8 Gbps) optical IOs, usually placed in-board (short track lengths) for:
 - Timing (TTC) path
 - Real-time data path
 - Readout data path
- large IO FPGAs and different boards hosting custom inter-FPGA connections topology is mandatory to share data between processing elements (large environment used by trigger algorithms) and minimise trigger latency (some data duplication at source).
- ATCA blades.
 - sometime non-standard form-factor (Zone-3 projection into rear area to make room for fibre routing).
 - backplane sometimes used for TTC and readout, GbE
 - configuration and control FPGA through GbE
 - optional System-On-Chip for configuration, control and monitoring



Algorithms: example EM

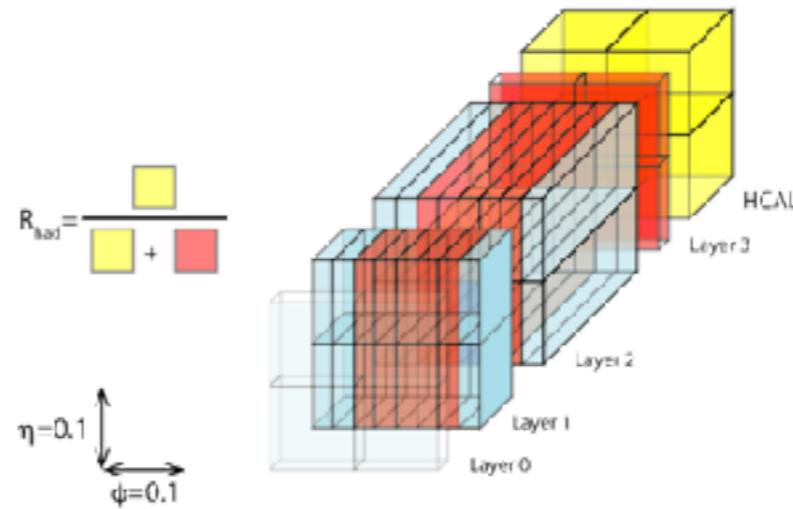
EM shower width:

$$R_{\eta} > R_{\eta}^{\text{Thresh}}$$



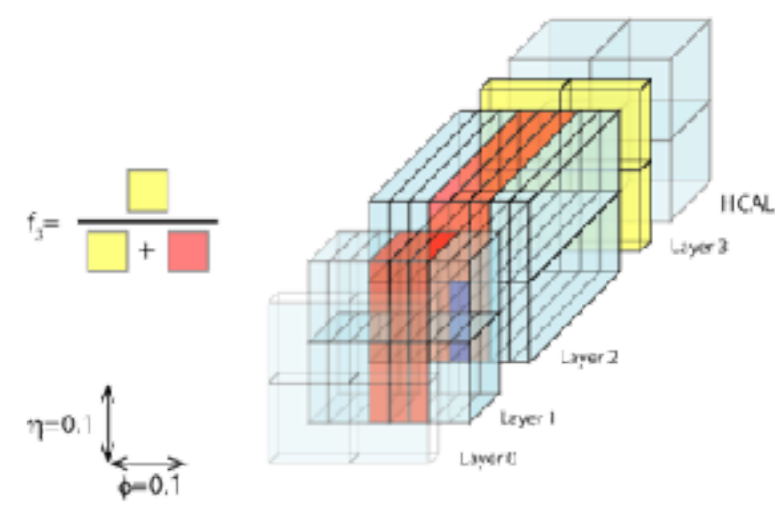
ECAL vs HCAL energy deposits:

$$R_{\text{had}} < R_{\text{had}}^{\text{Thresh}}$$

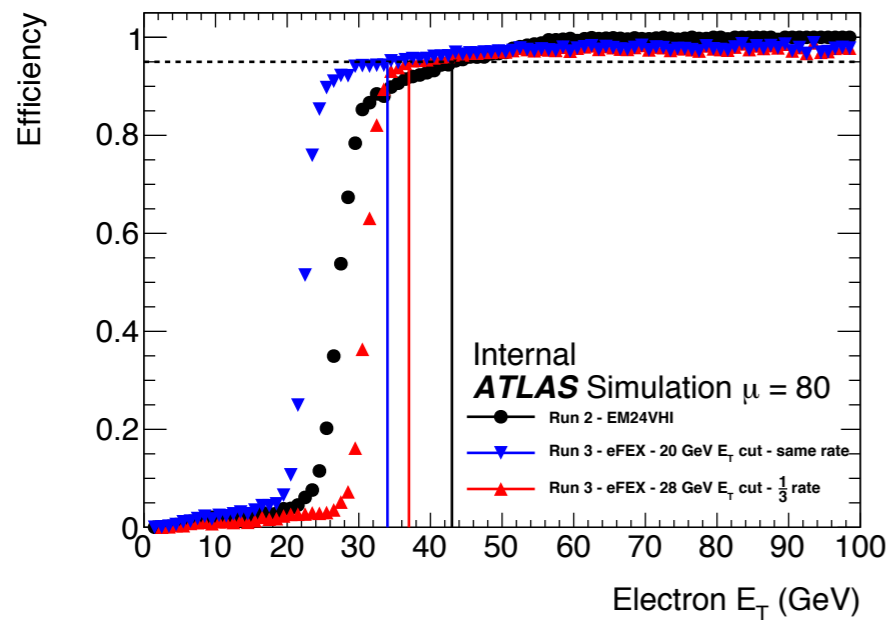


Shower depth in ECAL:

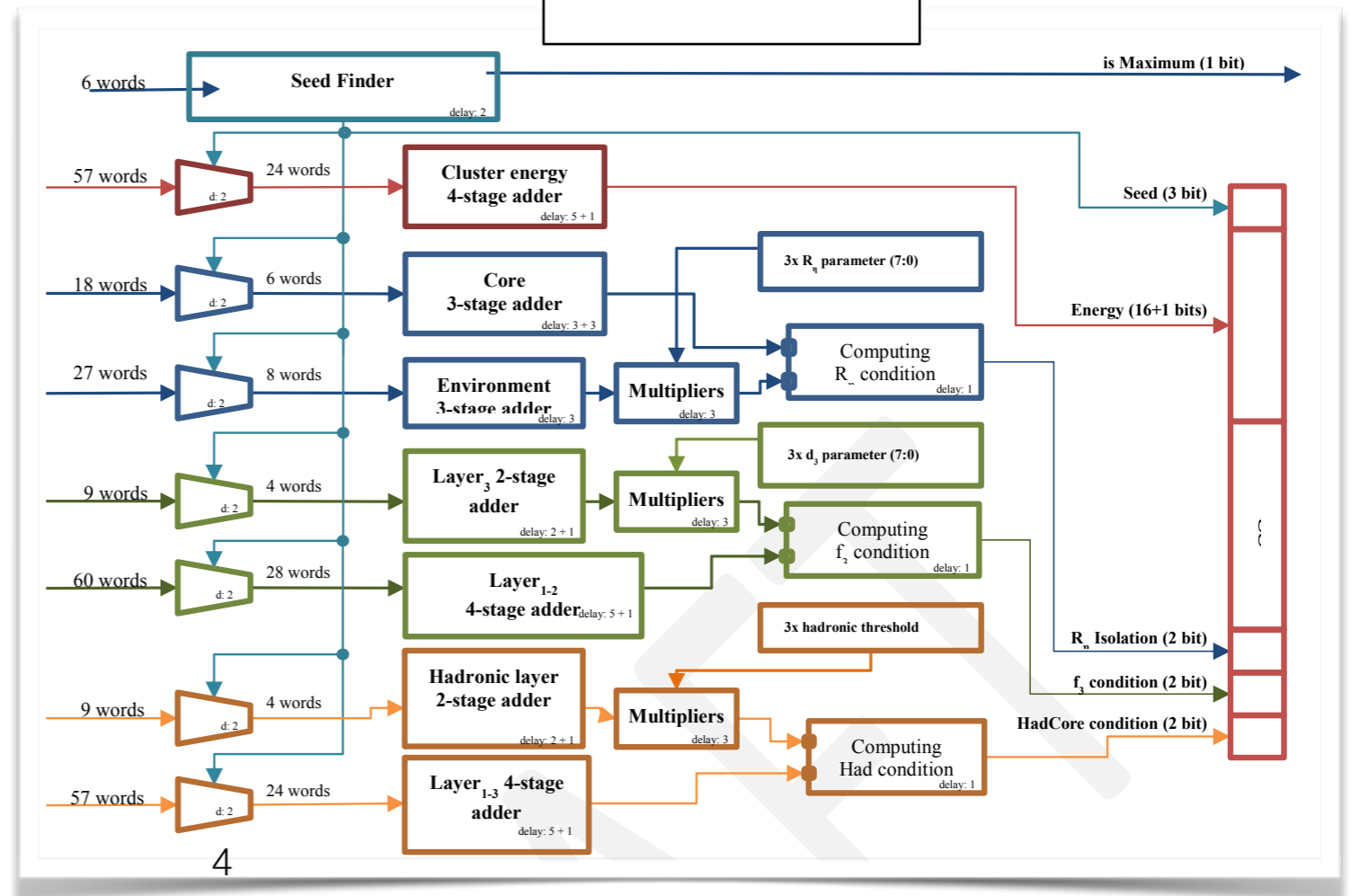
$$f_3 < f_3^{\text{Thresh}}$$



e/gamma efficiency Turn-on vs Run2



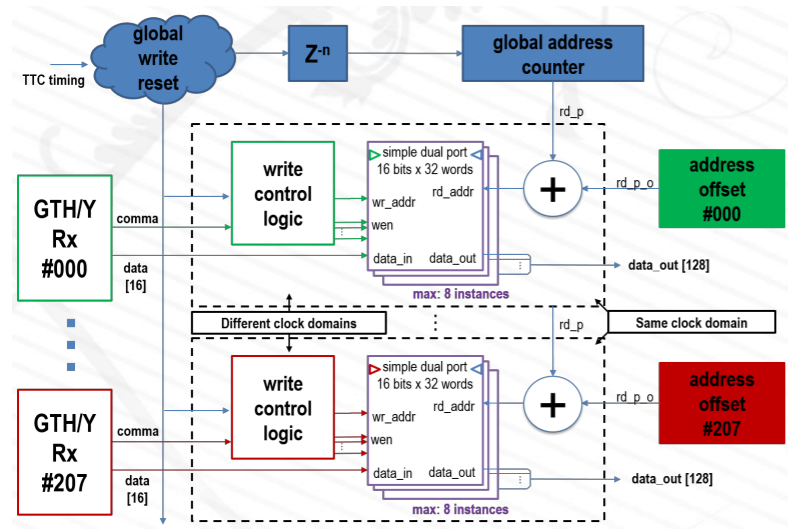
The eFEX EM module



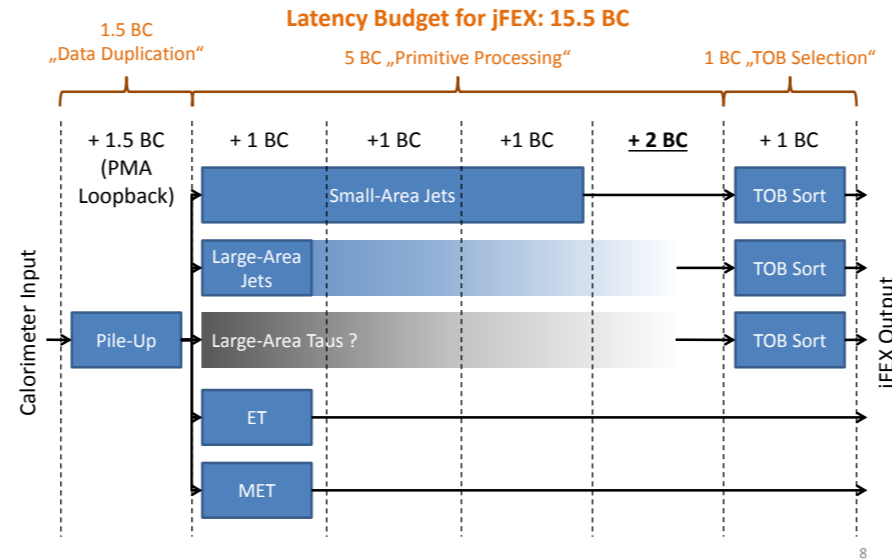
Processing

- strict requirements on realtime path, synchronisation to a common LHC clock domain and fixed latency.
- trigger data readout based on FELIX (Phase-2 ready).

MUCTPI sync+align block



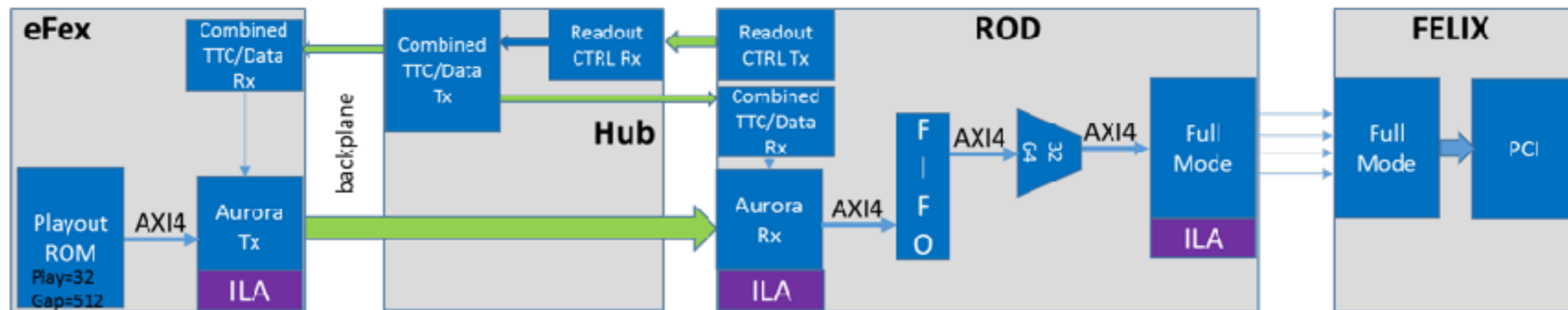
jFEX latency budget



eFEX resource utilisation

Site Type	used	Available	Utilization %
Slice LUTs	94550	346400	27.30
Slice Registers	143469	692800	20.71
F7 Muxes	4385	216600	2.02
F8 Muxes	145	108300	0.13
IBUFDS_GTE2	16	40	40.00
Block RAM	382	1180	32.37
DSP	144	2880	5.00
GTHE2_CHANNEL	64	80	80.00

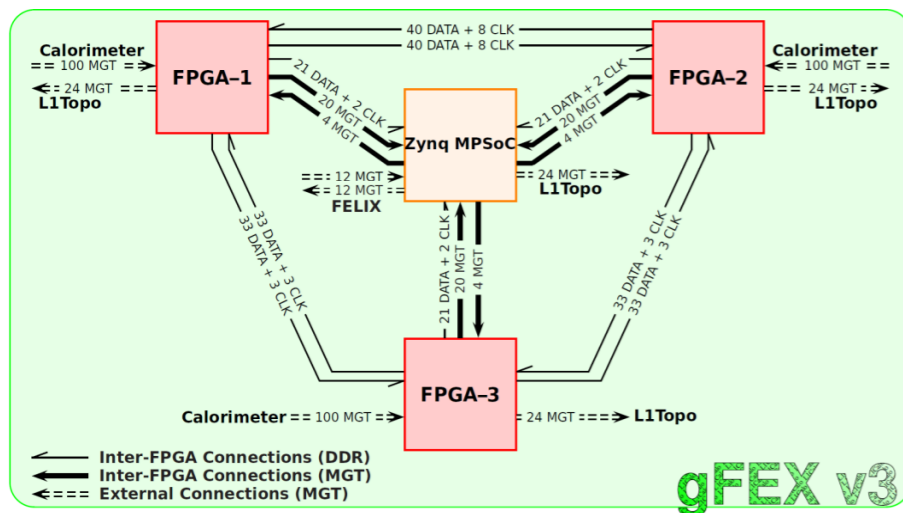
eFEX, jFEX and L1Topo readout through ATCA backplane



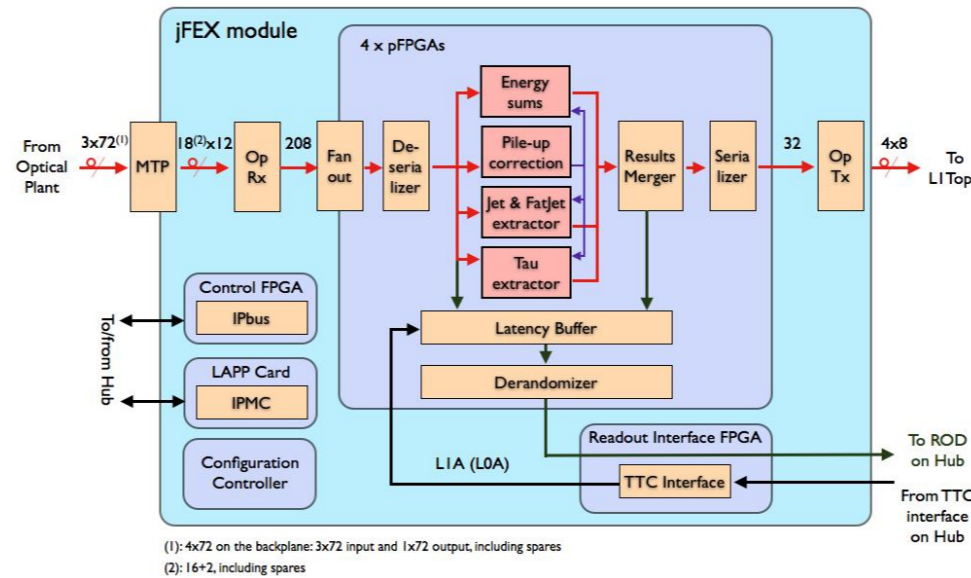
Inter-FPGA data and control paths

- use of MGT, LVDS (lowest latency), use of higher protocols (IPBus or AXI) for configuration and control.

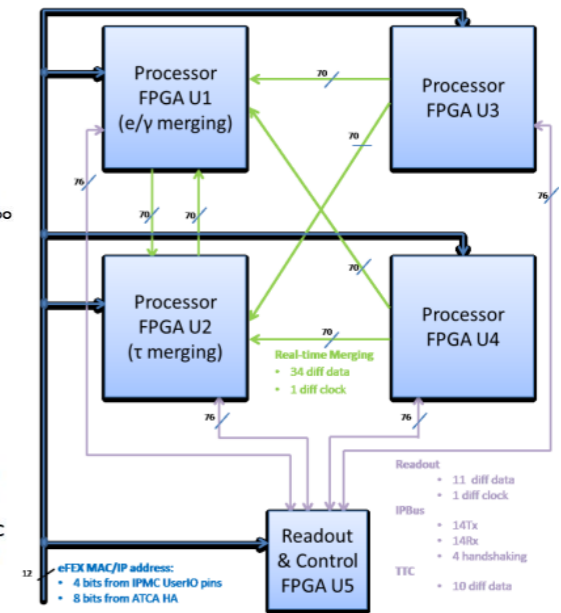
gFEX doesn't share input data, and uses LVDS to share partial results



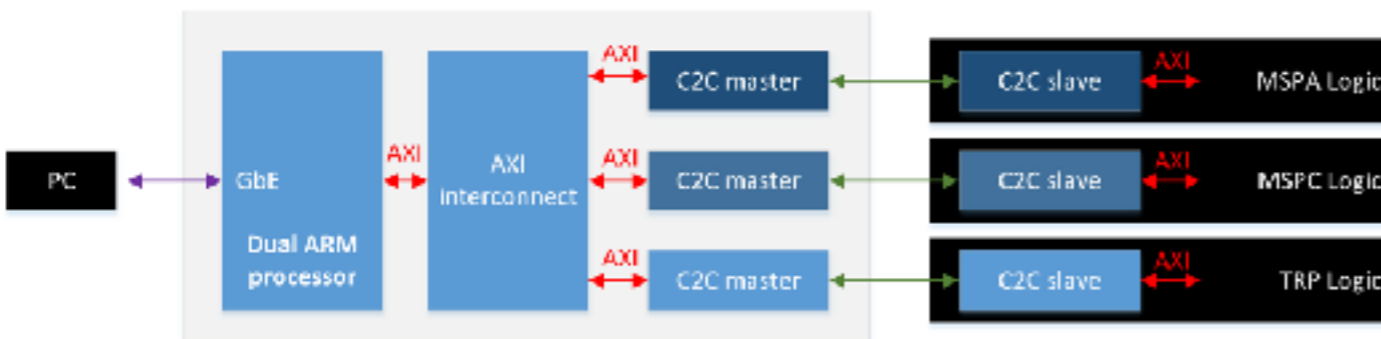
jFEX PMA loopback to share input data, and doesn't share partial results



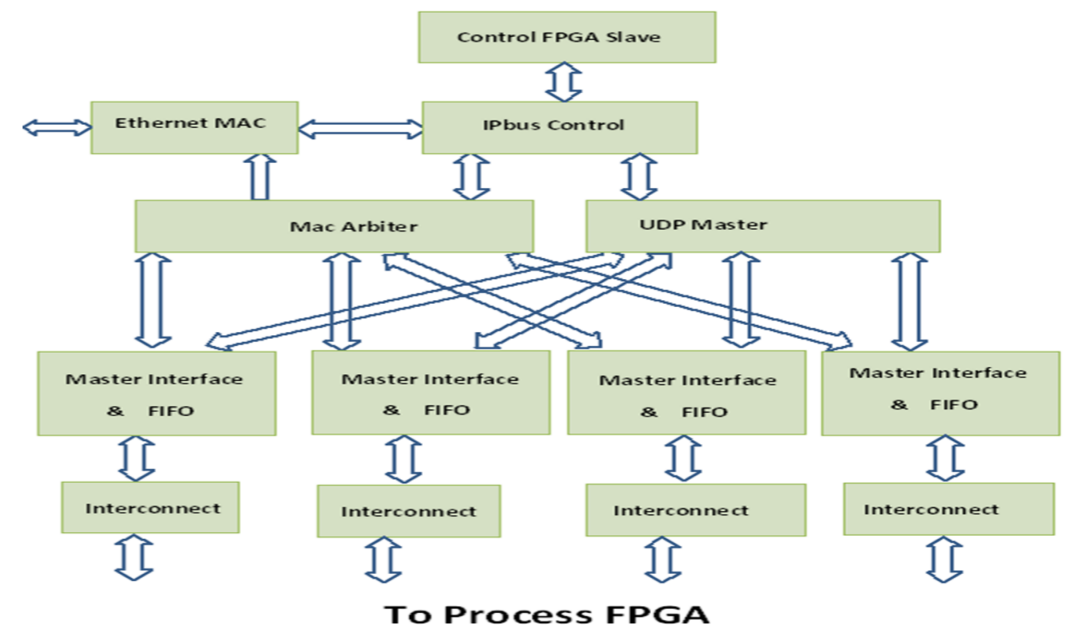
eFEX use discrete fan-out buffers to share multi-Gb/s input data, and LVDS to share partial results



MUCTPI AXI C2C inter-FPGA communication

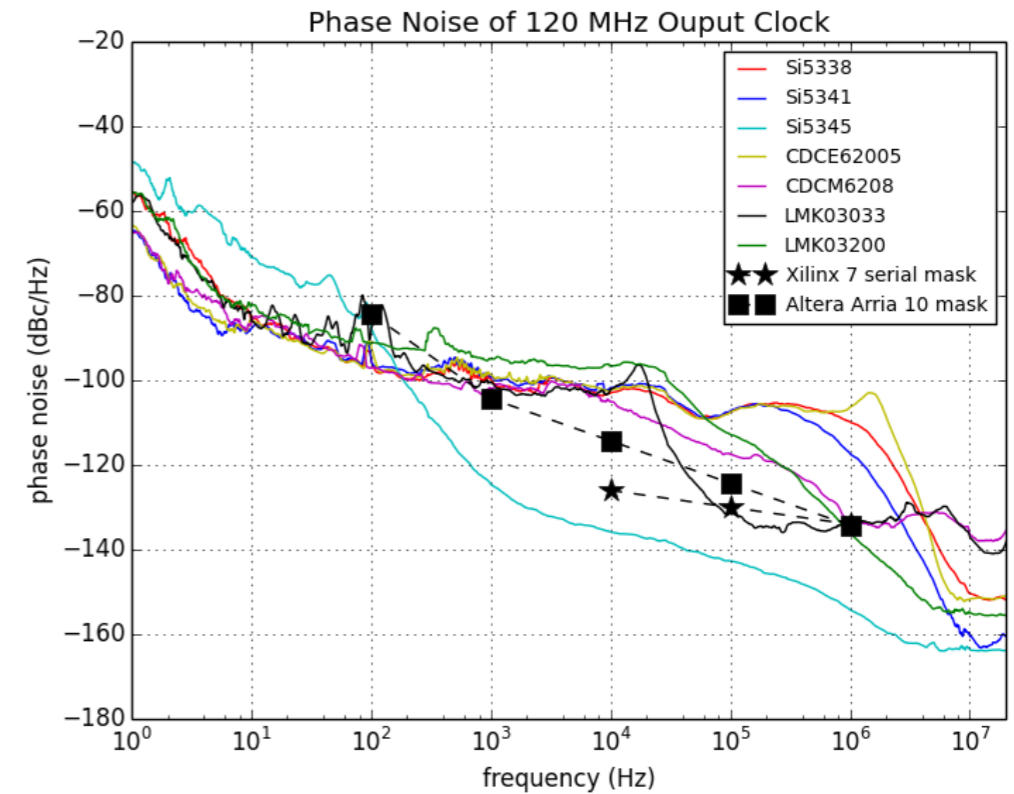


eFEX IPBus inter-FPGA communication



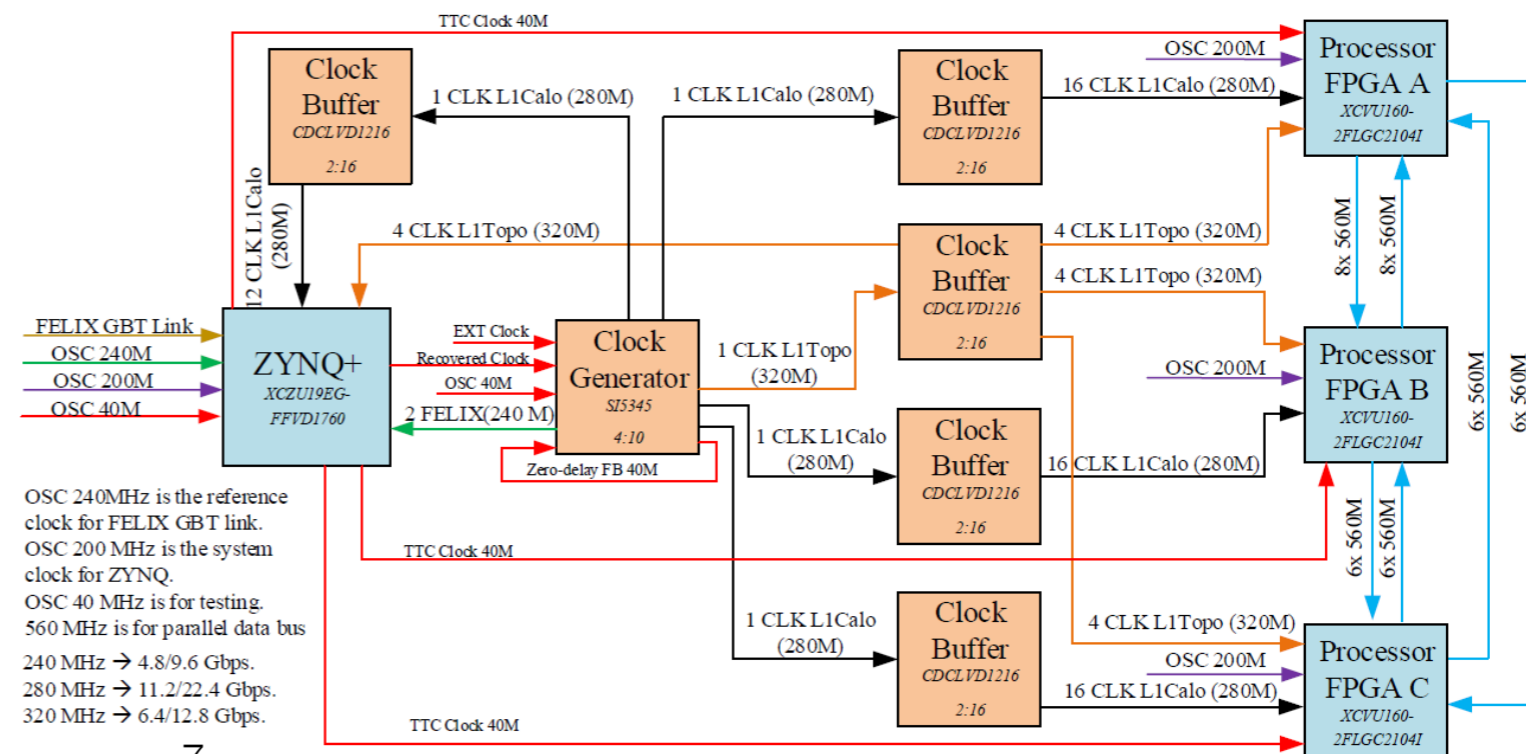
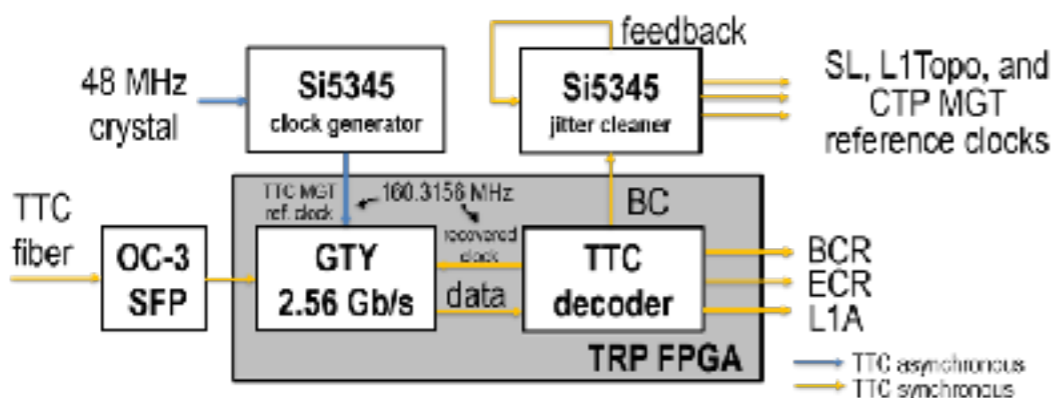
Clock quality and distribution

- fixed-latency realtime-path requires common LHC clock:
 - GBT link from TTC distribution (FELIX)
 - one FPGA with local oscillator(s) recover TTC clock phase.
- A low-jitter clock is particularly important for the MGTs
- use of jitter cleaner is a common solution adopted in all ATCA designs:
 - jitter cleaner chip (Si5345, zero-delay feedback mode) filters 40 MHz clock before distribution to other FPGAs.
- clock quality proven in multi-gigabit transmission tests (see next slide).



gFEX GBT RX clock distribution

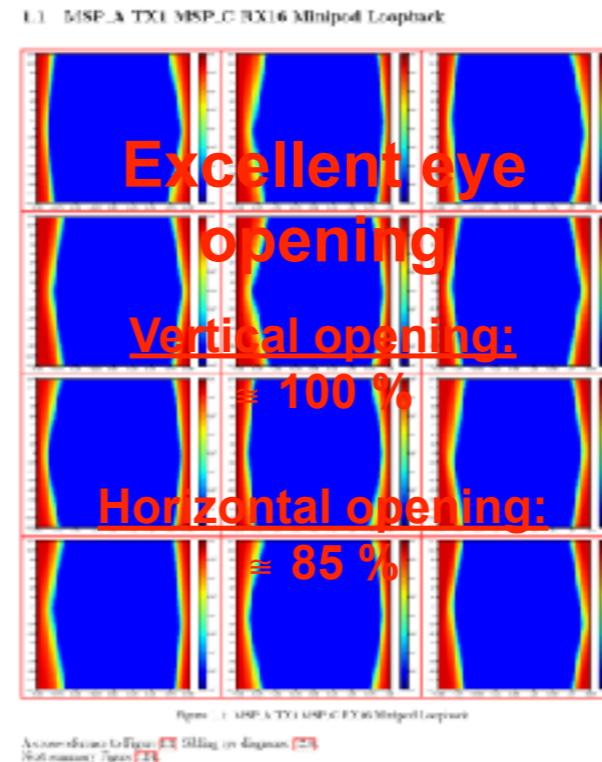
MUCTPI TTC RX clock distribution



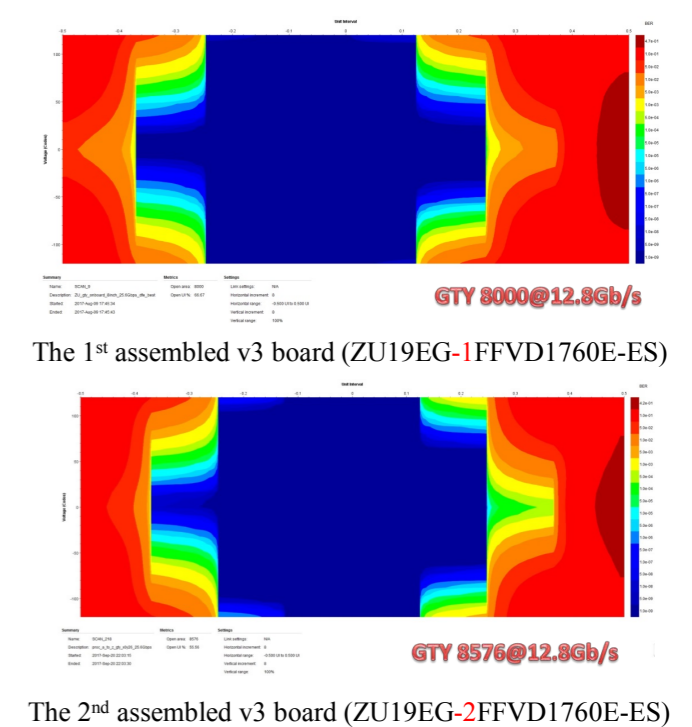
Multi-Gigabit Transceivers

- MGT optical used to feed detector data, electrical for inter-FPGA communication, along realtime path and also used in readout paths.
 - Speeds vary between 4.8 and 25.6 Gb/s.
 - BER < 10⁻¹⁶ (ROD ATCA mezzanine, 6.4 Gb/s through ATCA backplane)
- Extensive tests have proved the capability of these boards in handling hundreds of links running concurrently.
 - very few traces need to be re-drawn due to PCB layout issues.
 - factors in the success of the PCBs: the use of signal integrity simulation in the design of the boards, and the use of low Dk materials.

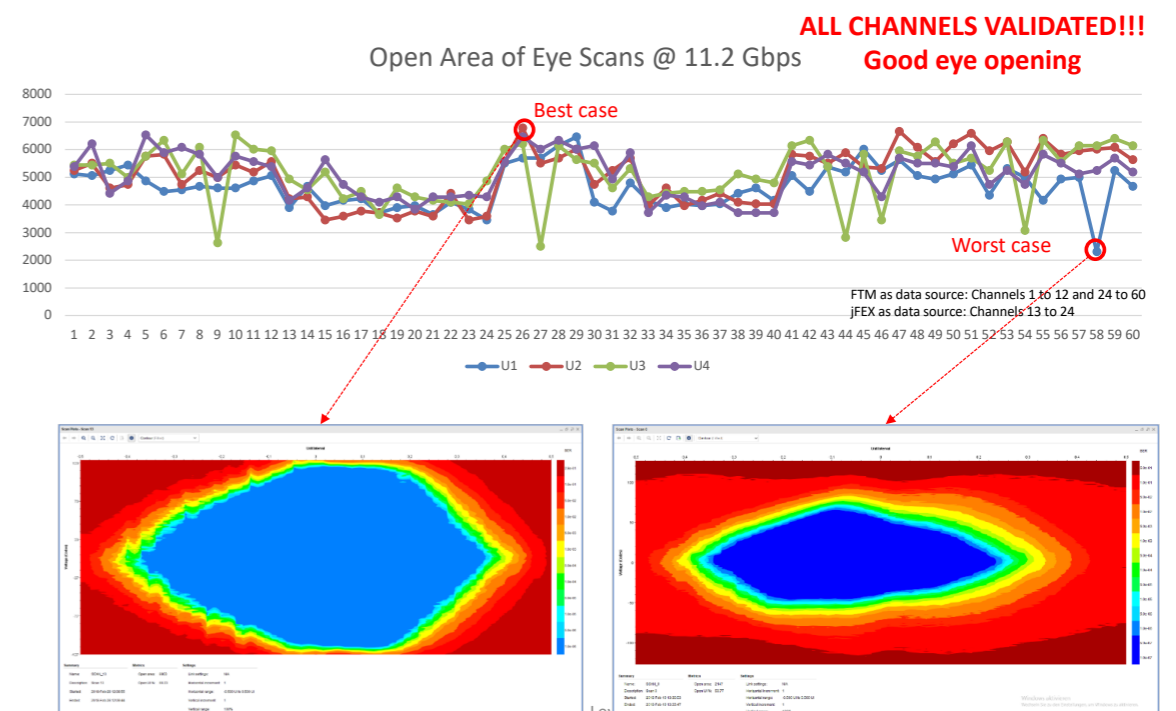
6.4 Gb/s MUCTPI Optical



25.6 Gb/s gFEX Electrical



11.2 Gb/s jFEX Optical



Processor Component	Input links per board	Output links per board
eFEX	144	48
jFEX	240	48
gFEX	312	108
L1Topo	236	48
MUCTPI	208	65
NSW TP	146	28

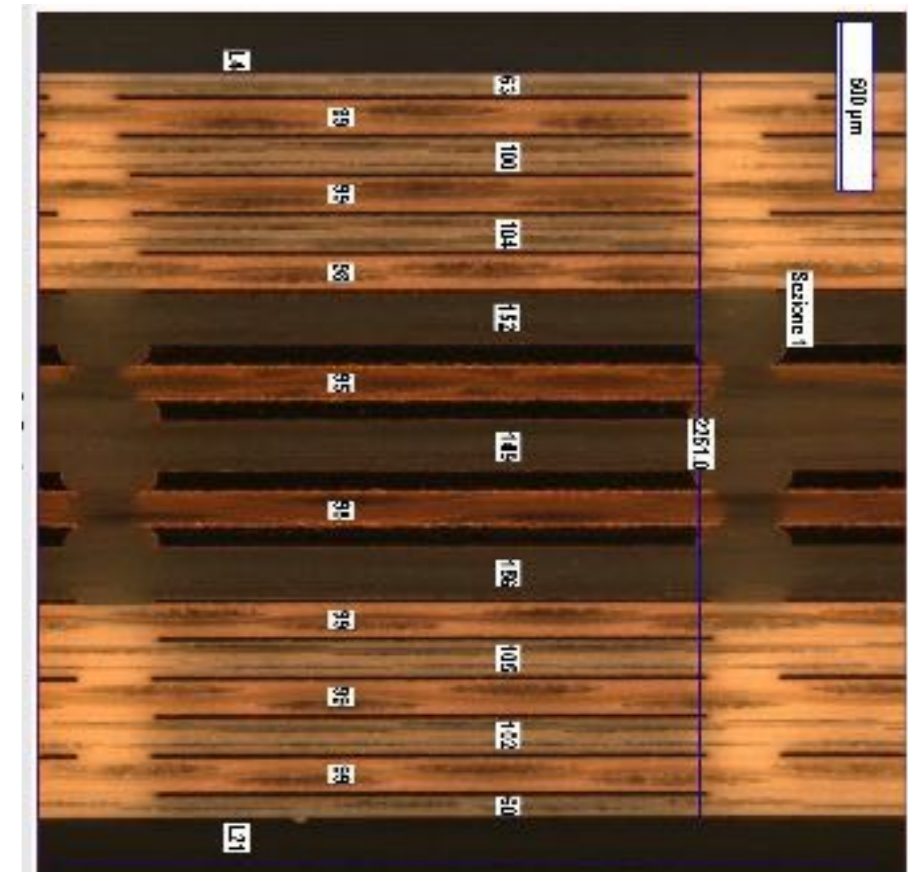
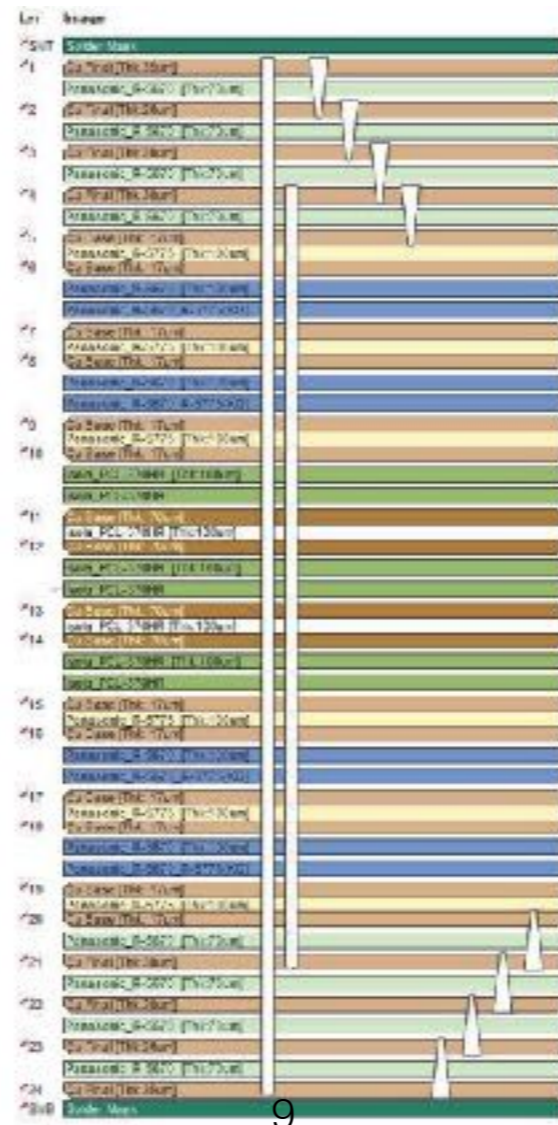
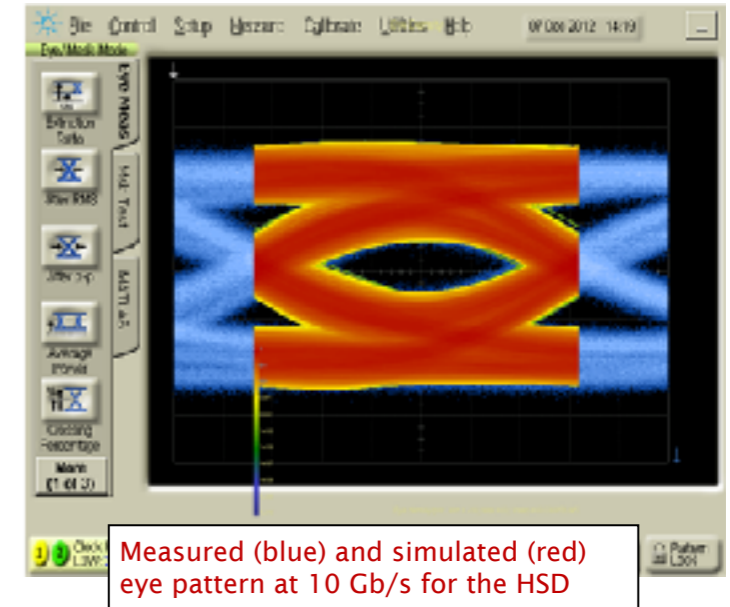
PCB material

- Many PCB manufacturing issues found with early prototypes due to manufacturers having limited experience with new materials (e.g Megtron-6, Isola FR408HR):

- low initial yield, many months delay on early prototypes.

- problems now solved through:

- dialogue with manufacturers
- signal integrity simulations
- early manufacture of PCB to test stack-up
- TDR/TRT test of bare boards
- Dual source for manufacturing.



JFEX stack-up

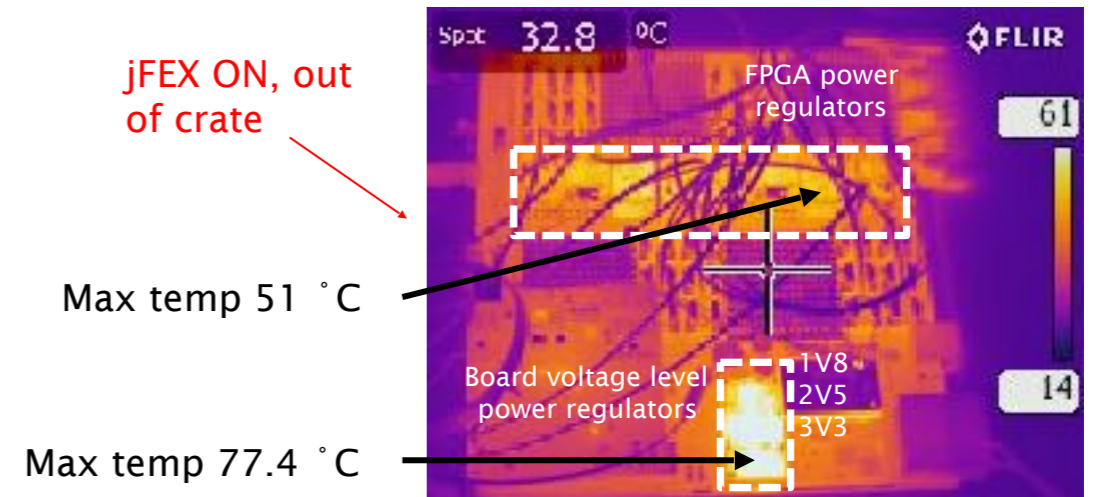
Power

- Blade Power Consumption has been a major worry, high-power 300W(gFEX)-400W(eFEX, jFEX) range expected, primarily because of the cooling it requires, rather than our ability to bring the power to the board.
- Extensive measurements on prototypes confirm that power and temperature are within estimations.
- Production versions are improving the design of the power planes to reduce the voltage drop. The layout of the eFEX for example also allows bus bars to be used to supply power to the far corners of the board, thus further reducing the drop if needed.

gFEX estimation

Power Consumption Estimation

	Number	Power Dissipation (W)	Total (W)
Processor FPGA	3	75	225
ZYNQ	1	50	50
MiniPODs RX	26	1.7	45
MiniPODs TX	9	2.0	18
Misc	1	25	25
gFEX	1		363
Efficiency	85%		427



❑ FPGA power is estimated by the XPE of Xilinx when running IBERT.

gFEX measurement, 342 W (363.8 @ 94% efficiency)

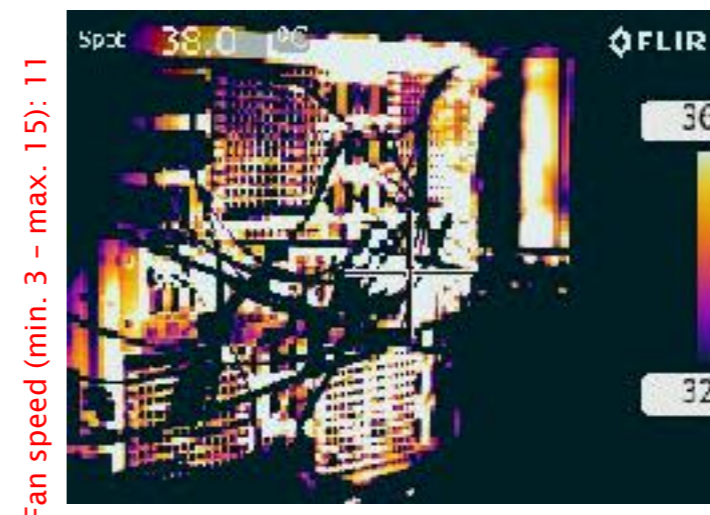
FPGA power consumption test with different firmware configurations

	FPGA	Link speed	MGT RX	MGT TX	VCCINT 0.95V	MGTAVCC 1.0V	MGTAVTT 1.2V	Totalpower
GT Firmware	VU160	12.8 Gb/s	48 GTH 52 GTY	12 GTH 12 GTY	6.6A 6.3W	14 A 14W	18.1A 21.7W	42 W
GT Firmware	VU160	12.8 Gb/s	48 GTH 52 GTY	12 GTH 32 GTY	6.6A 6.3W	14 A 14W	20.5A 24.6W	44.9W
IBERT	VU160	12.8 Gb/s	52 GTH 52 GTY	52 GTH 52 GTY	13.9 A 13.2W	17.8 A 17.8W	27.8 A 33.36W	64.4W

	FPGA	Link speed	MGT RX	MGT TX	VCCINT 0.85V	MGTAVCC 0.9V	MGTAVTT 1.2V	DDR4VDDQ 1.2V	Total power
IBERT	ZU19	12.8 Gb/s	44 GTH 28 GTY	44 GTH 28 GTY	13.5 A 11.5W	7A 6.3W	10.5 A 12.6W	2.4A 2.9W	33.3W

- ❑ 20W/pFPGA less for GT firmware.
- ❑ Nick's firmware is about 2.2W/pFPGA, the total power should be around 310W.
- ❑ In addition, the VU9P is about 10W/pFPGA less than VU160. So the total power consumption for the production board should be around 280W.

JFEX in shelf, ALL MGTs enabled



XADC	
FPGA	Temp. (°C)
U1	83.20
U2	69.30
U3	54.30
U4	62.30

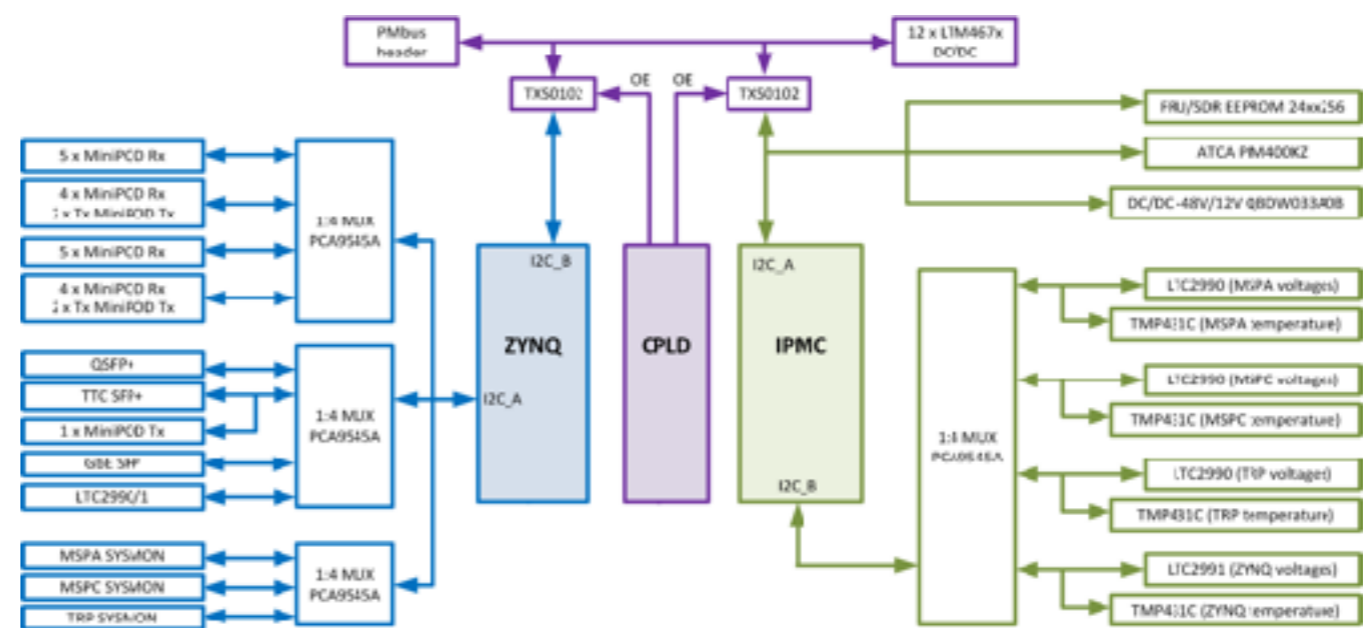
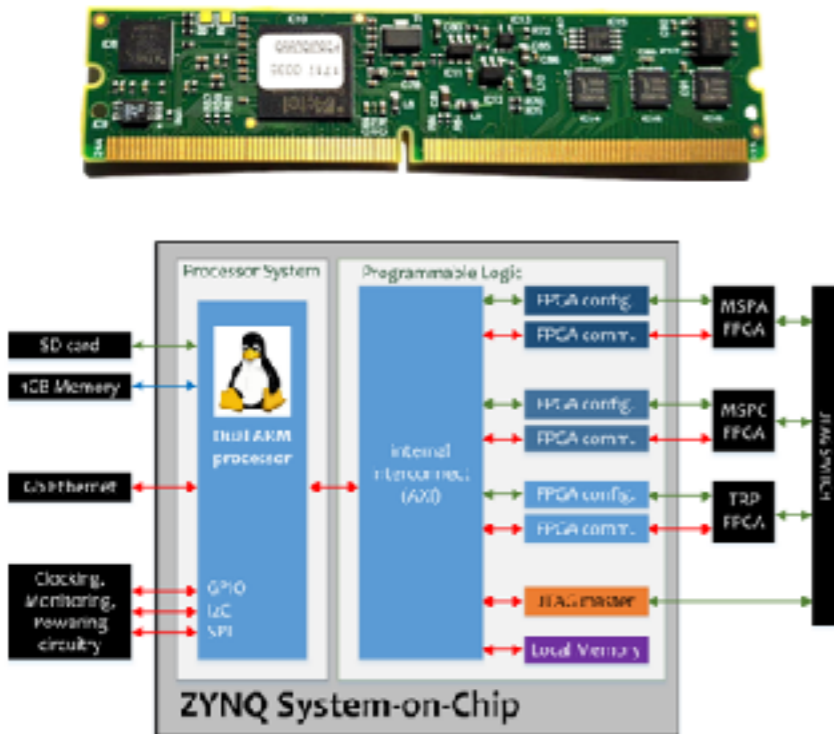
Shelf and blade Control and monitoring

- IPMC (CERN or LAPP) mainly used for basic blade control and monitoring only.
- In addition and complementary to the IPMC, some designs use an SoC to control and monitor the hardware of an ATCA blade
 - see MUCTPI and gFEX as examples below.

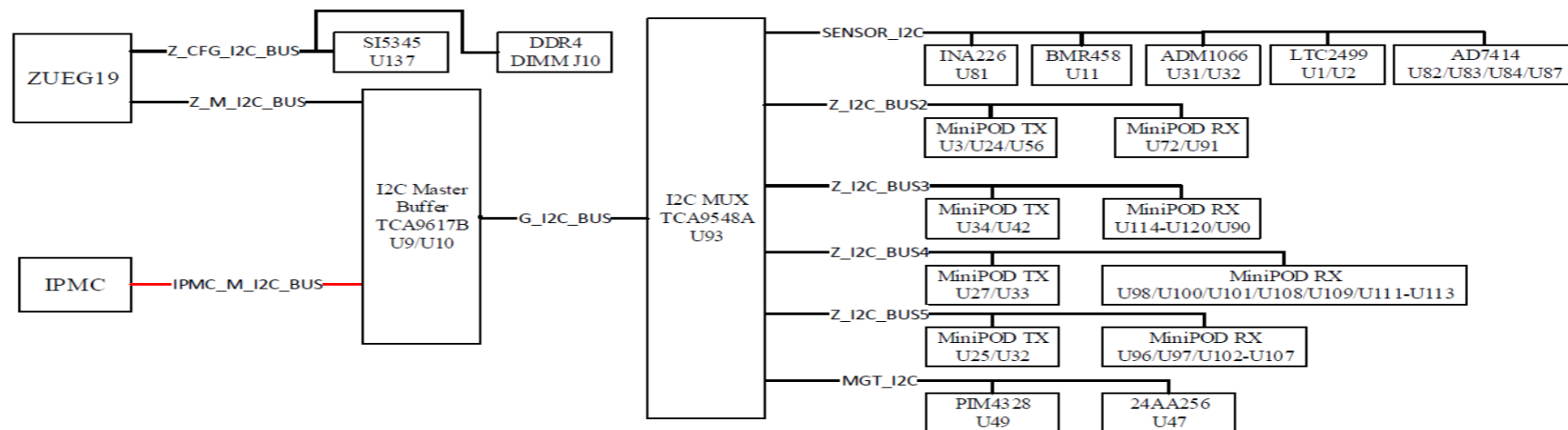
please follow R. Spiwoks presentation coming later today !

MUCTPI

CERN IPMC mezzanine



gFEX



FPGA technologies

- Choice mainly driven by number of Multi-Gigabit Transceivers available, preference for latest generation of Xilinx devices (Ultrascale+).
- The choice of FPGA represents the leading generation of FPGAs when the designs were frozen

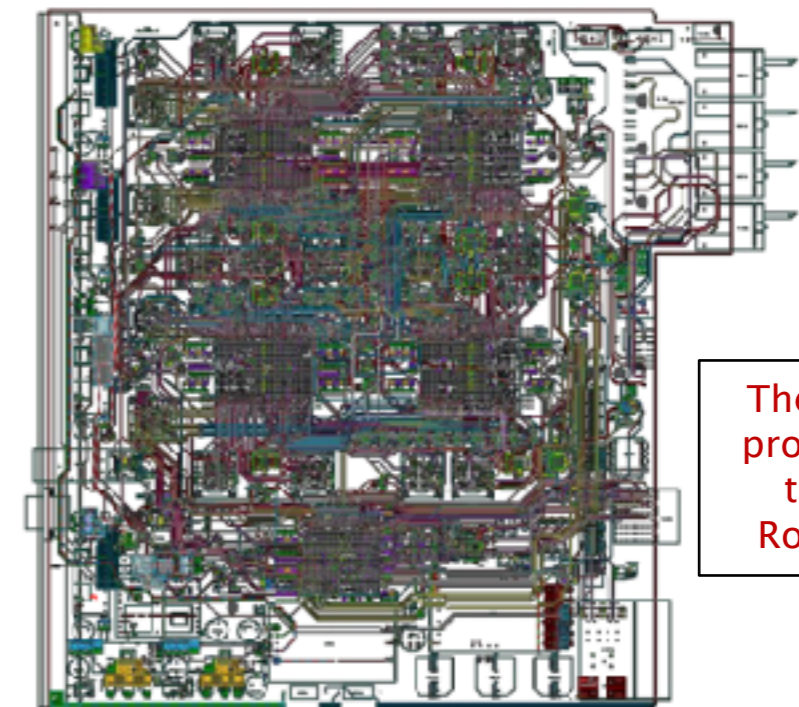
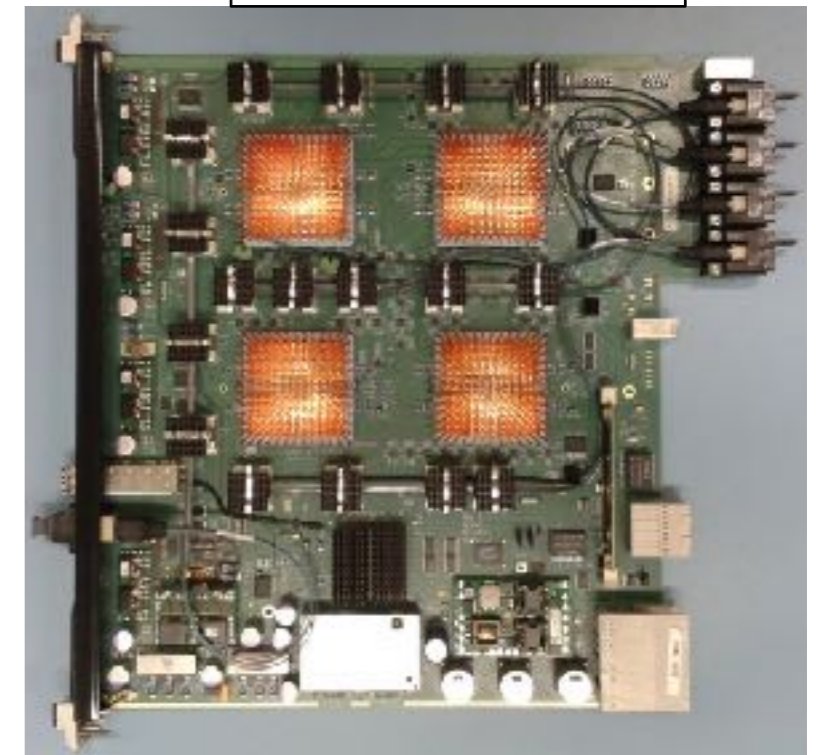
Blade	Processing FPGAs	Control FPGA
eFEX	4 x Xilinx V7550T	V7330T
jFEX	4 x VU9P	Artix7 mezzanine
gFEX	4 x VU160 -> VU9P	ZYNQ7000 -> ZYNQ+
MUCTPI	2 x VU160 -> VU9P	KU095 ZYNQ7000 (-> ZYNQ+)
NSW	4 x Virtex790 (on two mezzanines) 2 x V6 -> KU	ZYNQ7000

The Electron Feature Extractor
Identifies isolated e/γ & τ candidates
Higher granularity & segmentation in depth c.f. current system
24 modules in system

eFEX Hardware

- The eFEX module
 - 4 Processing FPGAs (XC7VX550T)
 - 424 links at ≤ 11.2 Gb/s
- 3 full prototypes assembled & fully tested
 - High-speed links: BER $< 10^{-14}$ @ 11.2 Gb/s
 - Power & temp. with 1st approximation of full firmware load
 - 280 W / module, 67 °C
 - IPBus: no errors over 10^7 block transfers
- **Passed Final Design Review, December'17**
 - Hardware + Firmware
 - Approved changes:
 - Increase power to FPGAs
 - Accommodate outputs to Phase-II Global Trigger Processor
- Currently Re-routing PCB
- Next Milestones
 - Receive Pre-Production modules July'18
 - PRR Dec'18

The eFEX Prototype



The eFEX prototype track Routing

eFEX Firmware

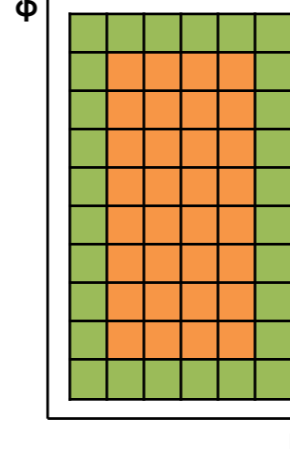
- Real-time path implemented
 - MGT Rx & Tx
 - Synchronisation
 - e/γ algorithm
 - TOB sorting
- Control interface (IPBus) implemented
- Readout path
 - Trigger OBjects readout path implemented
 - In development
 - Bulk (input) data path
 - Interface to L1Calo ROD
- Utilisation
 - ~50% logic, 70% RAM
 - (with Phase-II readout parameters)
 - Latency 11.1 BC
 - (c.f. 11.5 BC budget)

■ Environment
■ Core area

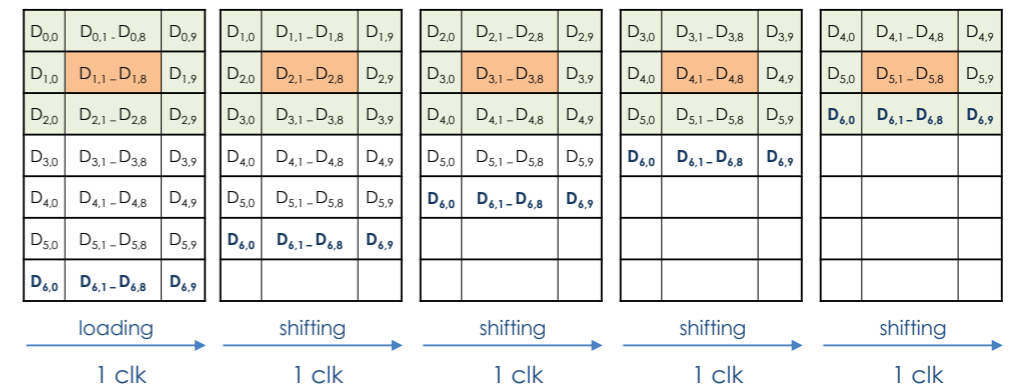
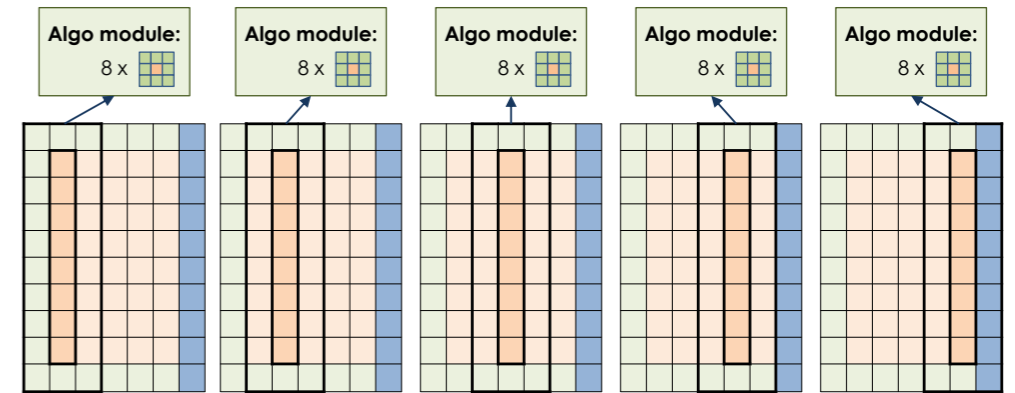
Algorithm Core



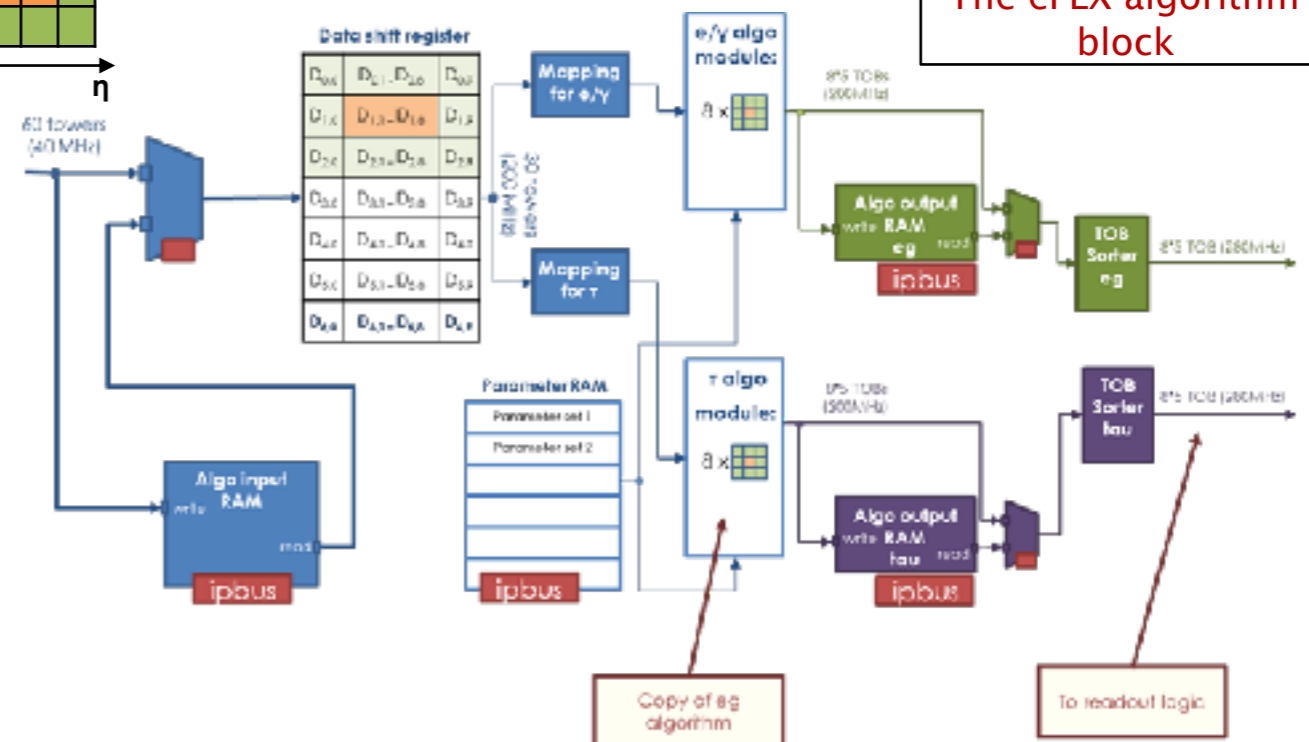
eFEX FPGA



e/γ algo module, Multiplexed at 200 MHz



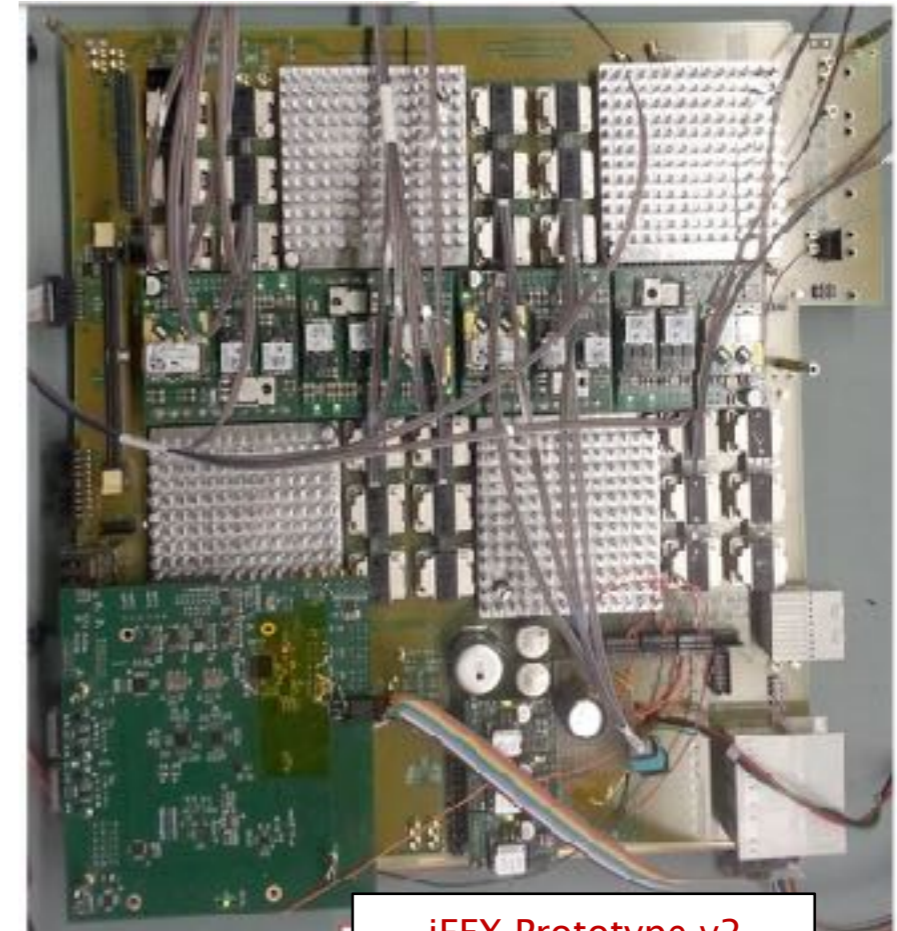
The eFEX algorithm block



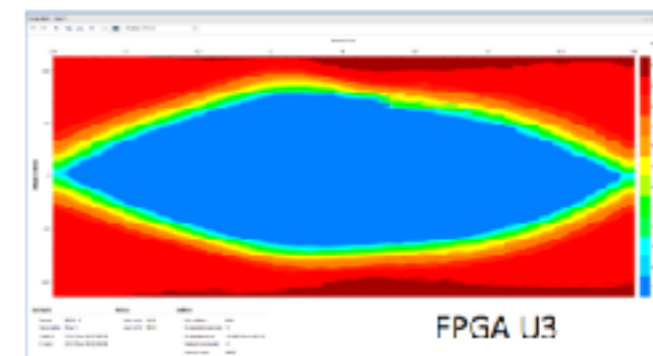
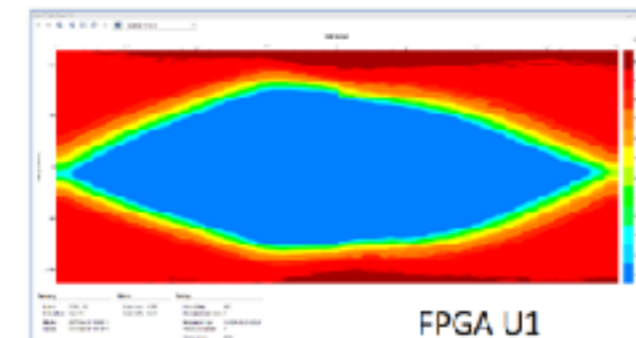
The Jet Feature Extractor Identifies jets (large τ), ΣET , ET_{miss} 6 modules in system

jFEX

- 1st prototype received Dec'16
 - PCB assembly issues due to board thickness/complexity
 - Single Processor FPGA fitted &
 - Tested successfully, but need to test all FPGAs concurrently required...
- Final Prototype received Nov'17
 - New PCB stack-up
 - All 4 Processor FPGAs (XU+) fitted
- Tests completed
 - Control mezzanine v2.1 almost fully tested
 - Power mezzanine: ripple within spec: <10mV
 - All MGT IO (optical & electrical) tested individually
 - Input MGT, output MGT, @ 12.8 Gb/s, BER < 10^{-15} (error free)
 - Inter-FPGA MGT links @ 12.8 Gb/s, BER < 10^{-15} (error free)
 - PMA loop-back
- Current tests
 - Completing soak test of all IO concurrently (signal integrity & power)
- Milestones
 - Final Design Review (hardware + firmware) May'18



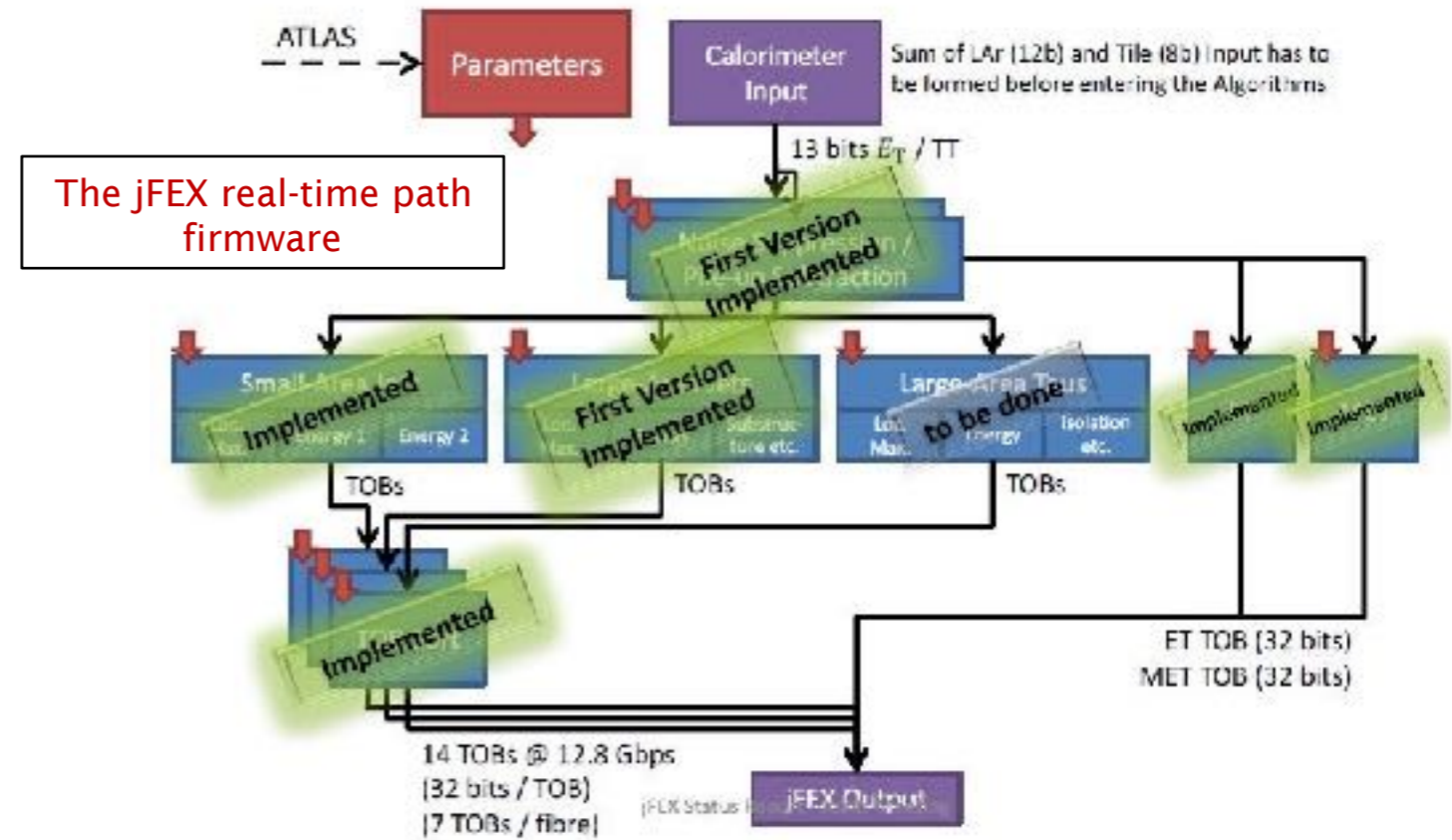
jFEX Prototype v2



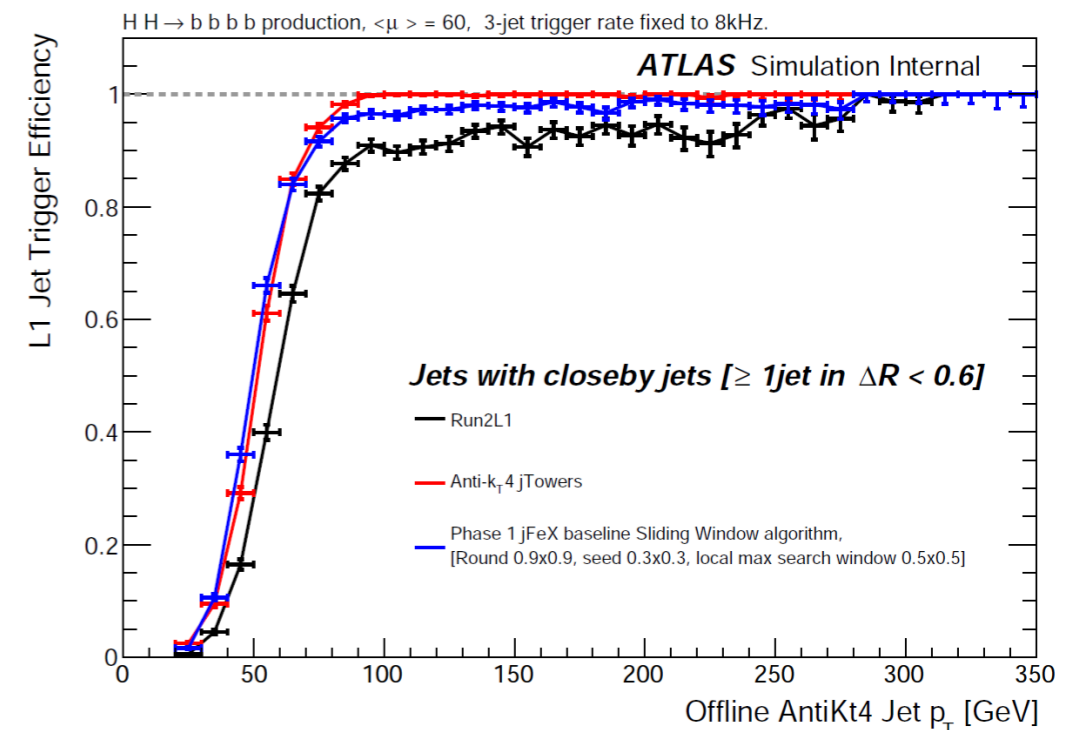
Eye patterns
measured on jFEX
prototype 2 @ 12.8
Gb/s

jFEX Firmware

- Real Time Path:
 - Individual blocks done
 - Integration ongoing
- Control path:
 - Implemented on IPBus,
 - I2C & IPMC underway
 - TTC outstanding
- Readout:
 - In development
- Algorithms implemented on dev board:
 - Small-area jets (0.9 x 0.9)
 - Gaussian weighting
 - Large-area jets (1.7 x 1.7)
 - Sub-structure to be done
 - ET & ETmiss
 - Average pile-up calculation and suppression
 - TOB generation, sorting and counting
- Latency: 7.5 BC (Algo)+ 6 BC (RTDP)
= 13.5 BC (c.f. 15.5 BC budget)
- Utilization: 16.6% LUT



jet turn-on with close-by jets



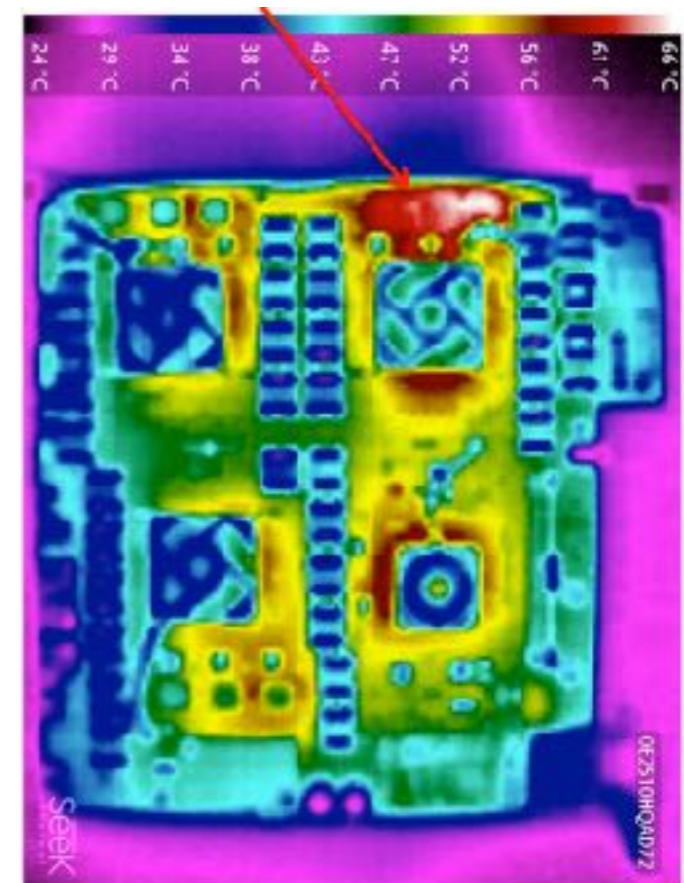
gFEX Hardware

Global Feature Extractor
Large radius jets, pile-up estimation and subtraction, ETmiss, centrality.
Processes entire calorimeter on a single module

- 3 generations of prototype with FPGA upgrades
 - v1 (XV7, 2015) → v3 (XU/U+, Jul'18)
 - Tests completed
 - BER < 10^{-15} , 310 W, < 66°C...
- Passed PRR, Dec'17
 - Based on prototypes v3a (3 x VU) & v3b (1 x VU+)
 - Remaining questions from FDR addressed
- Production underway
 - 3 complete modules will be built
 - The first fully-assembled gFEX production module has now been received and is under test at BNL
 - Results so far are good
- v3 prototype plans
 - Integration tests in LAr EMF, setting up.
 - Run parasitically in beam in 2018 (TBC)



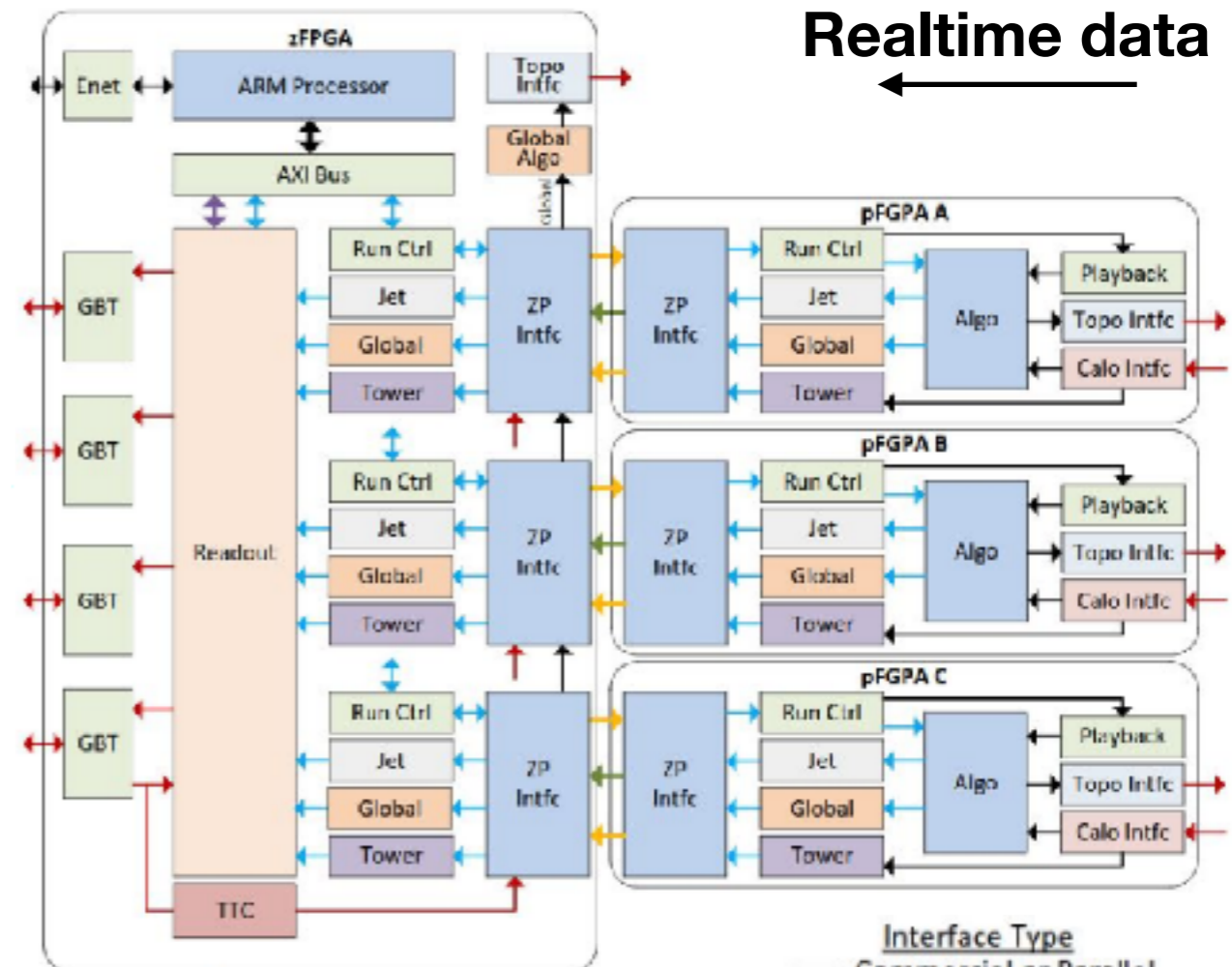
The gFEX Prototype V2, module 2



gFEX Firmware

- Firmware complete but needs porting from v3a prototype (3 x VU160 Processor FPGAs, 1 x ZU19 FPGA)
- → Ultrascale+ FPGAs
- → Vivado 7.3 s/w
- Working to improve robustness of inter-fpga communications
- Processor FPGA utilisation ~4%

Site type	Used	Fixed	Available	Util %
CLB LUTs	39767	0	926400	4.29
LUT as Logic	39767	0	926400	4.29
LUT as Memory	0	0	219840	0.00
CLB Registers	61271	0	1852800	3.31
Register as Flip Flop	53669	0	1852800	2.91
Register as Latch	7332	0	185280	0.40
CARRY8	45	0	125760	0.04
F7 Muxes	2753	0	503040	0.55
F8 Muxes	452	0	251520	0.18
F9 Muxes	0	0	125760	0.00

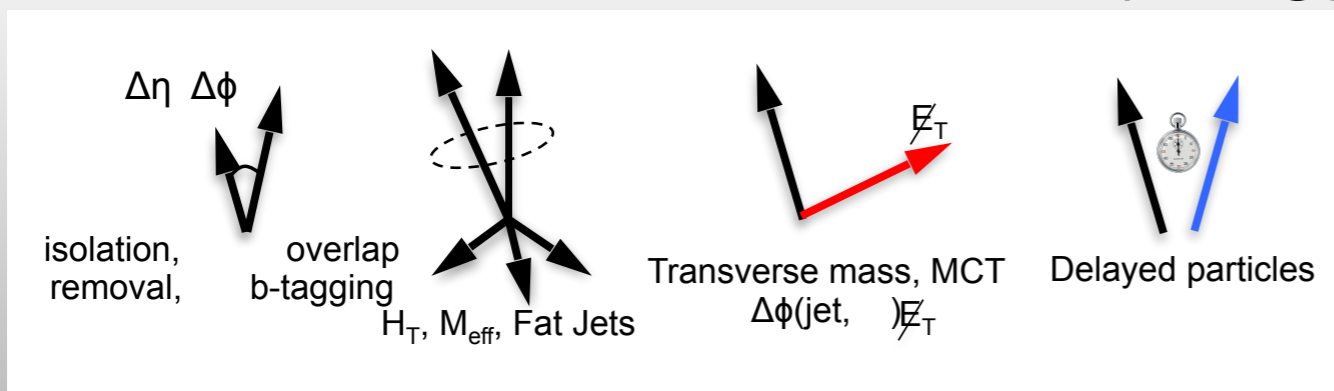


Realtime data ←

Latency (BC)	pFGPA	zFPGA
2.0	receiver & deserialization	
1.0	demultiplexing & synchronization	
5.0	primitive processing (Algorithms)	
1.0	TOB selection for output	global fragment transfers
1.0	TOB selection for output	global TOB processing
1.0	multiplexing	multiplexing
2.0	transmitter & serialization	transmitter & serialization
2.0	optical fiber to L1Topo (10 m)	optical fiber to L1Topo (10 m)
15.0	Total latency for gFEX	

Event selection based on event topology

Phase-I Topo

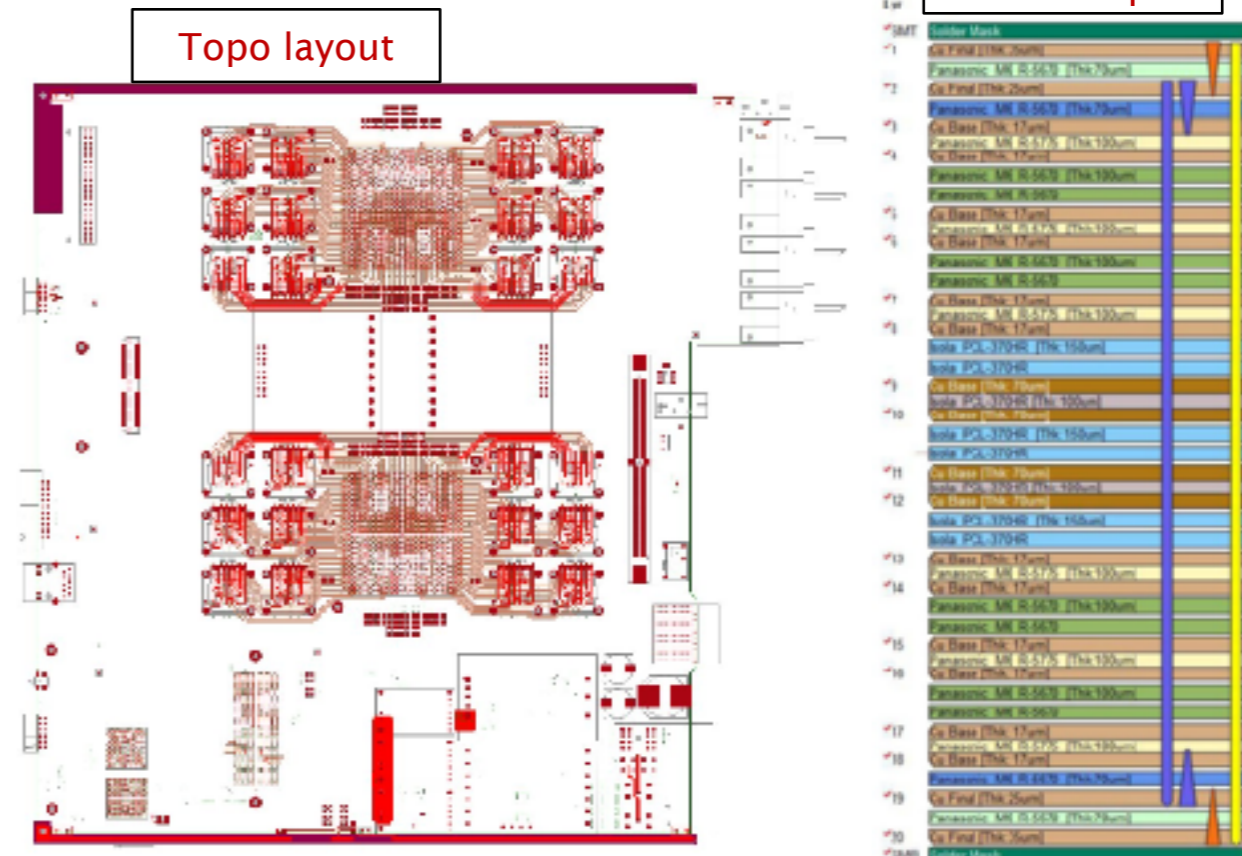


Topo placement guide

- Re-design of current Topo HW based on jFEX
 - 3 ATCA Modules
 - 2 XU+ FPGAs per module
 - x3 processing power of current Topo
 - 118 MGT inputs per FPGA (@ ≤ 12.8 Gb/s)
 - Electrical and optical output to CTP
- PDR/FDR passed (Nov' 17)
- Schematic entry completed
- PCB layout in progress
 - Stack-up reduced from jFEX 24 \rightarrow 20 layers
- Milestones
 - Receive prototypes Jun' 18
 - PRR Sep' 18



Topo PCB stack-up

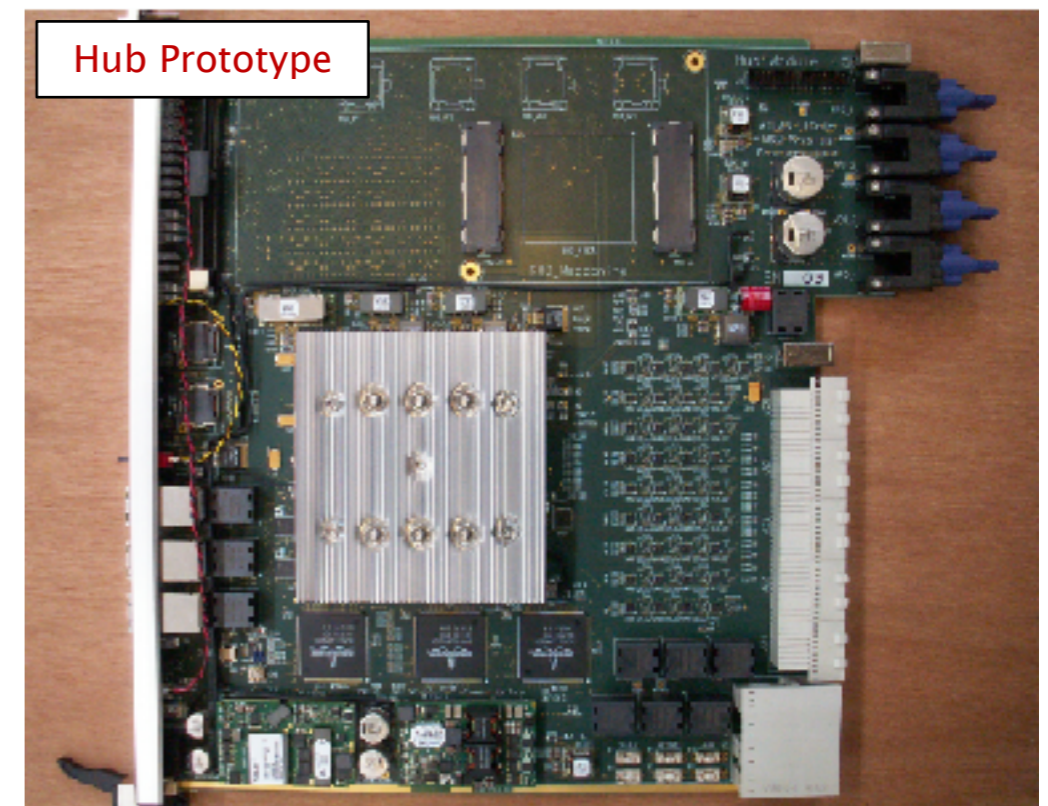
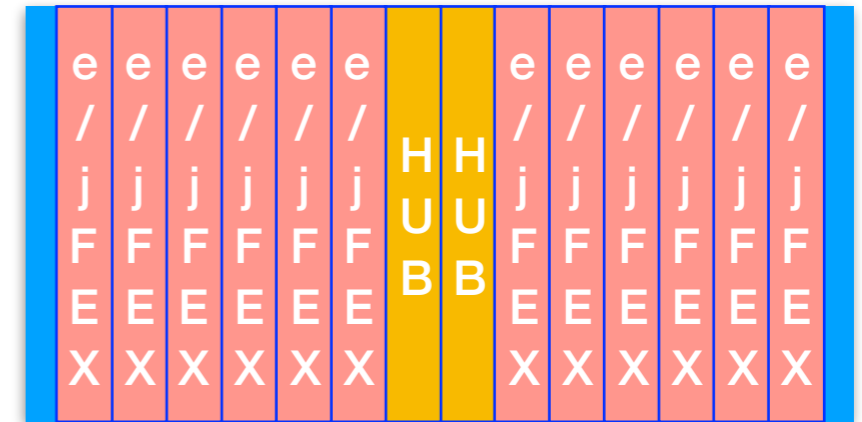


Control and clock hub in eFEX & jFEX shelves Houses L1Calo ROD

Hub

- Backplane use:
 - Zone-1 X 100 kbps Intelligent Platform management Bus / slot
 - Zone-2 Fabric: 8 channels Node-Hub (<10 Gbps) (slightly non-standard..)
 - One signal pair is used to carry the 40 MHz TTC clock from the Hub-ROD
 - One signal pair is used to carry re-coded TTC commands (6.4 Gb/s)
 - Six signal pairs are used (or reserved) to carry readout data to the Hub-ROD (6.4 Gb/s)
 - Zone-2 Base: 4 channels Node-Hub and Hub1-Hub2 (GbE)
 - Zone-3 Optics: MPO connectors to MiniPods
- 8 Hub Prototypes manufactured May'17
- Tests completed
 - Ethernet connectivity
 - Inc. switch bandwidth characterisation
 - Clocking (inc. transition under TTC change)
 - MGT backplane links verified for all channels, all Hubs, using IBERT
 - BER 10^{-16} @ 6.4 Gb/s
- To do: test signal integrity of Hub with full ATCA shelf of source modules (May-Jun'18)
 - Hub Test Module in manufacture
 - Carrier for commercial Trenz TE0782
- Firmware
 - IPBus implemented
 - Combined TTC-ROD-Hub MGT downlink to e/jFEX under test
- Milestones
 - Production Readiness Review Jul'18

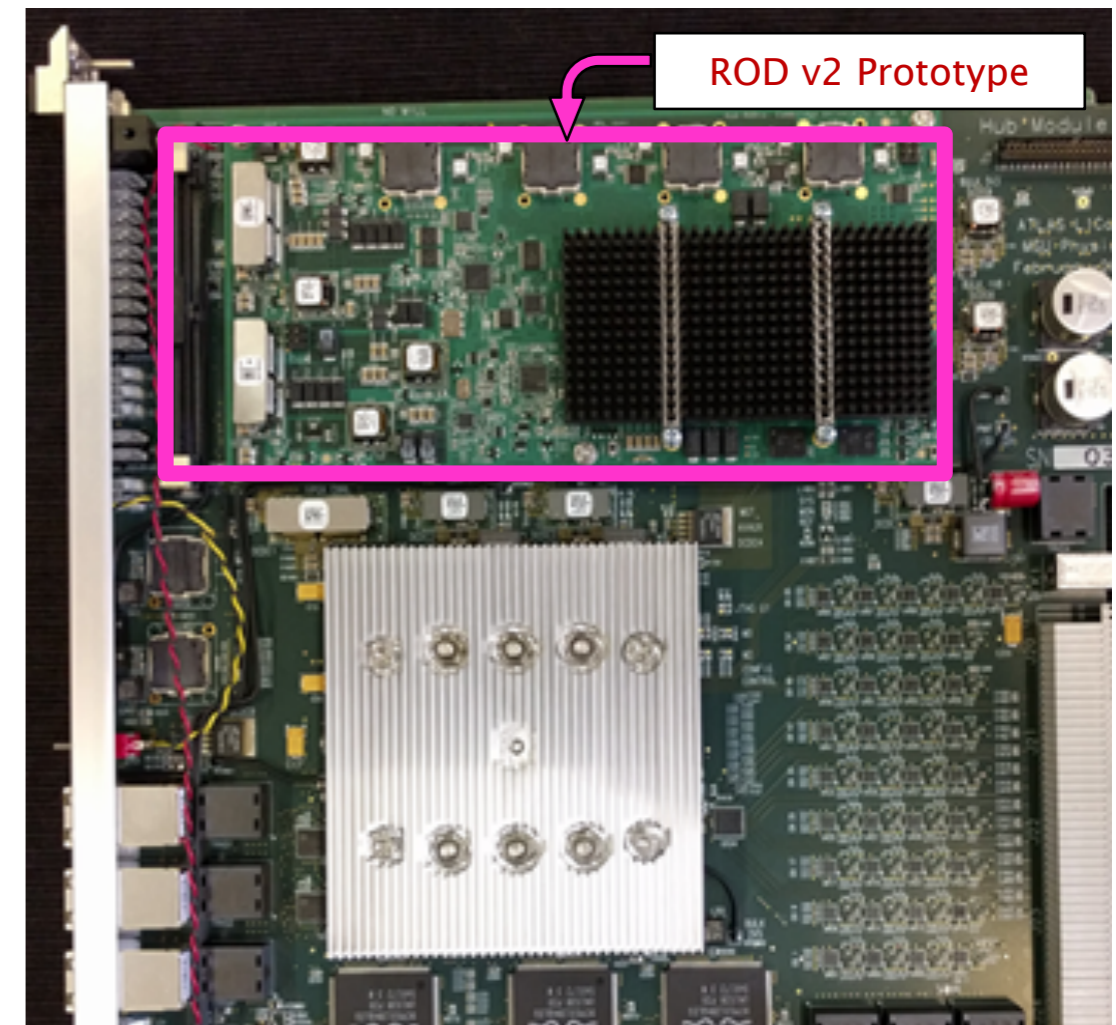
FEX/Hub shelf



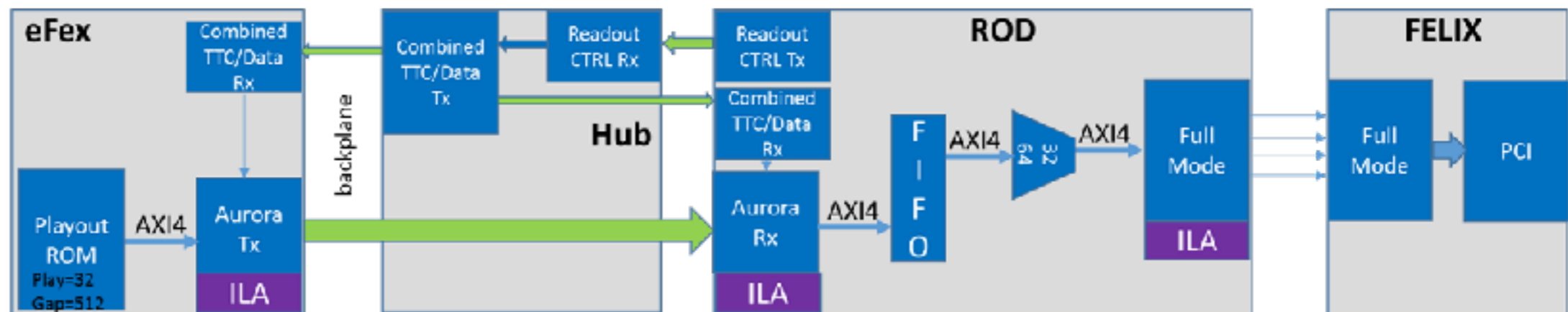
L1 Calo Readout Driver
Collects & buffers data across eFEX,
jFEX & Topo shelves;
transmits to FELIX

L1 Calo ROD Hardware

- Prototype v2 received Nov'16 (5 modules)
- All interfaces verified in hardware
 - MGT input across backplane (6.4 Gb/s) & optical output (9.6 Gb/s)
 - BER $<10^{-16}$ @ 6.4 Gb/s
- Pre-production design complete
 - (minor changes from prototype)
- Milestones
 - Test readout slice: eFEX → Hub-ROD → FELIX, Ongoing (all ROD fw ready)
 - FDR Apr'18



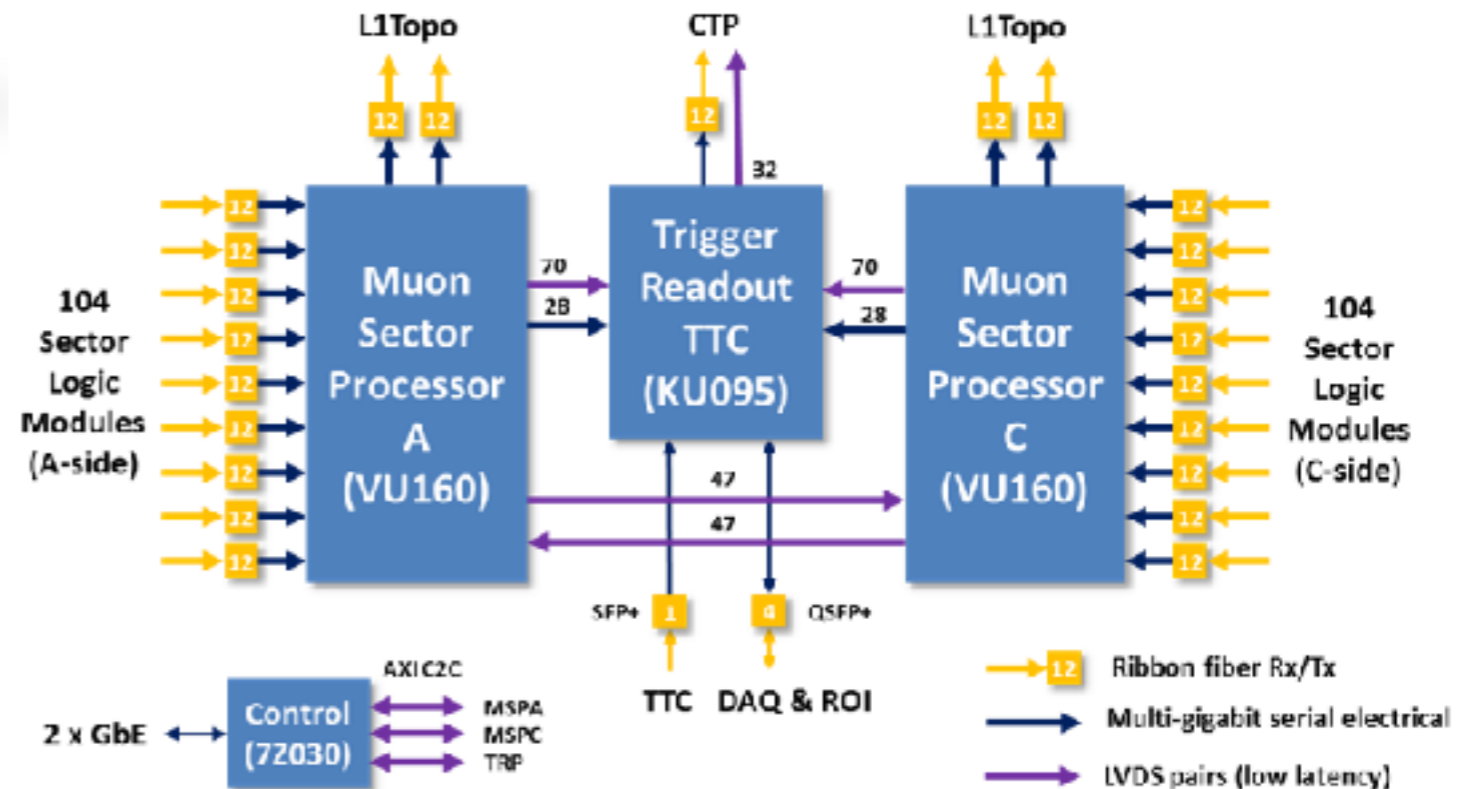
ROD readout slice test



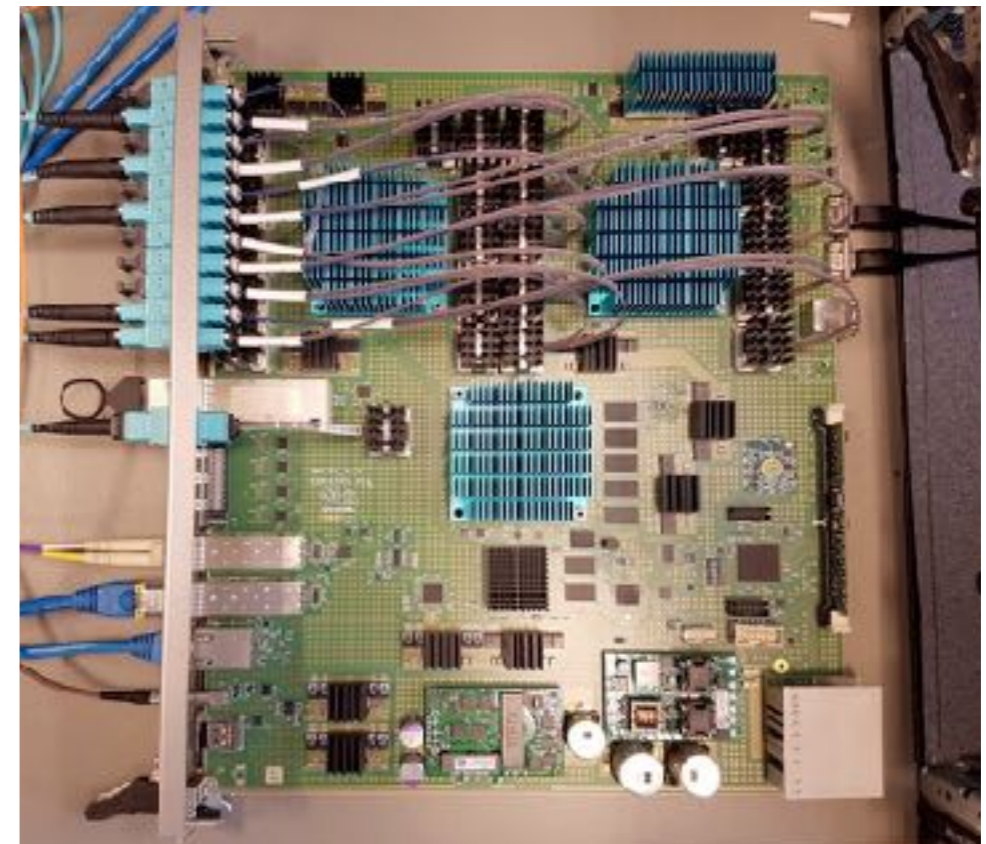
Interface between Muon Barrel and Endcap triggers and Central Trigger Processor (one module)

- Hardware status:
 - fully assembled MUCTPI prototype available
 - 100% of board tested successfully
 - only minor modifications required
 - PCB layout already modified
 - 2 partially assembled prototypes (w/o high-end FPGAs) available for SW/ FW development
 - hardware FDR passed in January 2018
 - integration test with TGC and RPC SL done

MUCTPI



MUCTPI 1st prototype



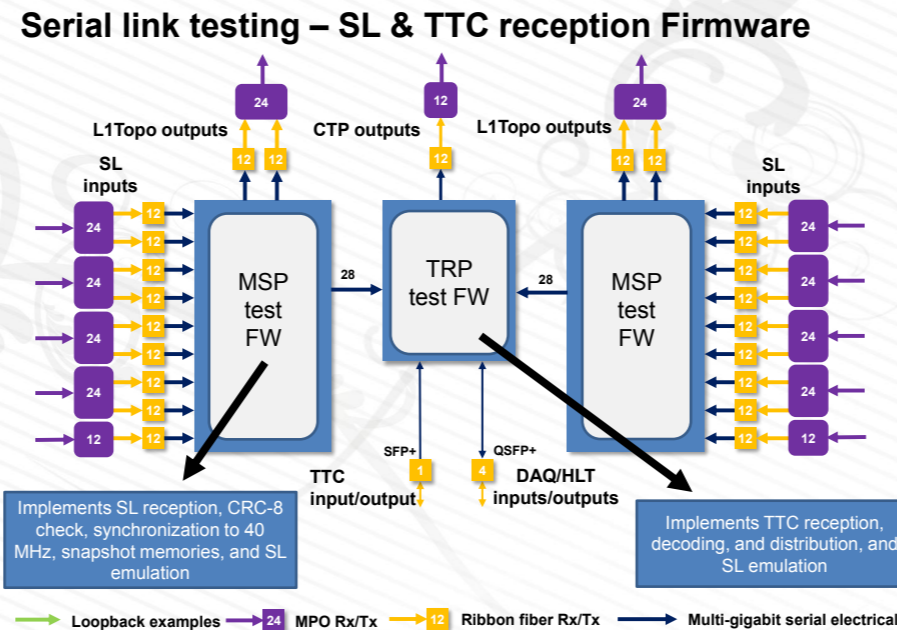
MUCTPI

- Firmware status:

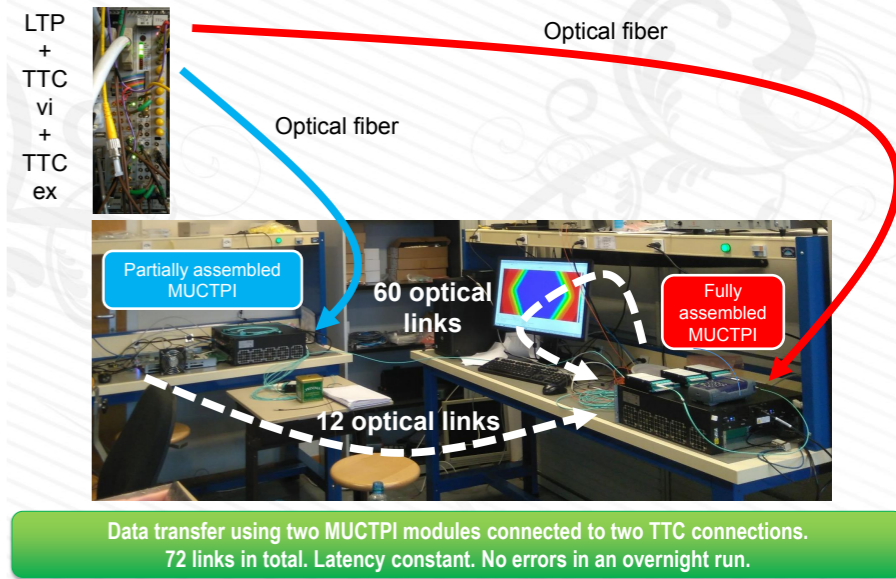
- board testing and infrastructure firmware ready
- SL reception and TTC interface firmware ready and tested
- trigger processing, readout and monitoring firmware, to be developed
 - builds on experience from current MUCTPI firmware

- Software status:

- The Zynq runs embedded Linux prepared using Yocto and the Xilinx metalayer
- software to control the HW (I2C, SPI, etc.) and FPGAs (using AXI Chip-2-Chip) developed and tested
- An ATLAS Run Control Application was ported to the PS using embedded Linux
- ROOT was ported to the Zynq and can be used to produce histograms
- low/high level SW ready to expand as the firmware evolves



SL Reception Firmware Lab Test



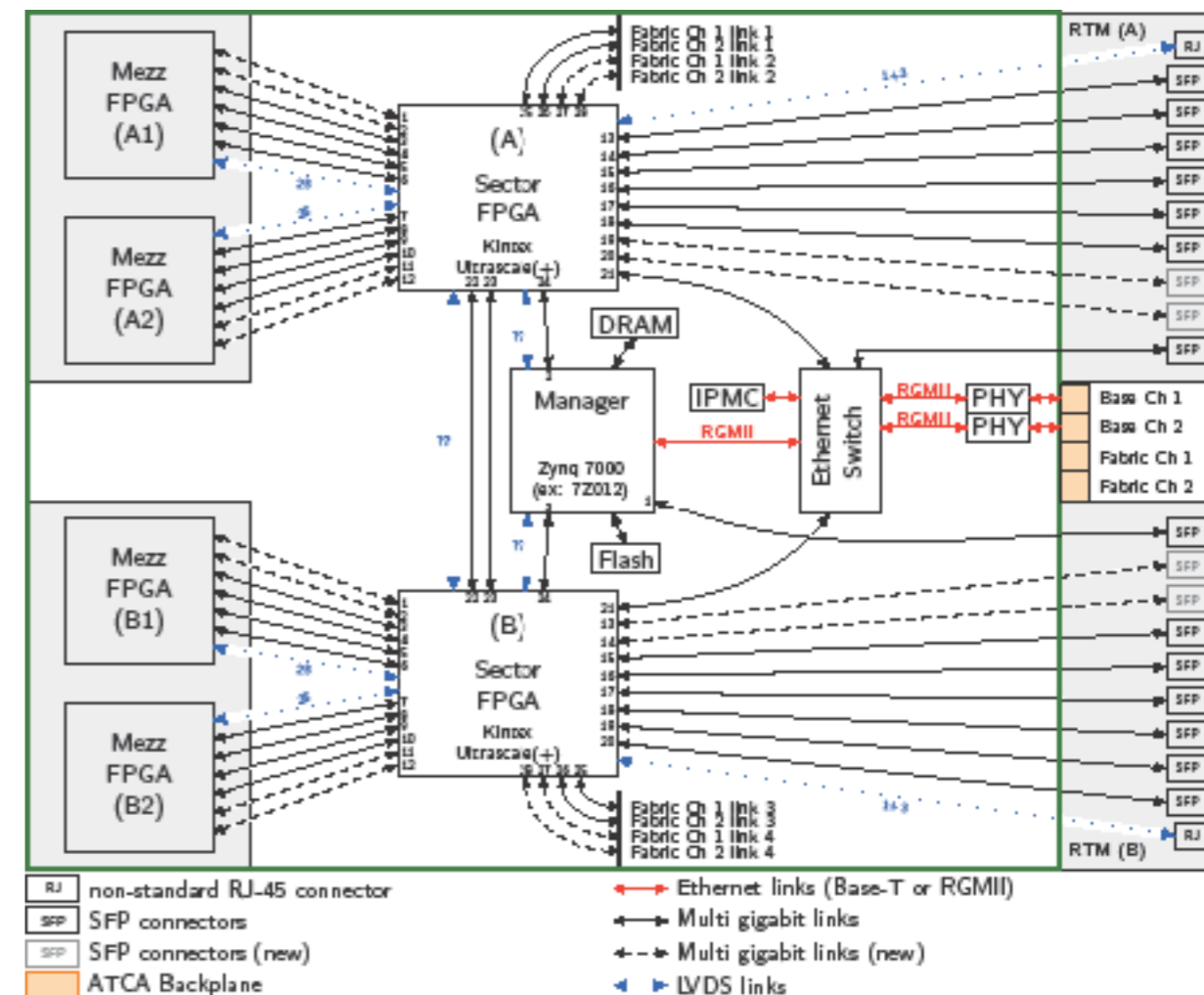
- 2nd prototype with Ultrascale+ FPGAs available and tested : Q2 2018
- 3rd prototype with Ultrascale+ MPSoC (64-bit, quad-core) available and tested: Q3 2018

Long-term Milestones

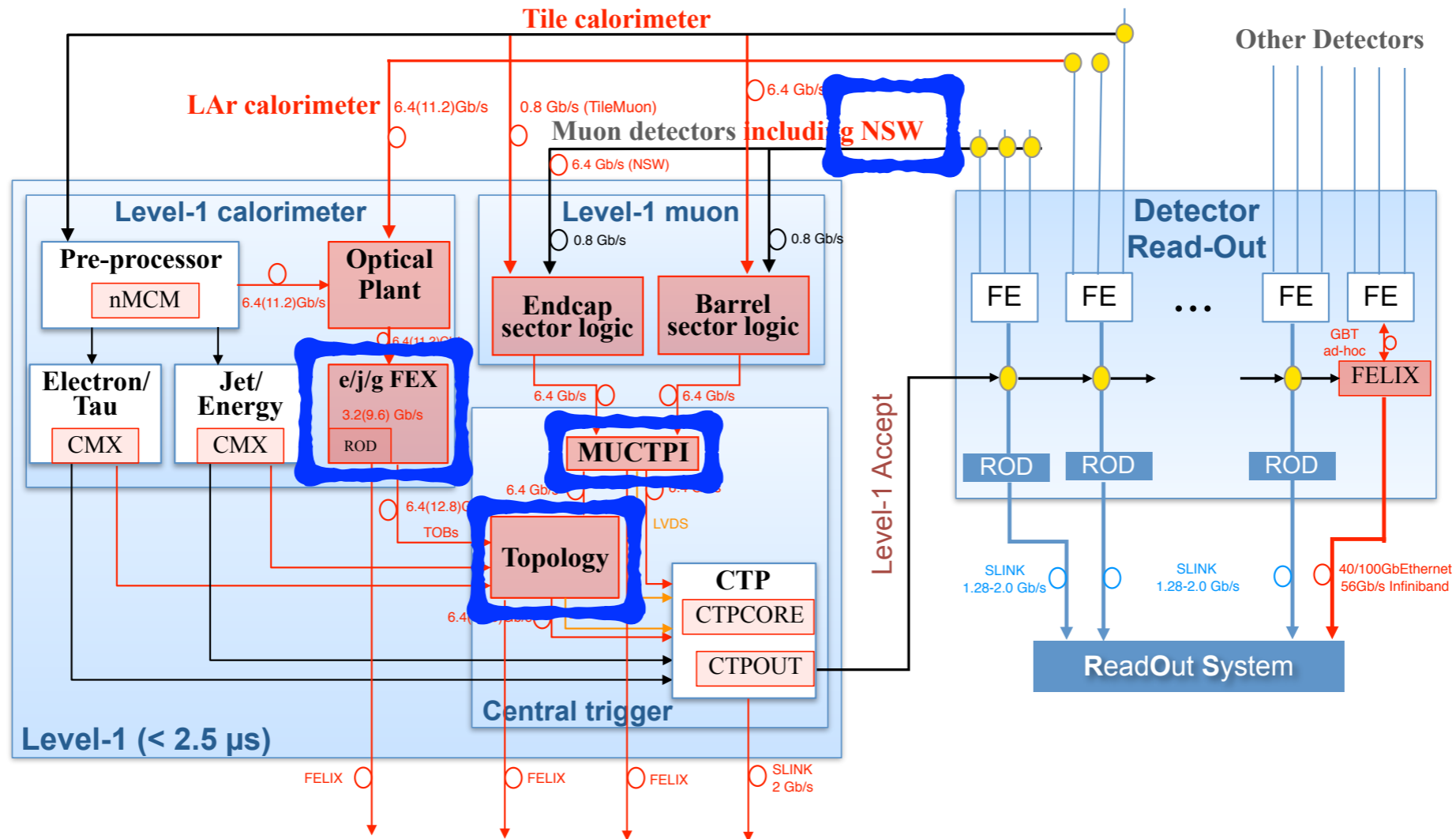
Milestone	Date	
FDR	(Q2-2017)	HW Jan-18
	Q4-2017	FW Q3-18
PRR	(Q1-2018)	Q4-18
	Q2-2018	
Production completed and tested	Q4-2018	Q1-19

ATLAS NSW trigger processor

- ATCA cards with two custom mezzanines and RTM
 - 16 cards required
- 4 Virtex 790 for trigger algorithms + 2 Ultrascale Kintex + Zynq 7000
- 144 fibers (minipods), input at 4.8G, 12x6.4G outputs to Sector Logic
 - MM inputs are from GBTx; sTGC uses custom scrambler
- Mezzanine: One FPGA per detector type (sTGC, Micromegas)
 - Different fiber lengths $\rightarrow \Delta t > 25 \text{ ns} \rightarrow$ must be aligned
 - 34 LVDS low-latency pairs between FPGAs for merging segments
- FELIX link for TTC, configuration, output of Level-1 data, monitoring
 - BCs that are not triggered by Level-1 can be sampled and monitored via FELIX
- OPC for configuration of operating parameters:
 - Register file in the FPGA will be seen via an OPC server by both DCS and DAQ configuration SW
 - OPC server uses FELIX E-links for its connection to the Trigger Processor
- Existing carrier is Virtex6; being upgraded to Kintex Ultrascale



ATLAS ATCA-based system size

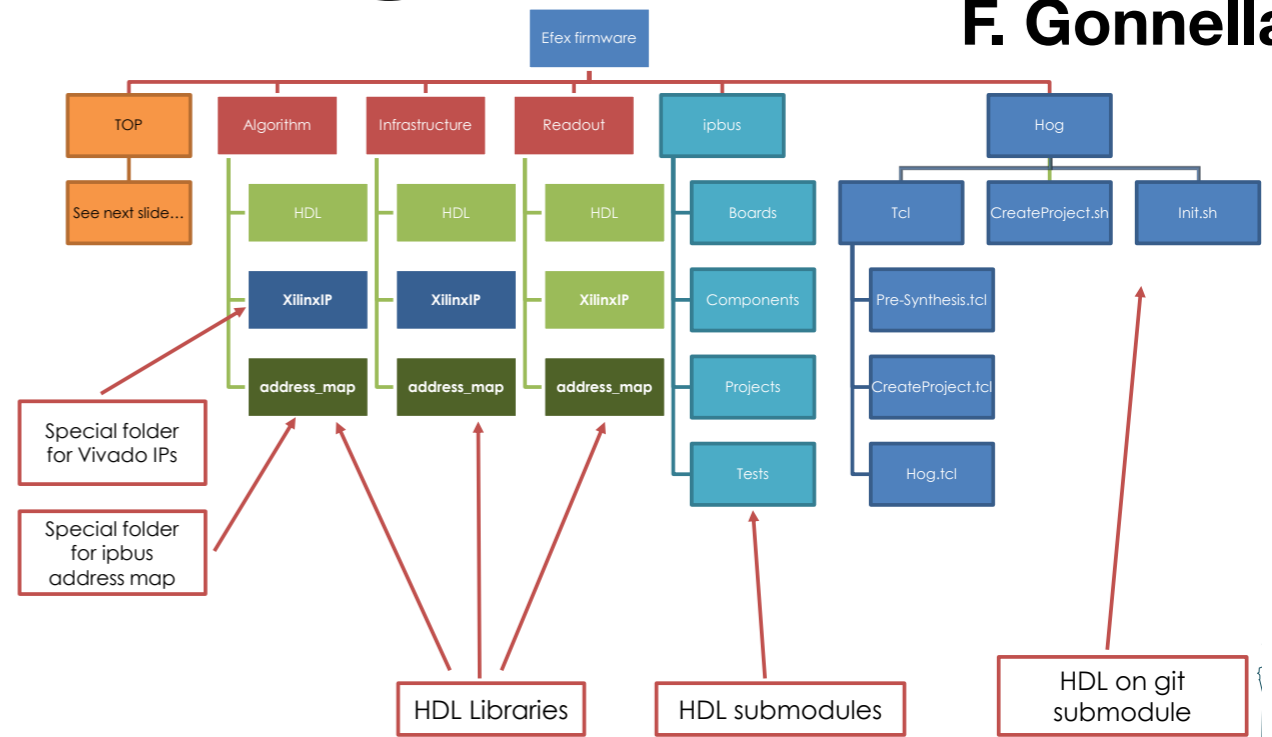


Subsystem	Component	Number of blades
Level-1 Calorimeter	eFEX	24
	jFEX	6
	gFEX	1
	L1Topo	3
Level-1 Muon	NSW	16
	MUCTPI	1

Firmware Management

F. Gonnella

- Collaborative Firmware Development is currently addressed in L1Calo (currently eFEX and gFEX, Hub and ROD projects next).
- Developers use Vivado normally, through an **Automated Workflow Engine** (AWE - gitlab.cern.ch/atlas-l1calo-efex/awe), python script running on CERN Openstack virtual machines:
 - Script-based flow, absolute control of HDL, constraint files, Vivado settings.
 - Use of git (HDL on git - gitlab.cern.ch/atlas-l1calo-efex/Hog):
 - AWE triggered by merge requests, produces bit files, reports, documentation.
 - git SHA embedded in firmware registers
 - Guarantee firmware synthesis reproducibility and traceability



eFEX board @efex commented 5 da
Project: control_fpga

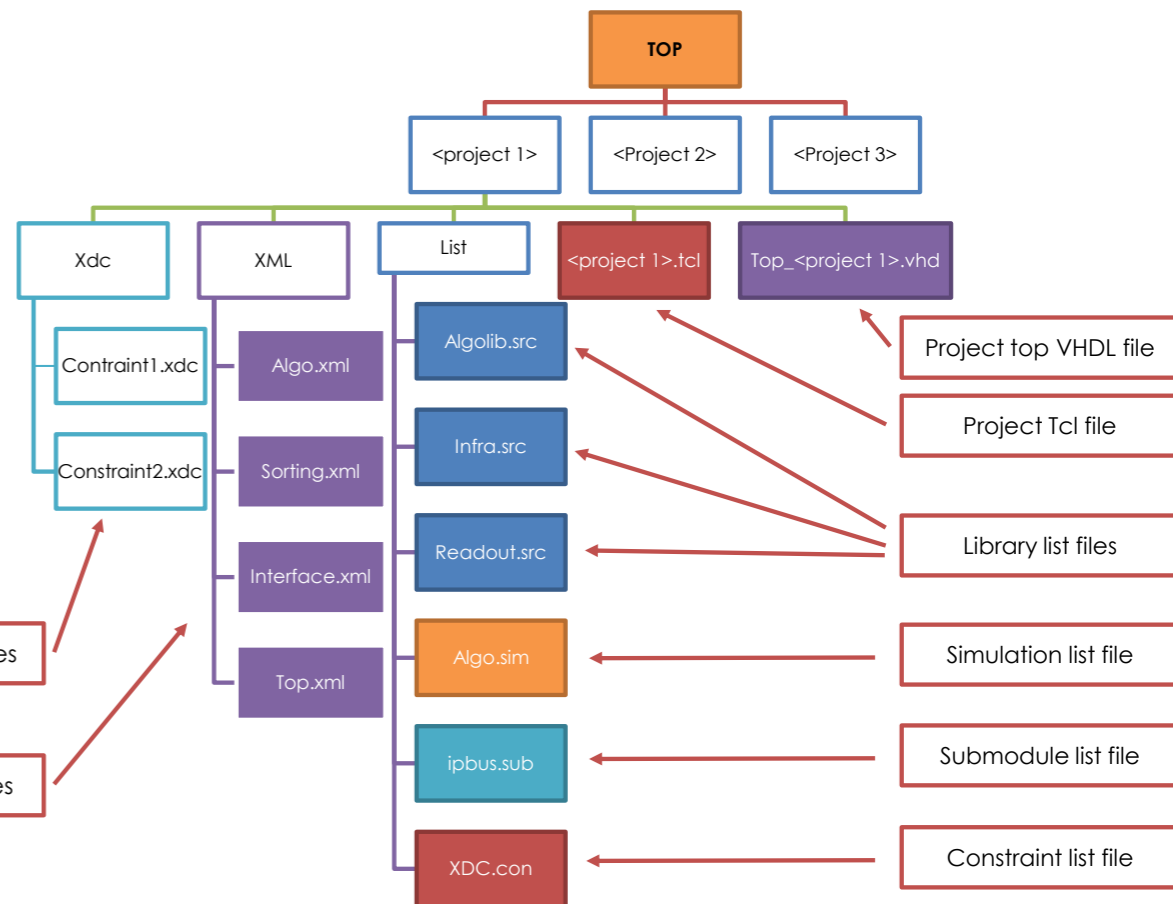
Work-flow was successfull

Timing report

Parameter	Value
WNS(ns)	0.225
TNS(ns)	0.000
TNS Failing Endpoints	0
TNS Total Endpoints	12762
WHS(ns)	0.050
THS(ns)	0.000
THS Failing Endpoints	0
THS Total Endpoints	12753
WPWS(ns)	0.264
TPWS(ns)	0.000
TPWS Failing Endpoints	0
TPWS Total Endpoints	5491

Constraint files

IPBus XML files



All user specified timing constraints

Conclusions

- ATCA blades in Phase-I ATLAS have been designed with specific functionalities in mind (EM, Jet, Topological processors), due to constraints from the legacy system.
- Different technical solutions have been adopted, but a pattern in the technical solutions adopted can be found, where:
 - a blade is carrier for few large processing FPGAs, with MGT optical IOs, inter-FPGA on board connections using low and high-level protocols.
 - A limited number of subsystems use the available ATCA infrastructure, like backplane connections, for trigger/readout functionalities.
 - Limited use of IPMC for configuration and control. Complementary SoC is added, control and monitoring done through GbE.
- The above pattern will be carried further in Phase-2, where general blade designs will be adopted for different functionalities and subsystems. The technology required in Phase-2 is essentially available now.
 - please see next talk by E. Hazen.
- Design effort is moving away from hardware design to firmware, where more and more person-power is needed in a collaborative environment suited to different design skills. This our current challenge. It will become more significant in Phase-2.