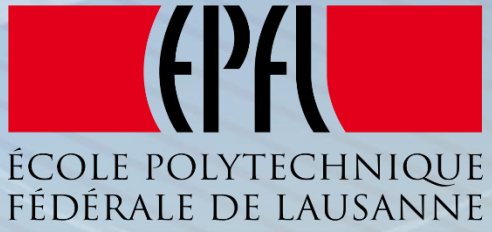
The ATLAS MUCTPI Upgrade for the Run 3 of the LHC

Marcos Vinicius Silva Oliveira

Supervisors: Stefan Haas (CERN), Alain Vachoux , Yusuf Leblebici (EPFL)

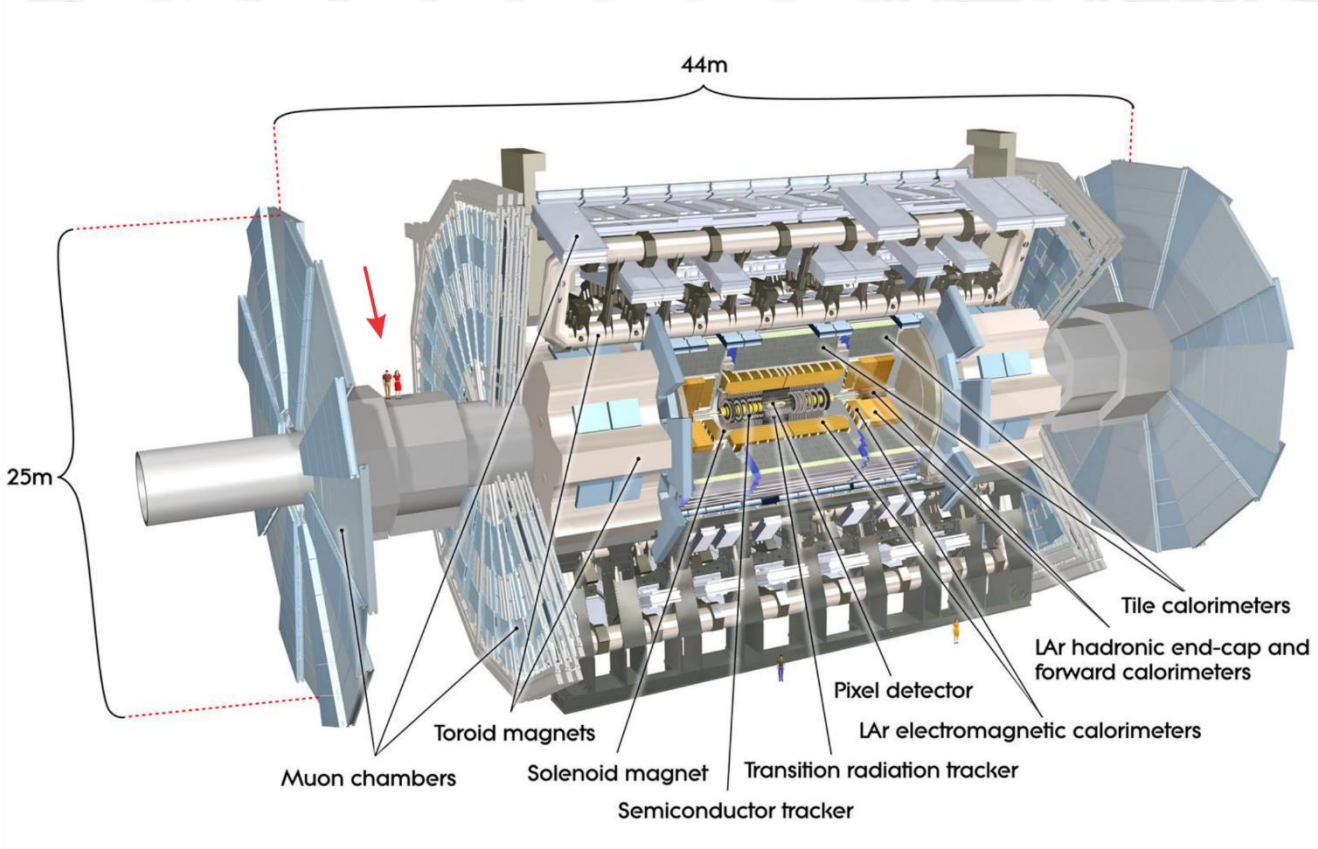


Outline

- Introduction
- MUCTPI architecture
- Preparatory work
- MUCTPI prototype
- Firmware & tests
- Summary & next steps

ATLAS

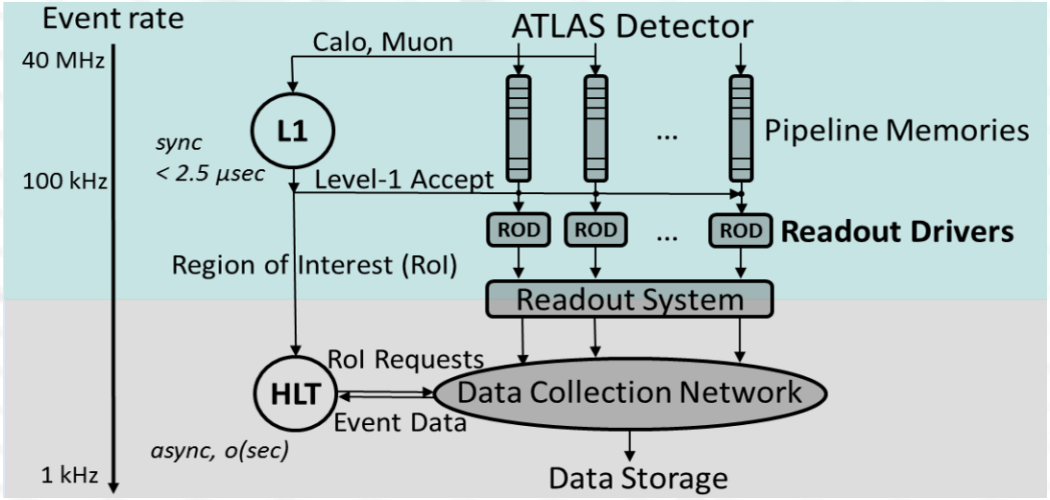
- One of the 4 LHC experiments, observes proton-proton collisions
- Process physics event data at high rates from thousands of channels
  - Large amount of data → requires on-line filtering, a.k.a. trigger



Trigger and Data Acquisition System (TDAQ)

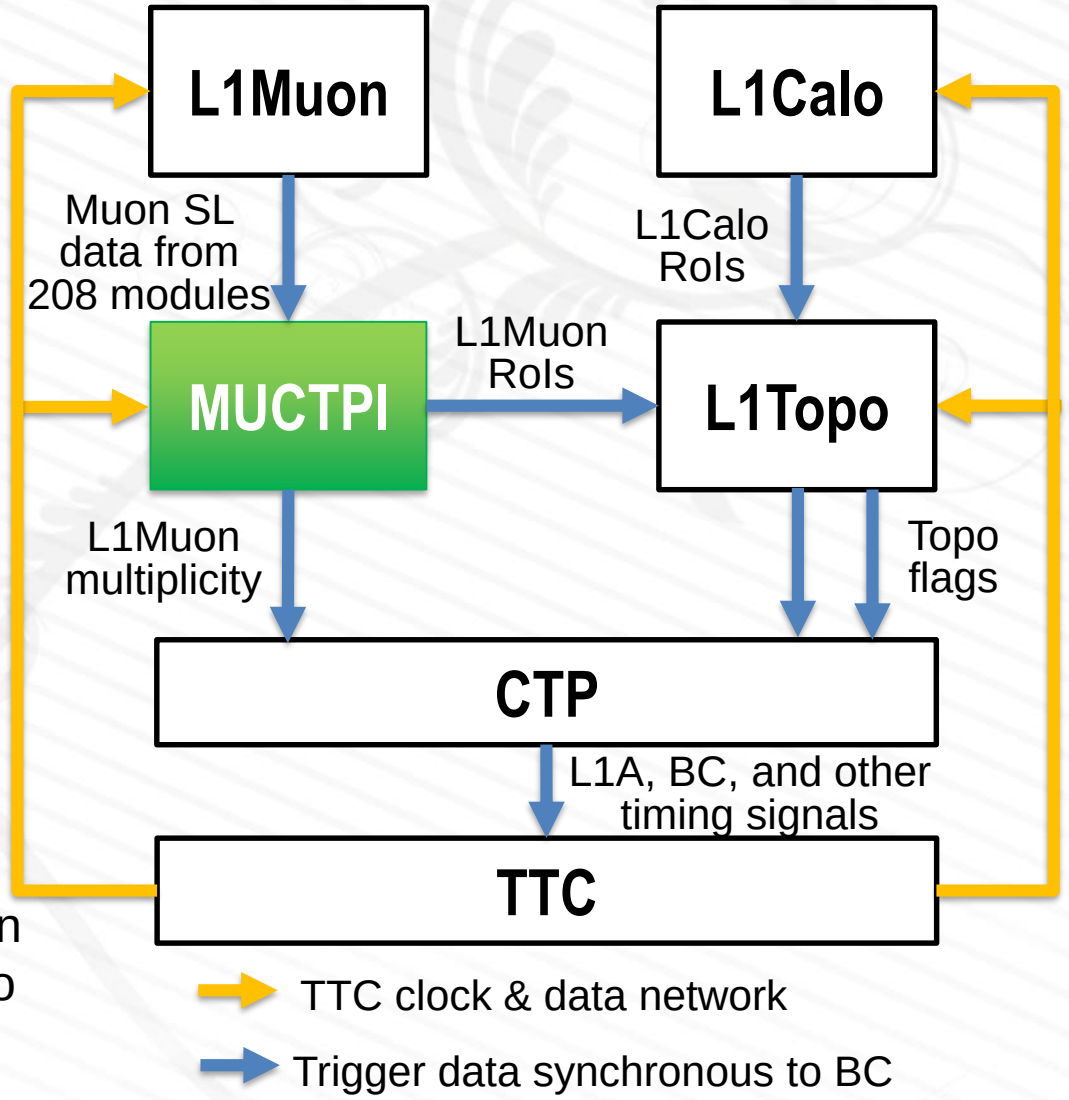
- Reduces BC rate of 40 MHz → 1 kHz (permanent storage)
- 2 levels:

Level-1 Trigger	High-Level Trigger (HLT)
40 MHz → 100 kHz (real time processing)	100 kHz → 1 kHz (asynchronous processing)
Low-latency (2.5 us) (on-detector buffer size) → fast event selection based on reduced-granularity calorimeter & muon data	**Lower input event rate** (thanks to Level-1 Trigger) → Relaxed-latency (> 1 s) → event selection based on complete detector information
Custom electronics, **FPGA** for off-detector → high processing capacity and reprogrammability	commercial computers, network switches, and custom **software**



Level-1 Trigger System

- Low-latency, high-throughput
- Data transfer based on system-synchronous clocking technique
  - Requires fixed-latency processing and data transfer
- MUCTPI:
  - Combines L1Muon data
  - Count and sort muon candidates according their transverse energy
  - Avoids double counting
  - Encodes & sends muon position & energy information to L1Topo

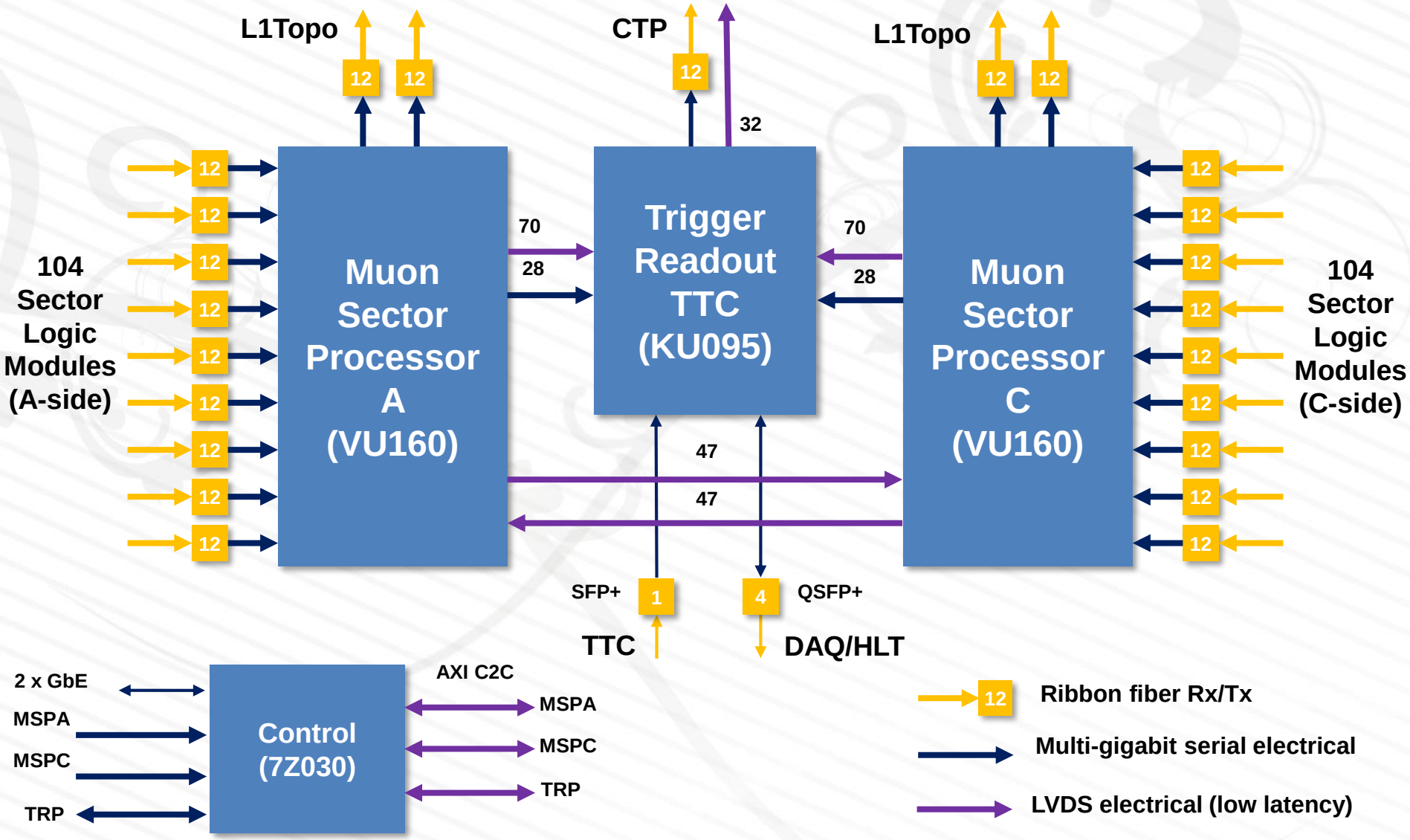


MUCTPI upgrade

- LHC luminosity increases → trigger needs to become more selective
- Better performing algorithms → more data (BW increases), processing units receives information from larger parts of detector (data concentration increases)
- For achieving it: use of high-speed serial optical communication, higher FPGA densities
  - High-speed serial communication → BW increases, latency increases, both compared to parallel electrical transmission
    - More muon candidates can be received (up to 4 instead of 2)
    - Finer energy threshold (Up to 4 bits (Phase-I), up to 6 bits (Phase-II) instead of 3 bits)
    - Finer position (RoI of up to 16 bits instead of 8, Phase II only)
  - Higher integration → latency decreases, data concentration increases both compared to current implementation in a full 9U VME crate with 18 boards
    - full 9U VME crate (18 boards) → single ATCA blade
    - Overlap handling across octants
    - New algorithms (muon-only topological trigger)

MUCTPI architecture

MUCTPI architecture

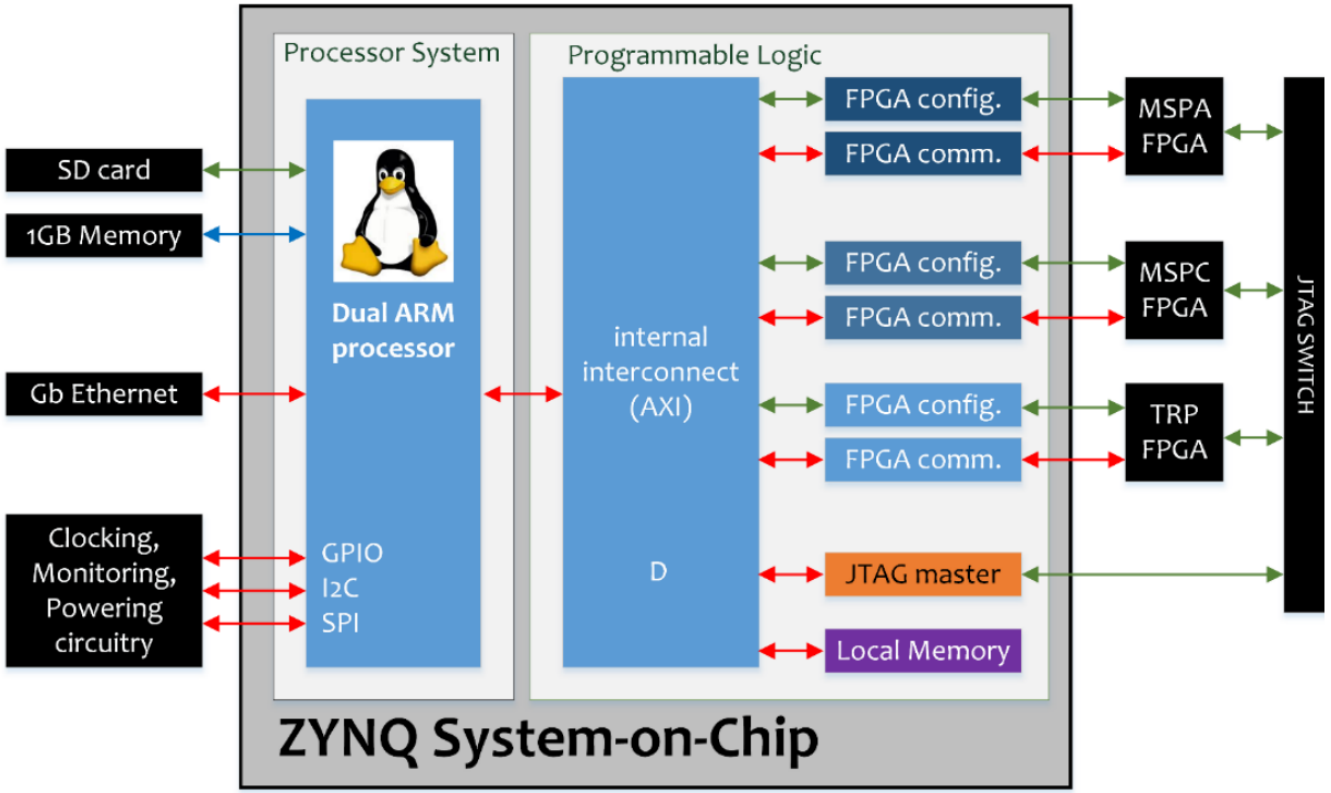


MUCTPI architecture

- 2 **Muon Sector Processor (MSP)** FPGAs (Xilinx Virtex Ultrascale VU160)
  - 1 FPGA handles ½ of the muon trigger
  - Muon sector data reception & timing alignment
    - Uses 9 RX MiniPOD to receive 104 MGT Rx
  - Muon trigger object output to L1Topo
    - Uses 2 TX MiniPOD to send (48 MGT Tx)
  - Overlap handling: suppress double counting of single muons
  - Monitoring: rates & per-bunch histograms per sector
  - On-chip playback & snapshot memories
- **Trigger and Readout Processor (TRP)** FPGA (Kintex Ultrasclale KU095)
  - Receive and merge information from 2 MSP FPGAs
  - Calculate global muon candidate multiplicities
  - Implement muon-only topological algorithms
  - Send trigger multiplicities and flags to CTP
  - DAQ readout, HLT output
  - Event monitoring
  - TTC reception, decoding and distribution

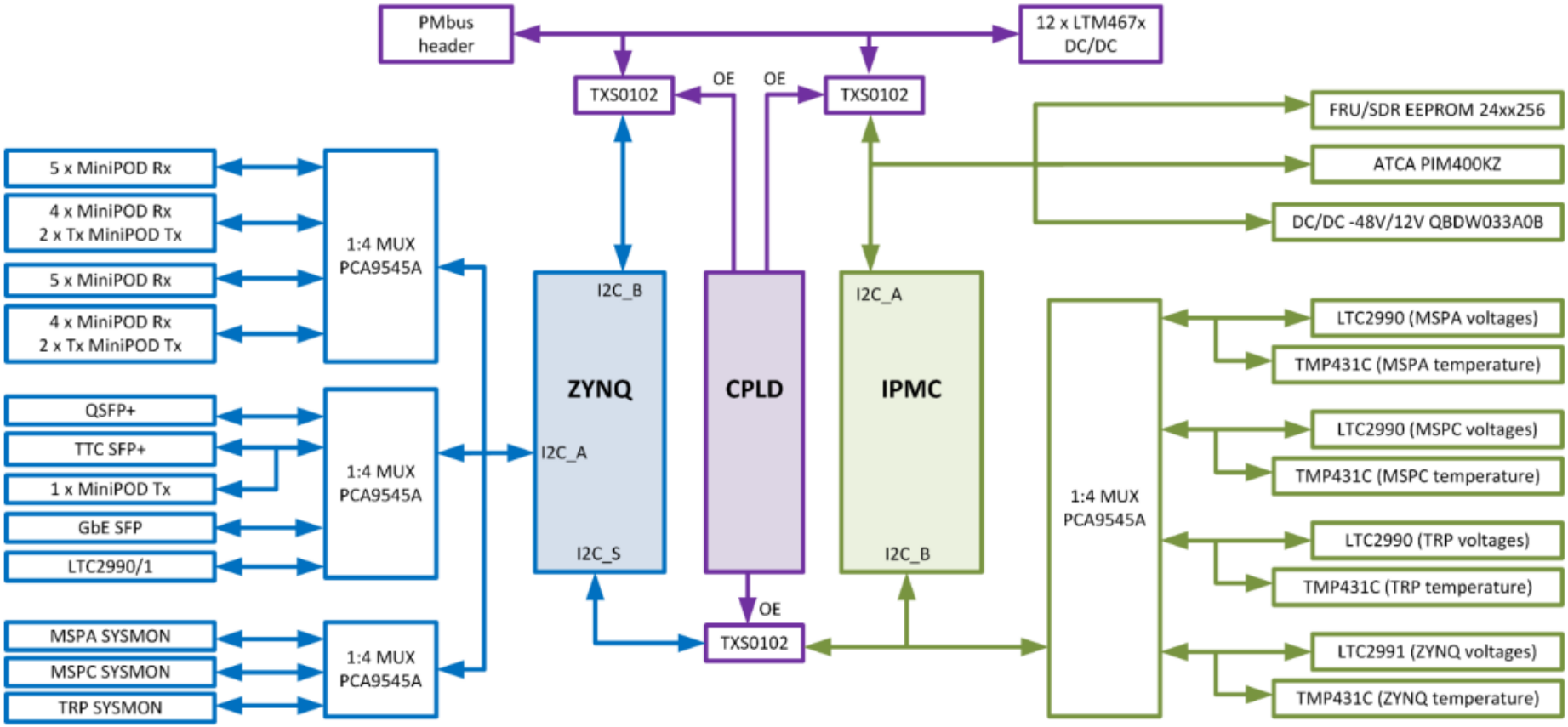
MUCTPI architecture

- Xilinx Zynq dual-core ARM SoC with programmable logic running embedded Linux OS and Gigabit Ethernet connection with PC
- Zynq handles configuration, control and monitoring of the board
- Tools for testing and debugging the hardware available



MUCTPI hardware monitoring

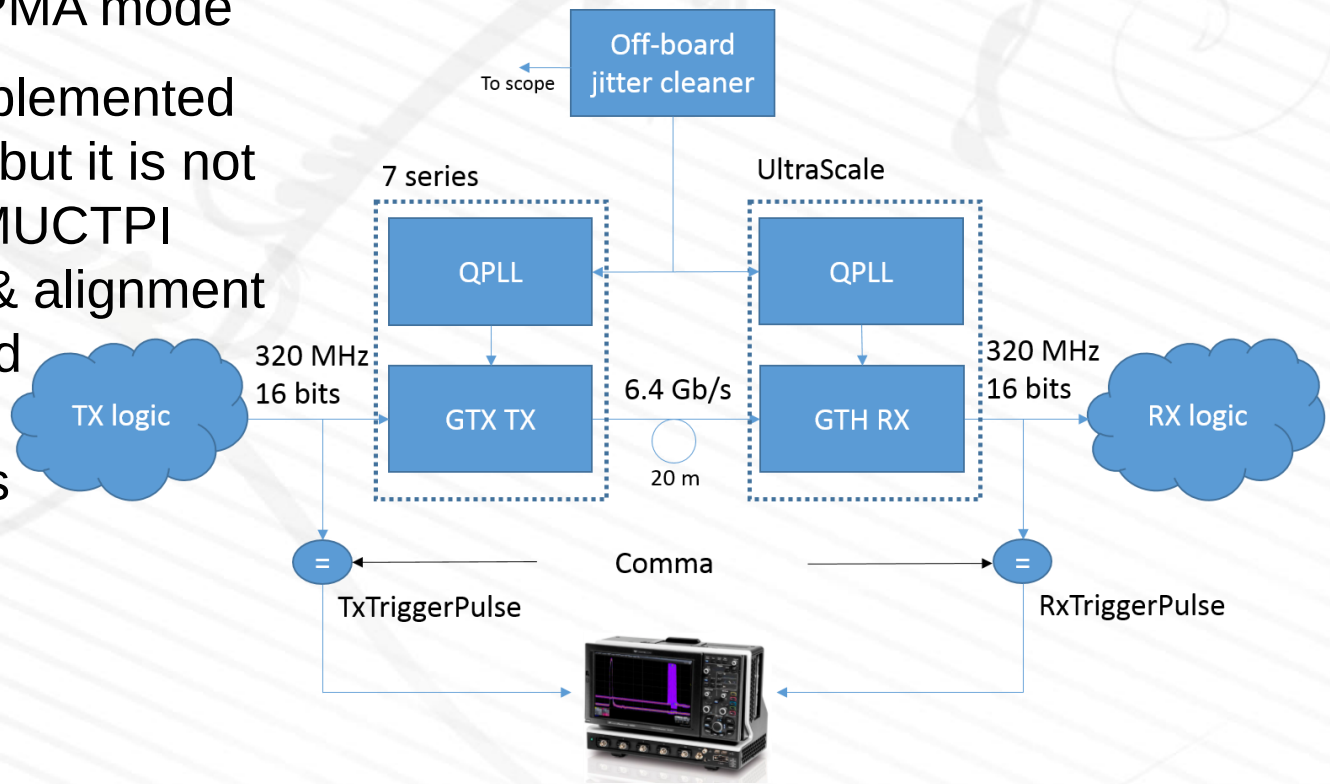
- I2C network to monitor board status (~500 values)
  - Temperatures, currents, voltages, optical modules, alarms …
- Independent monitoring path for Zynq and IPMC



Preparatory work

MGT latency measurement & optimization

- Measurement of not fixed latency in most of MGT operation modes
- Found TX & RX configuration for minimum latency and reduced latency uncertainty to 1 RXUSRCLK period
- Latency variation can be can be reduced near to 0 using manual RX clock & data alignment in PMA mode
  - This could be implemented for the SL inputs but it is not required as the MUCTPI synchronization & alignment circuit is designed to absorb small latency variations



MGT power estimation / verification measurement

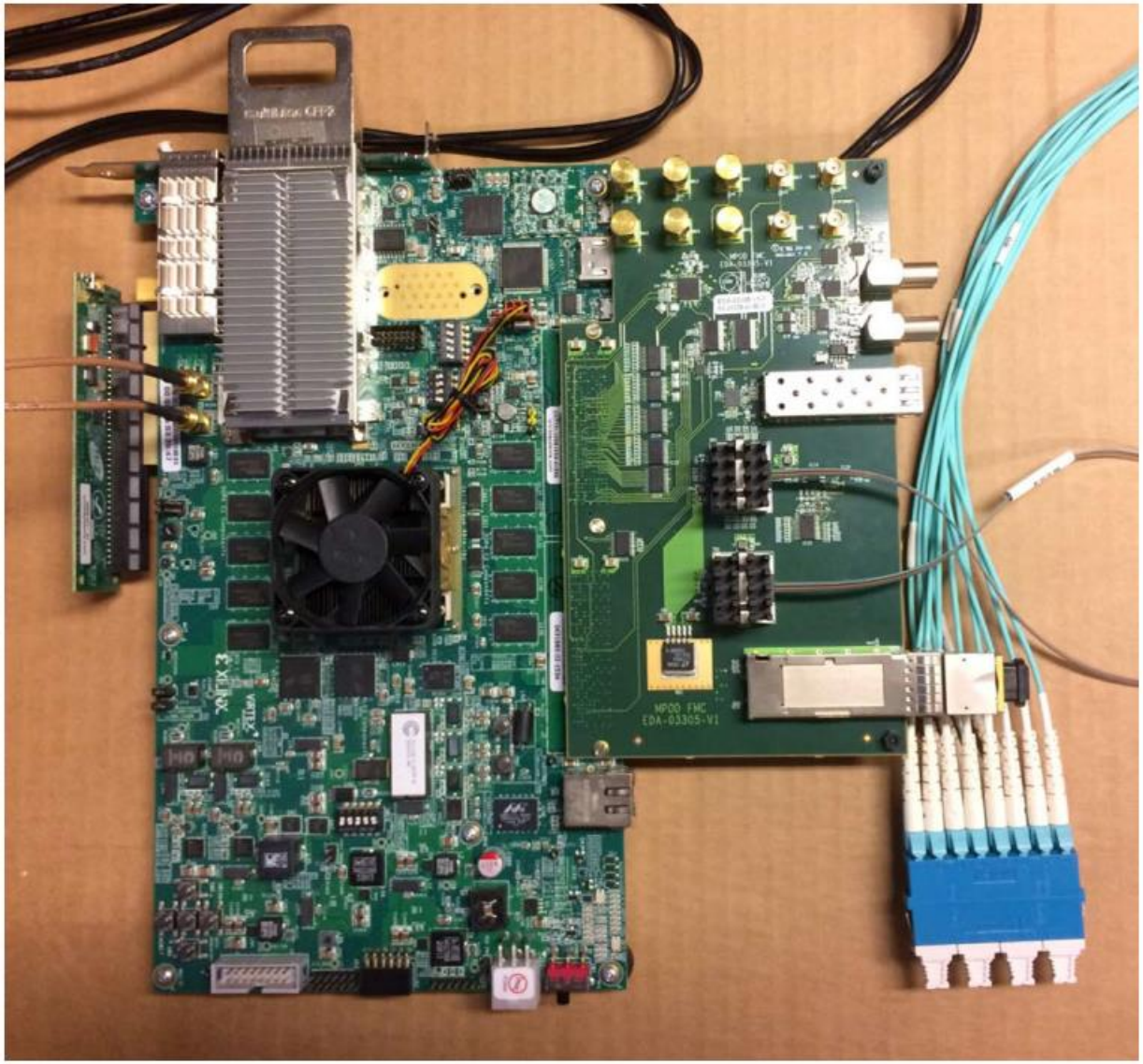
- Xilinx Power Estimation tool used
  - Accuracy (± 20 % for production devices)
  - Known problems in estimation for previous versions → verify values
- Test system based on Xilinx VCU 110
  - 104 transceivers connected (the same as MUCTPI FPGA device)
  - IBERT IP firmware with all transceivers running at 6.4, 9.6, and 12.8 Gb/s
  - Example:

				MGTAVCC			MGTAVTT		
			Scenarious	Estimated	Measured	Mismatch	Estimated	Measured	Mismatch
104 GTHGTY (4 GTY links not locked)	9.6 Gbps	GTY CPLL (not available for MUCTPI FPGA @ 9.6 Gbps)	Near-end PMA DFE TX and RX	16.11	15.40	-4.41%	18.89	21.30	12.76%
			Near-end PMA TX and RX	14.03	14.40	2.64%	17.04	20.30	19.13%
			Near-end PMA GTH TX RX GTY RX only	12.08	13.80	14.24%	10.64	14.70	38.16%
104 GTHGTY (4 GTY links not locked)	9.6 Gbps	GTY QPLL1	Near-end PMA DFE TX and RX	15.65	15.60	-0.32%	20.17	23.60	17.01%
			Near-end PMA TX and RX	13.57	14.60	7.56%	18.32	22.60	23.34%
			Near-end PMA GTH TX RX GTY RX only	11.62	13.90	19.62%	11.92	17.00	42.62%

  - **Power supply system designed to handle extra power margin**
    - **We have used 25 A regulators for MGTAVCC and MGTAVTT**

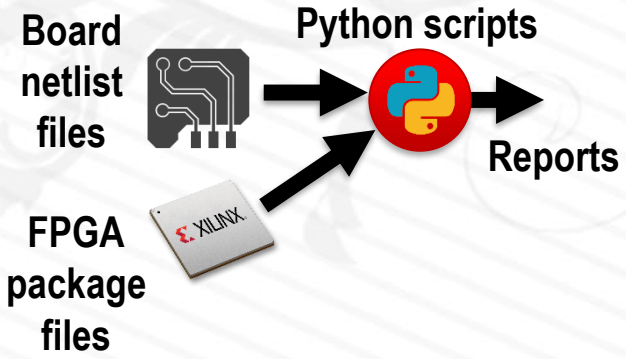
MUCTPI demonstrator

- Custom double-width FMC card
- Designed to test FPGA family, on-chip MGTs, 12-channel ribbon fiber optics receiver and transmitter modules (MiniPOD), and clock circuitry
- Designed SL reception demonstration firmware
- Used successfully for connection tests with TGC and RPC sector logic modules
- Latency measured as 4.5 BC period from SL 40 MHz to MUCTPI 40 MHz

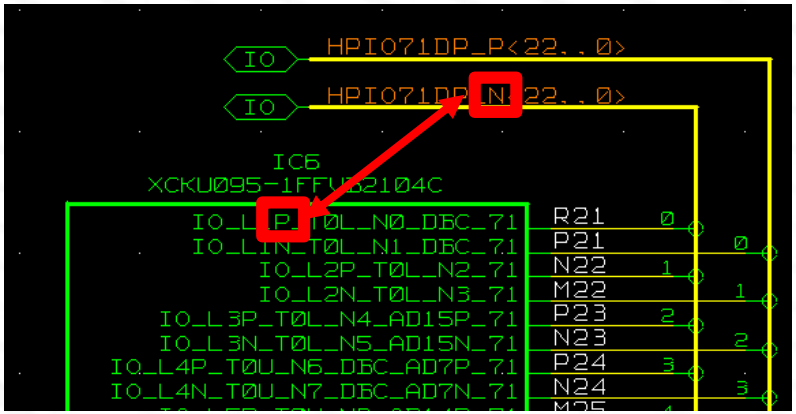


Schematics verification

- First step: reading schematics thoroughly
- 100+ schematic pages → very difficult to spot accidental swaps
  e.g.: P/N differential pair polarity inversions
- Created automated python tool to check:
  - FPGA components description from CERN cadence library
  - P/N differential pair polarity inversions
  - i2C SDA/SCL swaps
  - JTAG TCK, TDI, and TDO swaps



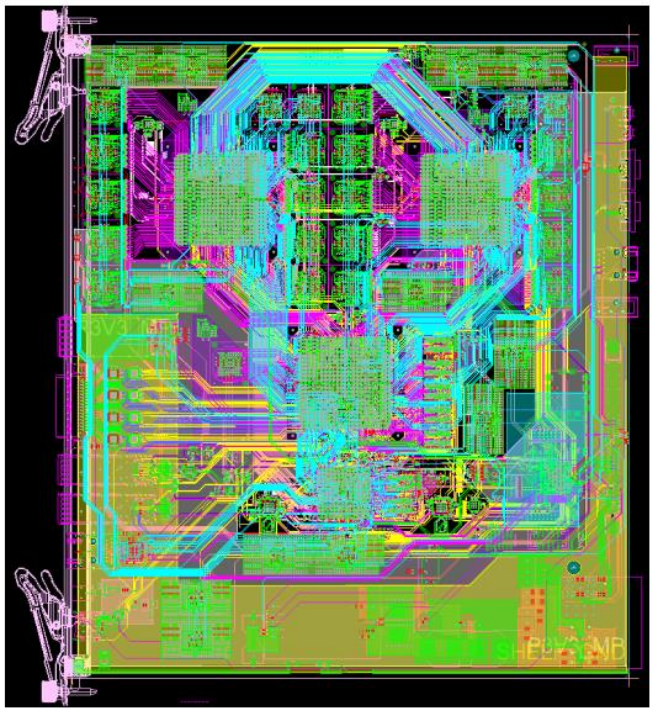
- Detected 140 P/N accidental differential pair polarity inversions
  - **We were able to fix them before PCB production**



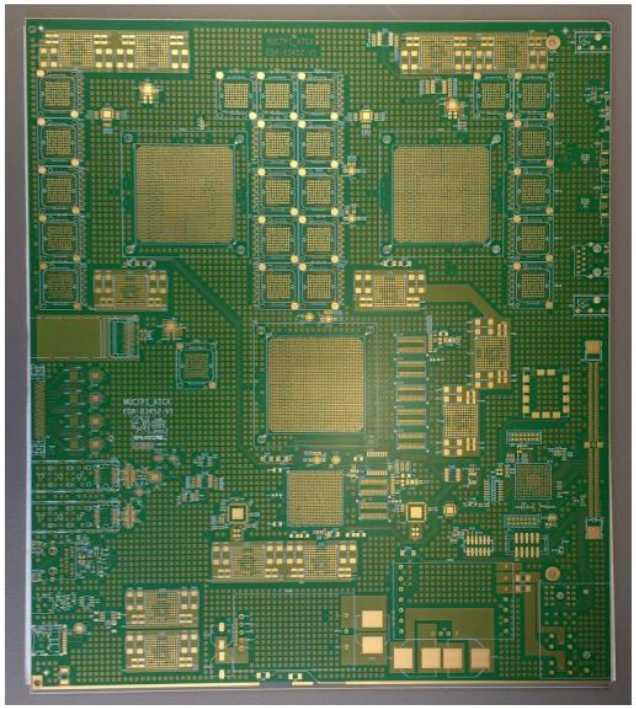
MUCTPI implementation

MUCTPI prototype

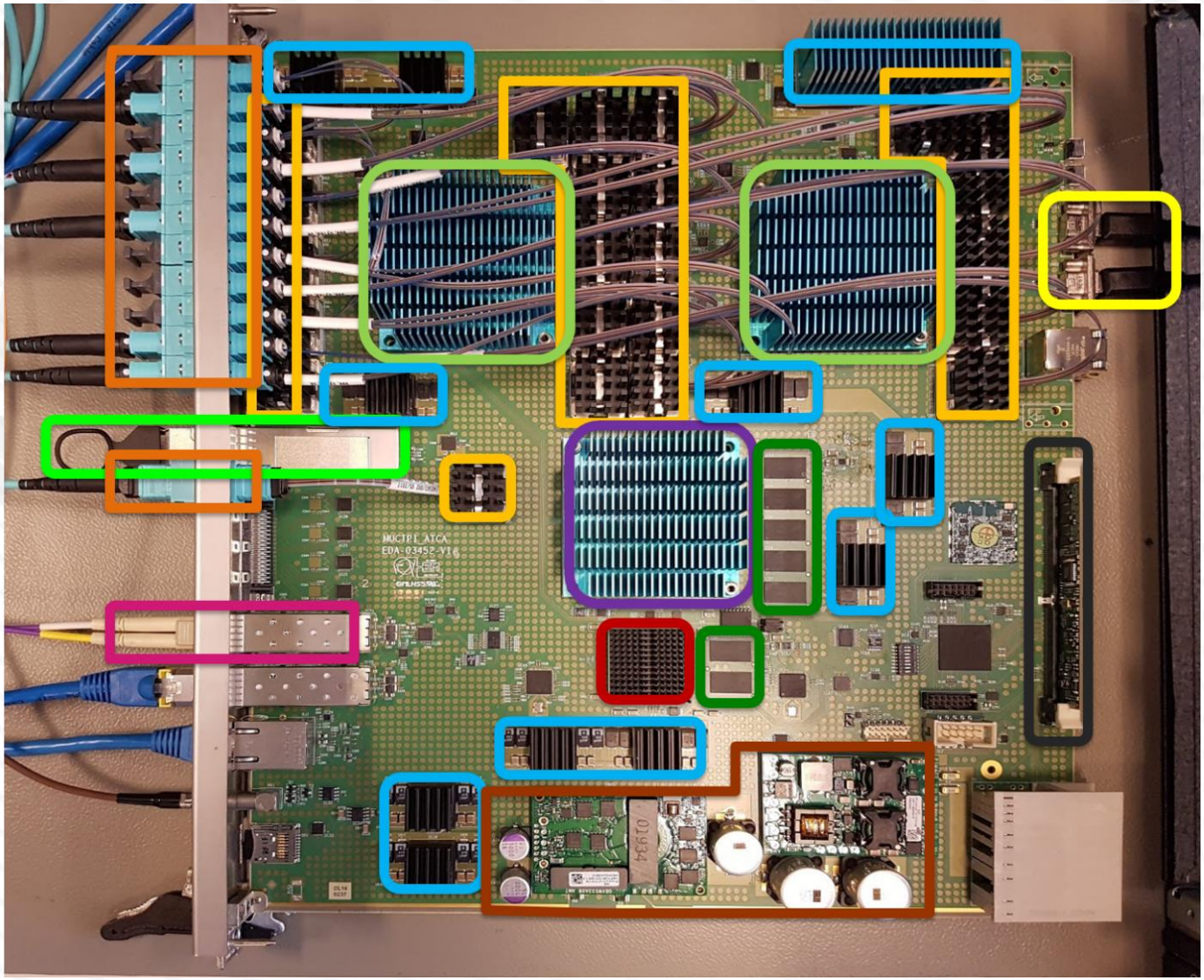
- Routing
  - 23 12-wayMiniPODs (18 Rx & 5 Tx)
  - 3 x 2104-pin Ultrascale FPGAs
  - ~330 MGT pairs (6.4 to 12.8 Gb/s)
  - ~240 LVDS pairs (1.28 Gb/s)



- PCB
  - 22 layer PCB
  - Megtron-6 low loss material
  - Blind vias for high-speed track layers



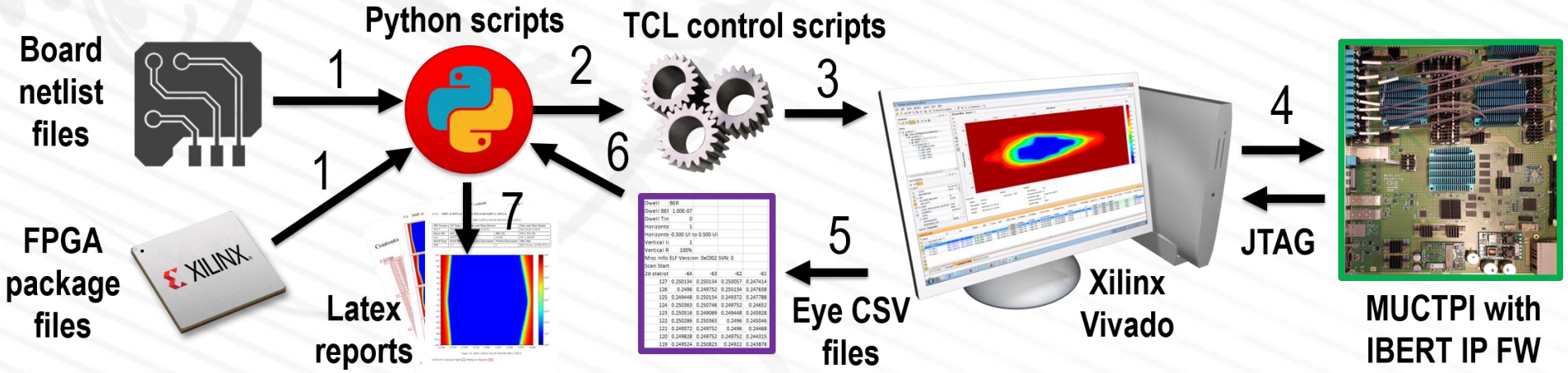
MUCTPI prototype



Firmware & testing

Testing of MGT serial links

- Swap between MGT channels and polarity inversions were allowed
- Firmware: Xilinx GTH & GTY IBERT IP used
- Software: Python scripts were generated to:
  - **Extract interconnectivity information from the back-annoted board design**
  - **Automate the interconnection between links in Vivado**
  - **Polarity configuration**
  - **Test running in Vivado**
  - **Eye-diagram compilation for all the links running at 6.4, 9.6, and 12.8 Gb/s**



10 days error-free PRBS-31 data transfer over 116 serial links running at 12.8 Gb/s, BER < 10^{-15} (95 % confidence level)

Report examples

Contents

1.1 MSP_A TX1 MSP_C RX16 Minipod Loopback

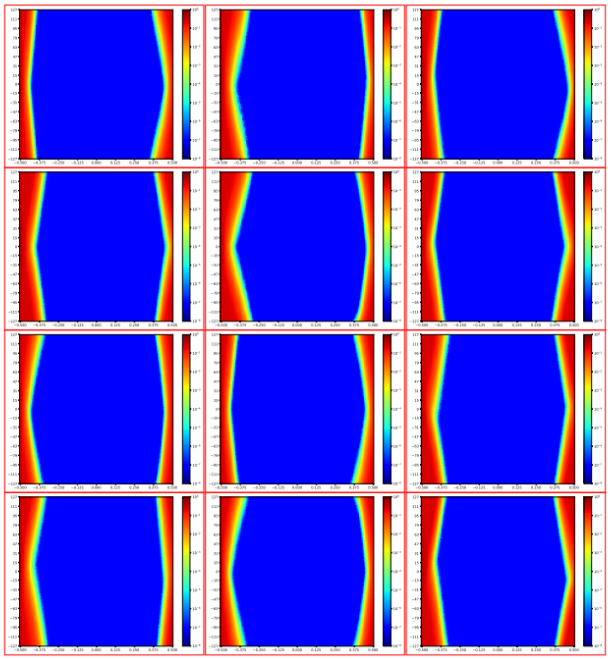


Figure 1.1: MSP_A TX1 MSP_C RX16 Minipod Loopback

A cross-reference to Figure 1.1. Sibling eye diagrams: 12.8.
Next summary Figure 1.14.

1.1.1 MSP_A_FPGA-TX1-00–RX16-00-MSP_C_FPGA

Table 1.1: MSP_A_FPGA-TX1-00–RX16-00-MSP_C_FPGA

SW Version	GT Type	Date and Time Started		Date and Time Ended	
2017.2	UltraScale GTH	2017-Jul-26 14:55:42		2017-Jul-26 14:56:53	
Reset RX	**OA**	**HO**	**HO (%)**	**VO**	**VO (%)**
true	25873	109	84.50%	255	100.00%
Dwell Type	**Dwell BER**	**Horizontal Increment**	**Vertical Increment**	**Misc Info**	
BER	1e-7	1	1	ELF Version: 0xC002 SVN: 0	

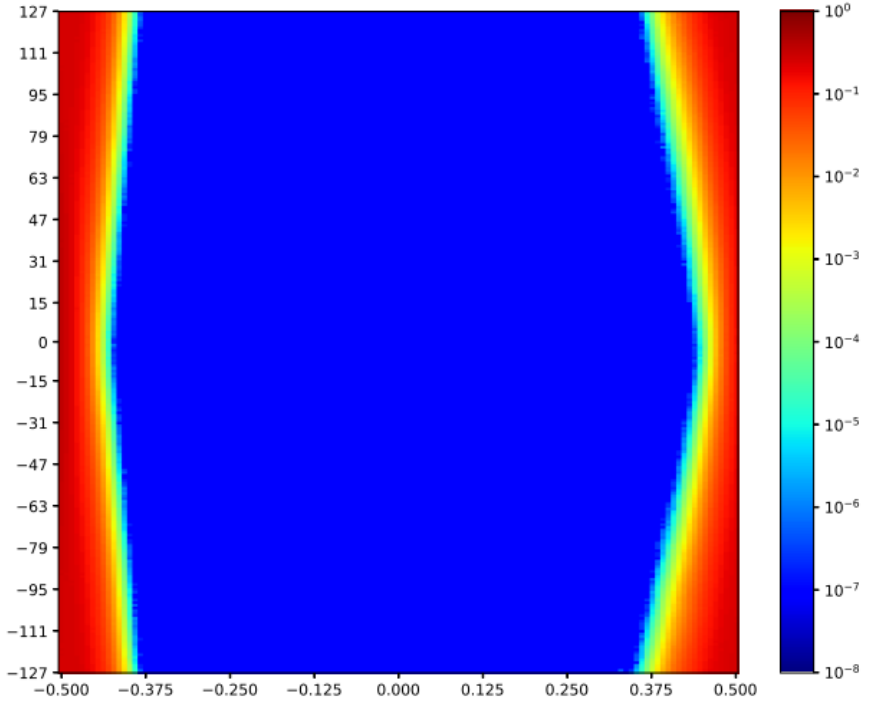
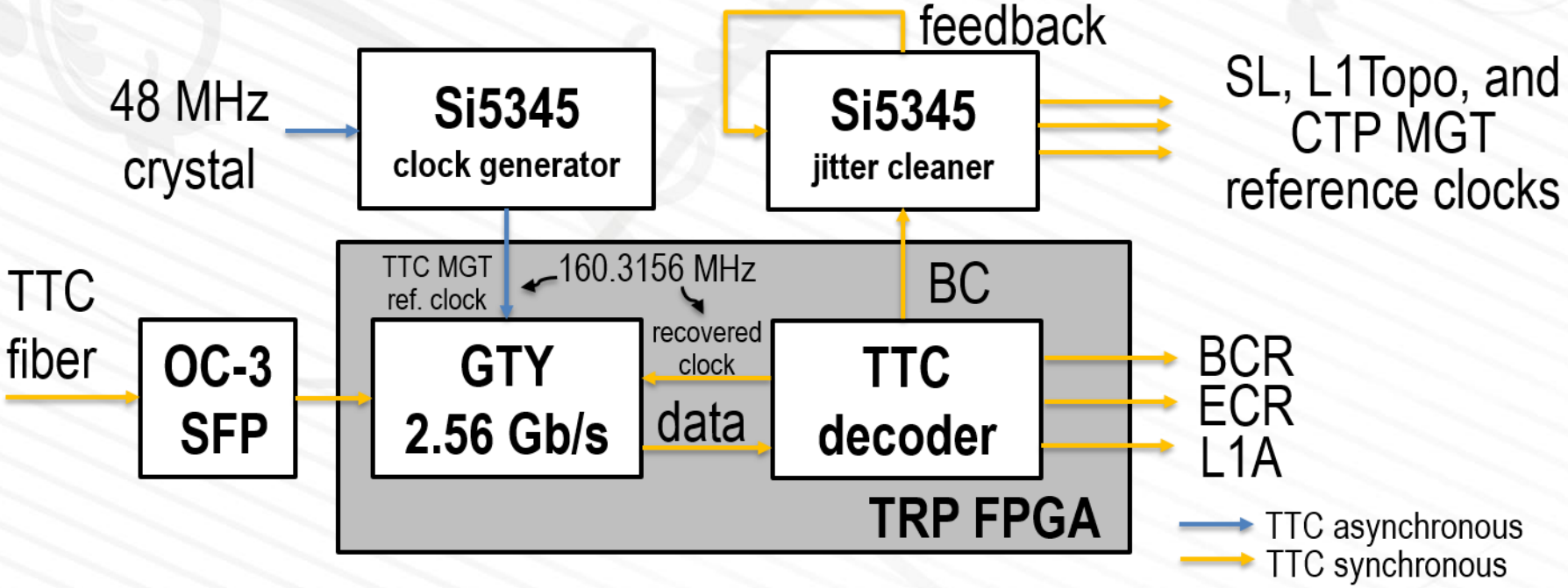


Figure 1.2: MSP_A_FPGA-TX1-00–RX16-00-MSP_C_FPGA

Call back to summary Figure 1.1. Sibling eye diagrams: 12.8.

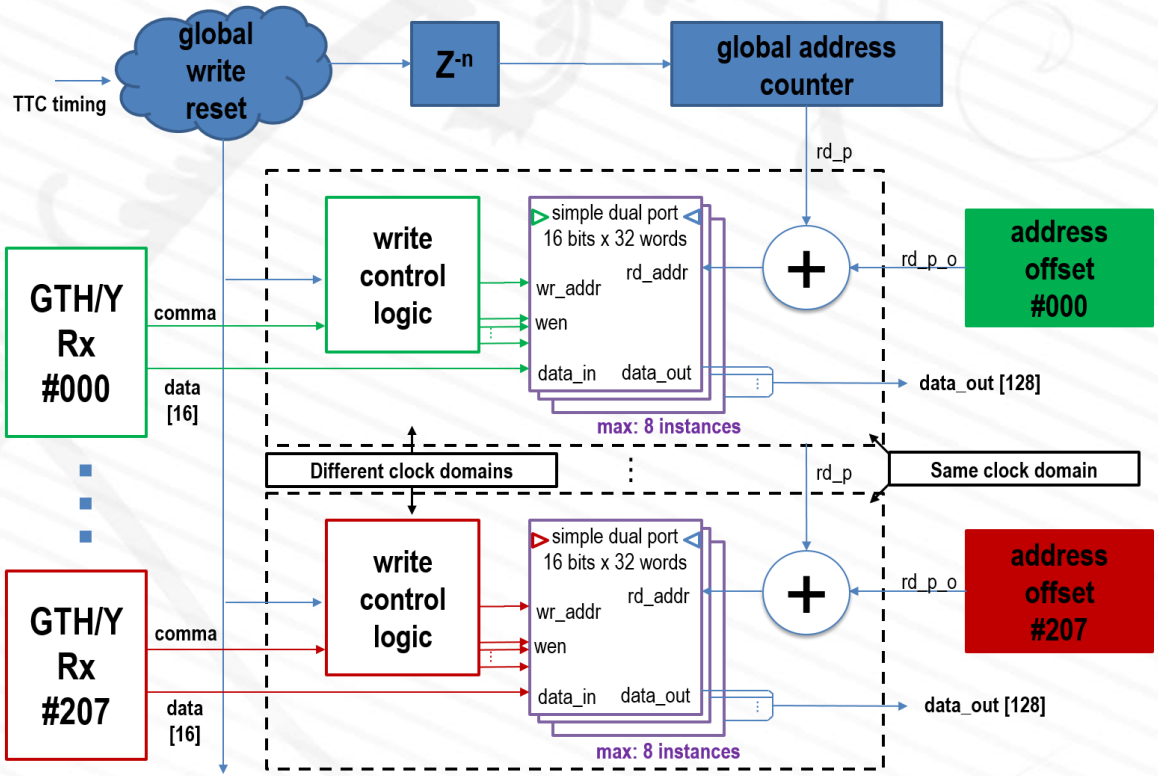
Timing, Trigger & Control (TTC) recovery

- Circuit based on external CDR ADN2814 will be used to recover the TTC
  - The same circuit as TTC FMC, widely used and tested at CMS
- An alternative was tested: General purpose transceiver and user logic to receive TTC (replaces TTC FMC)
  - OC-3 SFP transceiver module (optical → electrical transmission)
  - 2.56 Gb/s GTY oversamples 160 Mb/s TTC encoded signal
  - GTY outputs 160 MHz TTC recovered clock
  - User logic aligns recovered clock & data, decodes TTC, and generates 40 MHz
  - Jitter cleaner cleans TTC clock and generates required MGT reference clocks



SL input Synchronization & Alignment

- Has to compensate the input phase-skew
- Align the signals in multiples of the bunch-crossing period of 25 ns
- Write control logic detects BC frame boundaries
- Dual port memories transfer all 208 inputs from their respective clock domains into a single clock domain for combined data processing
- It can cope with phase variation of the received data and non-deterministic data transfer latency from FPGA transceivers by monitoring received data timing and setting logic delays in the write control logic
- Reduced version already tested with barrel and end-cap SL prototypes
- Complete version tested with 72 links running concurrently
- Will be used for connection tests with end-cap SL prototypes



Test results

- 3 boards were assembled with three different configurations:
  - 1) Zynq & Power supply system
    - Verify zynq and power supply
  - 2) Prototype without high-end FPGAs (MSP & TRP)
    - Used for software development and test pattern generation
  - 3) Fully assembled prototype

- Fully tested board results:
  - Power supply and cooling: **OK**
  - Zynq SoC and internal/external interfaces: **OK**
  - Board infrastructure monitoring (power supply, temperatures, optical modules, etc.): **OK**
  - High-speed serial optical input/output links: **OK**
    - Wide eye opening for SL inputs running at 6.4 Gb/s (~75 % of bit period)
      - Compatibility for Phase-II operation at 9.6 & 12.8 Gb/s was also confirmed
  - On-board high-speed serial links: **OK**
  - Clock distribution and TTC decoding and clock recovery: **OK**
  - LVDS links between FPGAs: **OK**
  - ATCA infrastructure and IPMC: **OK**
  - TRP FPGA DDR memory: **OK**

Summary & next steps

Summary & next steps

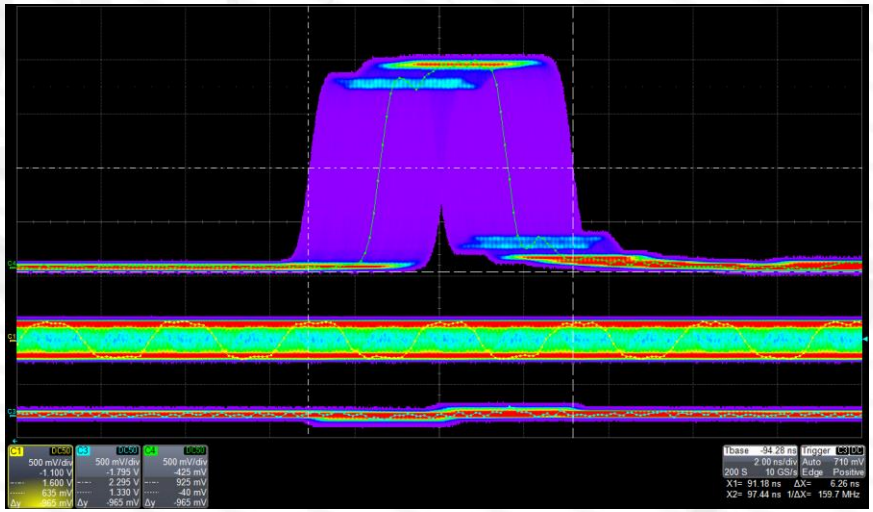
- Upgraded MUCTPI prototype tested successfully
- Tests without problems or difficulties thanks to extensive preparatory work using development kits
- Infrastructure firmware and board testing firmware ready
- Embedded Linux and application software running on Zynq working

- Next steps:
  - Fix remaining issues in PCB and launch a second prototype
    - pin-compatible Ultrascale+ FPGAs for the MSP
  - Connection tests with TGC sector logic and RPC interface board
  - Functional/algorithm firmware development
    - VU160 -> VU9P (+25% logic resources, 3x on-chip memory)
  - Expected in Q1’18

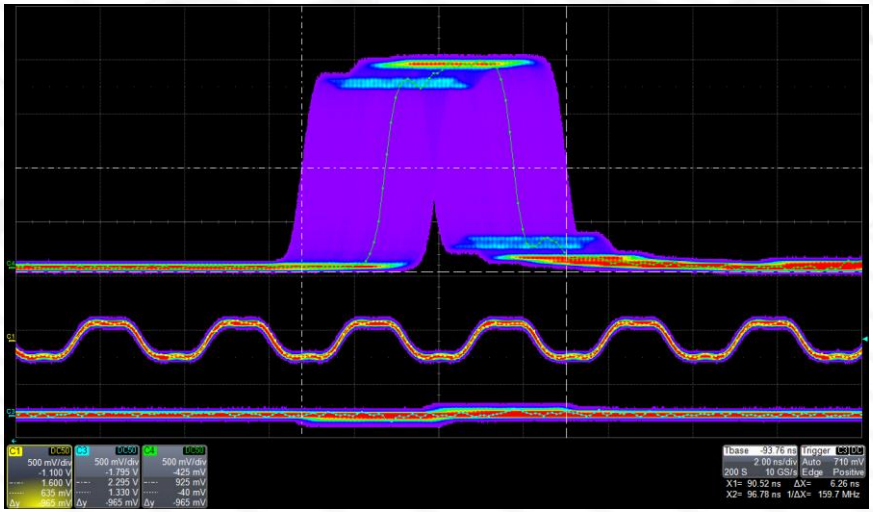
Back-up slides

Latency uncertainty reduced to 1 RXUSRCLK period

TX & RX buffer bypass **ON**
TXOUTCLK = TXPROGDIVCLK



TX & RX buffer bypass **ON**
TXOUTCLK = TXPLLREFCLK_DIV1



C1 = MGTREFCLK (320 MHz), C3 = TxTriggerPulse (Triggering), C4 = RxTriggerPulse
The TX latency becomes deterministic when the programmable divider is not used to generate the TXOUTCLK (paralell clock). Hence, the reference clock has to be the same as the desired TXOUTCLK. For the plots above, the transceiver reset is asserted every 3s, and thousands of waveforms are captured. The TX-RX latency has a skew of 3.125 ns in both cases, however the latency from the assertion of the REFCLK to the data being received in the RX can be longer when TXPROGDIVCLK is used, due to the variation of the latency in the TX.