A BMT Layer-1 technology demonstrator card and Links

Hardware and Firmware

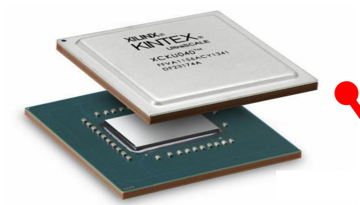
S. Mallios[1], K. Adamidis[1], G. Bestintzanos[1], C. Foudas[1], G. Karathanasis[2], P. Katsoulis[1], N. Manthos[1], I. Papadopoulos[1], S. Sotiropoulos[3], P. Sphicas[2]

(1) University of Ioannina
(2) University of Athens
(3) Institute of Accelerating Systems and Applications (IASA), Athens

Hardware

BMT Layer-1 demonstrator

Board Overview

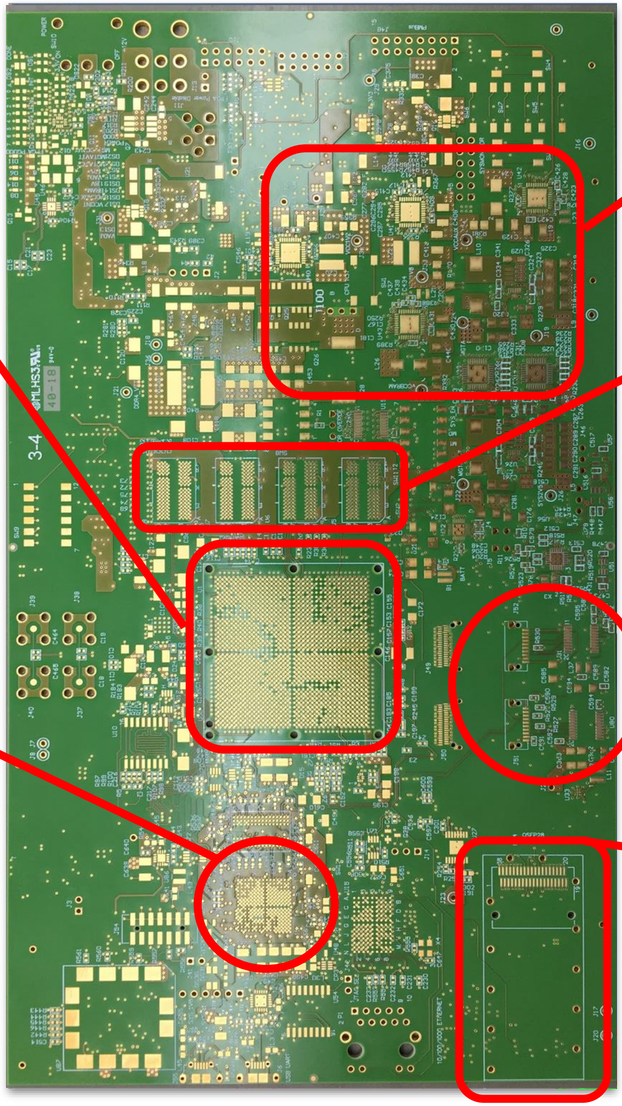


FPGA XCKU040

- 20 x [16 Gb/s GTH transceivers]
- Speed grade : -2
- Footprint : FFVA1156 (BGA surface-mount packaging)

ZYNQ XC7Z010
System-on-Chip (SoC)
for slow control over I2C

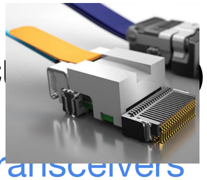
- Clock configuration (si570, Si5328 & Si5338)
- PMBus power supply monitoring
- Kintex FPGA SYSMON monitoring



Power Modules

2 GB DDR4 (4x512KB)

Samtec Firefly (EC



- 12x optical transceivers
- 16 Gb/s per channel

QSFP28 FTLC9551REPM

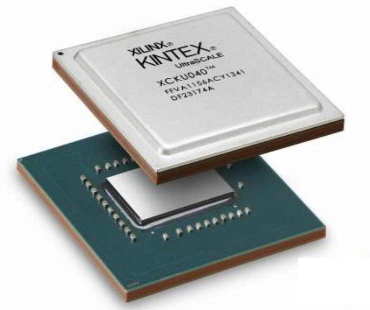
- 4x25Gb/s transmitter
- Max link length 100m on OM4 Fiber (MMF)
- Hot-pluggable QSFP28 form factor

Kindex Ultrascale FPGA

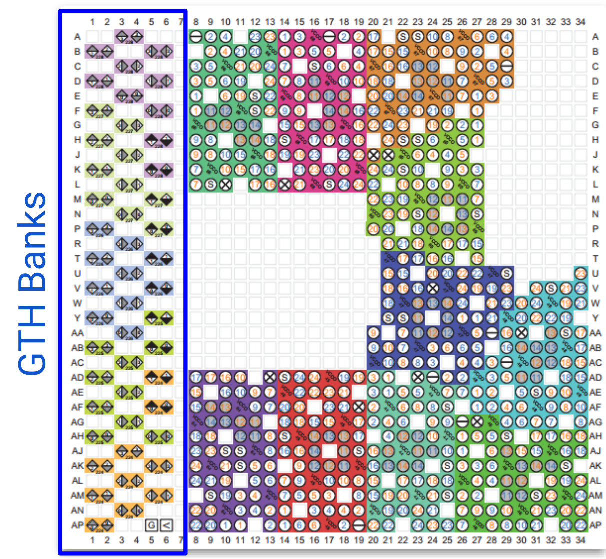
XCKU040-2FFVA1156E

- ❖ Kintex UltraScale XCKU040 FPGA
- ❖ Best Price/Performance/Watt at 20 nm
- ❖ 20 x [16 Gb/s backplane - capable transceivers]
- ❖ Temperature grade : Extended (0°C - 100°C)
- ❖ Speed grade : -2
- ❖ Footprint : FFVA1156 (BGA surface-mount packaging)

	XCKU040
System Logic Cells (K)	530
DSP Slices	1,920
Block RAM (Mb)	21.1
16.3Gb/s Transceivers	20
I/O Pins	520



Kintex UltraScale XCKU040 - 2FFVA1156E FPGA



XCKU040 I/O Bank Diagram

Advantages :

- + Good price/performance ratio (~ 2 KChf)
- + 16 Gb/s GTH MGTs (sufficient for the BMT Layer-1)
- + Used on the KCU105 development xilinx board
- + Big enough to run Barrel Muon Track Finder kalman algorithm

Disadvantages :

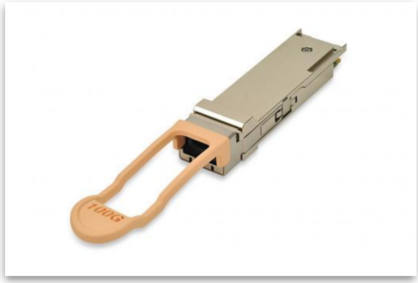
- - Relatively small to run more complex algorithms
- - Absence of GTY MGTs (up to 32.75Gb/s)

Optics Overview

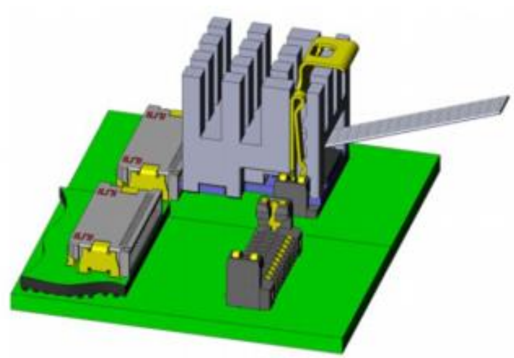
High Speed Optical Links

- 16 optical links - operating in excess of 16 Gbps
  - Twelve (12) optical FireFly links
    - optical flyover assembly, placed next to the FPGA.
    - 12 separate transmitter (TX) and receiver (RX) optical modules,
      - joined in a "Y" configuration and
      - terminate to a single 24 fiber MPO connector.
    - connectors placed mid-board and the data "fly" over the PCB, allowing easier routing.
  - Four (4) QSFP28 links transceiver module.
    - Part : Finisar FTLC9551REPM
    - Hot-pluggable QSFP28 form factor
    - Supports 103.1 Gbps of aggregate bit rate.
    - Rate is limited by the MGT's maximum speed to 64 Gbps.

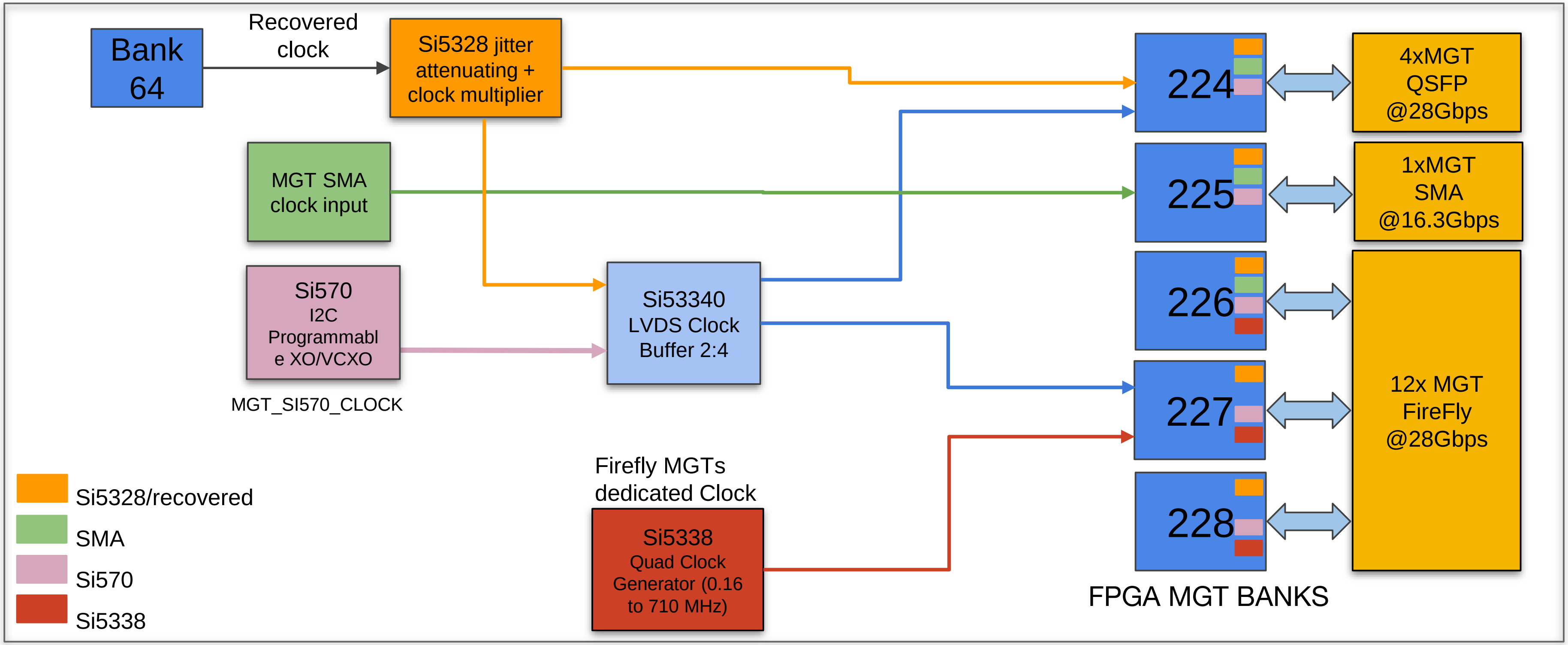
Samtec Firefly ECUO-Y12-16



Finisar FTLC9551REPM 100G QSFP28 Optical Transceiver

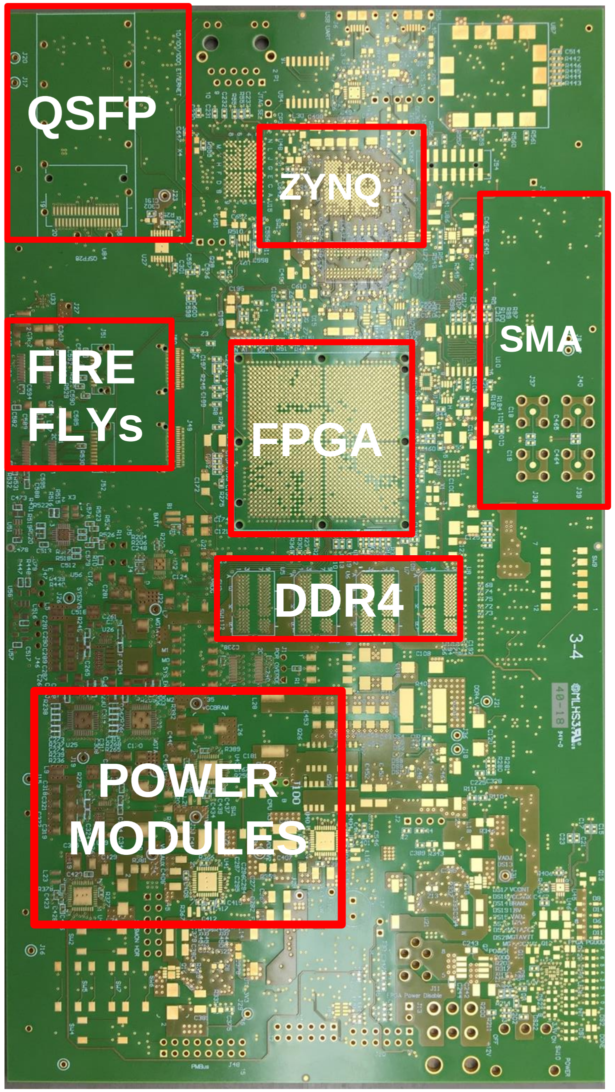


Miniature On-board Optical FireFly Micro Flyover System (source: Samtec)



Clock distribution tree

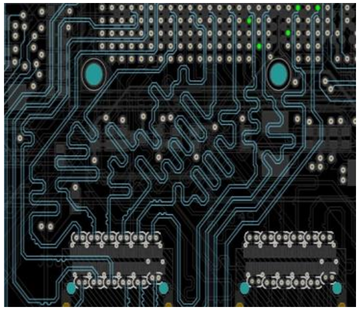
- 3 low-jitter programmable clock sources.
  - Dedicated, low jitter, quad clock generator (Si5338) for the high speed optical links.
  - Low-jitter frequency generator (Si570) is connected to the QSFP28 transceivers
  - Jitter attenuator (Si5328B) for the recovered clock.
- A fixed frequency clock source for reset and initialization FSMs
- SMA external clock input
- All programmable clocks are accessed through a dedicated I^{2}C bus.



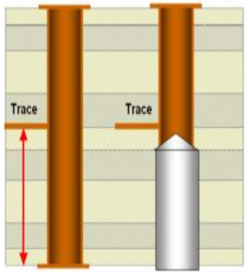
PCB Top Layer

PCB

- 16-layer stackup (ground-plane has been placed between each layer containing high-speed traces)
- Megtron-6 (Panasonic) substrate
  - excellent high-frequency performance and impedance properties.
- Backdrilling
- Serpentine routing to match route lengths of high speed differential pairs



Serpentine routing



Backdrilling

- Status of the PCB fabrication :
  - Received the bare-boards on the 20.10.2018
  - Sent for assembly
  - The assembly should take 2-3 weeks.
  - We expect to have two boars ready in the 2nd week in December

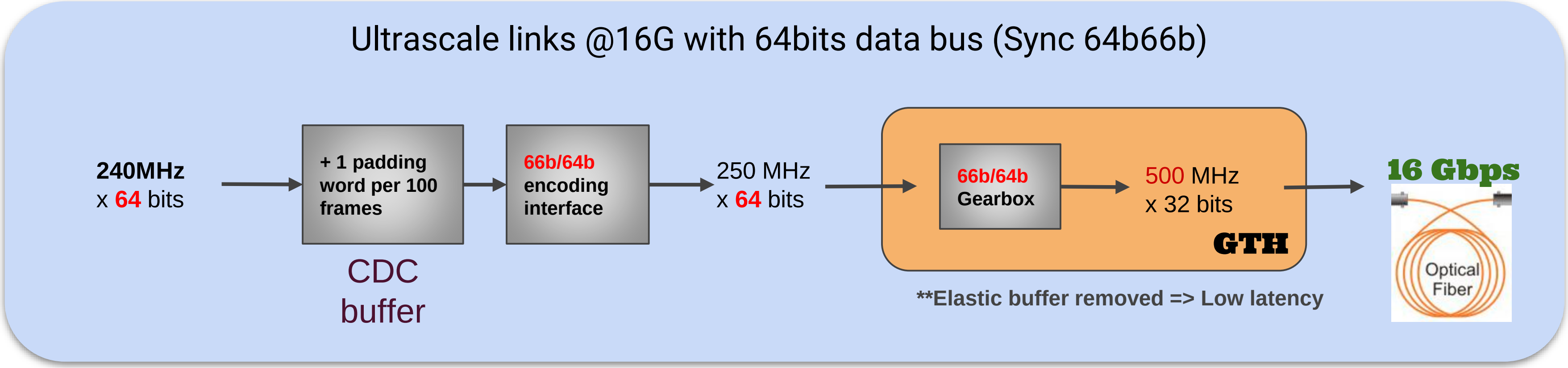
- 4 Bare Boards have been delivered to the assembly company.

- Two assembled boards are expected to arrive at CERN mid December 2018

- Post - assembly planning :
  - Perform basic hardware test to ensure that the individual devices and buses/interconnects are operational (power, clocks, and basic functional connectivity)
  - Optical/serial link validation (BER tests, board-to-board connectivity etc)
  - Implement advanced algorithms (i.e. Kalman barrel muon algorithm)

Firmware

16 Gbps asynchronous links

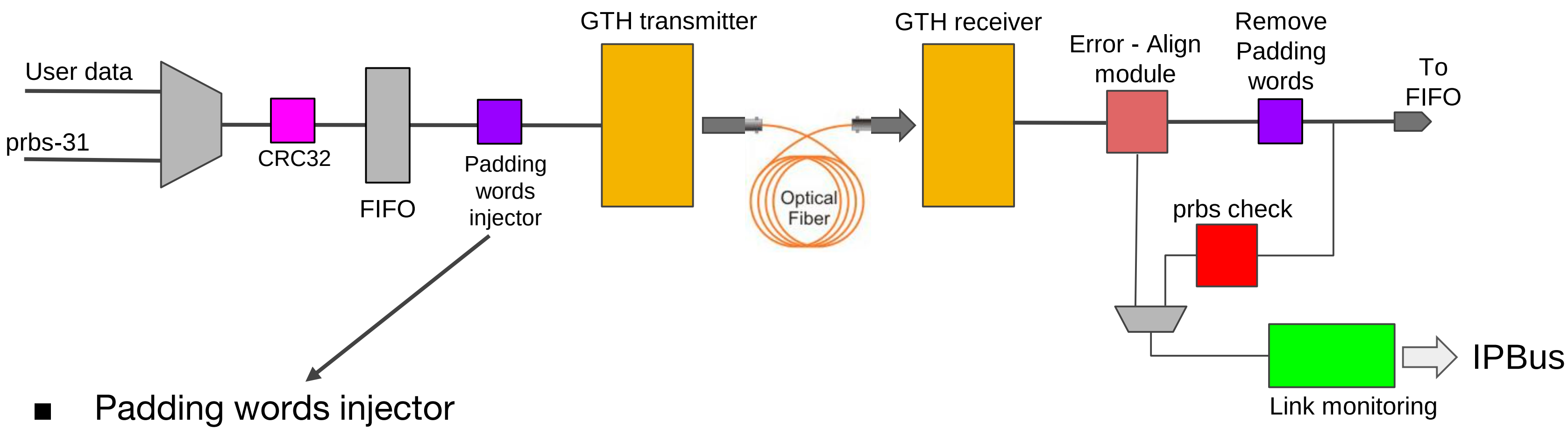


- Overview
  - Asynchronous design : algorithmic logic run in different speed than the link clock.
  - FIFOs to cross between clock domains.
  - Inject padding word to compensate the freq difference.
  - Tested with 240Mhz algo clock and 250 MHz link clock.

- Encoding
  - 66b64b with synchronous gearbox.
  - 2-bit header (coding bits).
  - 3.125% overhead.

- Latency
  - Elastic Buffer bypassed.
  - GTH latency 9 CLKs.
  - Total latency 21 CLKS (including clock domain crossing FIFOs).

Protocol - Overview



- Padding words injector
  - Controls the header and the data/control words. It
  - Checks the incoming valid bit / padding flag / crc flag and injects the corresponding words.
    - Valid bit = 0 & padding word = 0 → header = b“10” (can use for alignment)
→ data = IDLE
    - Valid bit = 0 & padding word = 1 → header = b“10” (used for alignment)
→ data = Padding Word
    - Valid bit = 1 & padding word = 0 → header = b“01” (cannot use for alignment)
→ data = user data
    - Valid bit = 1 & padding word = 1 → header = b“10” (used for alignment)
→ data = Padding Word
- We check the data quality when the header is b“10”
  - Single errors (soft) are monitored
  - Continues errors (hard) trigger a re-alignment procedure
- Illegal header values considered link errors
- Align_time(max) = 100(clks/pad)*63bit *8 = 50400 clks or 201600 ns or ~200us

Protocol - Packets

	Valid Bit	Header	CODE (8-bit)	PAYLOAD (56-bit)		
	0	10	0x55	56-bit IDLE word (??)	IDLE	Algo CLK domain
Start of Packet	1	01	64-bit DATA		DATA	
	1	01	64-bit DATA		DATA	
	1	01	64-bit DATA		DATA	
Fifo empty	X	10	0x78	56-bit Padding word	CDR	Link CLK domain
	1	01	64-bit DATA		DATA	
	1	01	64-bit DATA		DATA	
	.	..	...			
	.	..	...			
End of packet	1	01	64-bit DATA		DATA	
CRC word	0	10	0x99	0x000000 | 32-bit CRC	CRC	Algo CLK domain
	0	10	0x55	56-bit IDLE word	IDLE	Algo CLK domain
	0	10	0x55	56-bit IDLE word	IDLE	Algo CLK domain

Packet builder

- Based on the Aurora 66b/64b protocol
  - All transmissions performed using 64-bit blocks.
  - Aurora offers ten types of blocks that can be transmitted through an Aurora channel.
- We use 4 types of BLOCKs
  - User data block (normal data)
  - Padding block (CDR)
  - CRC block
  - IDLEs block

The functionality of the links was extensively tested

- using XILINX KCU105 ultrascale development board. For the tests an
- FMC loopback card was used to implement a copper loopback quad link.

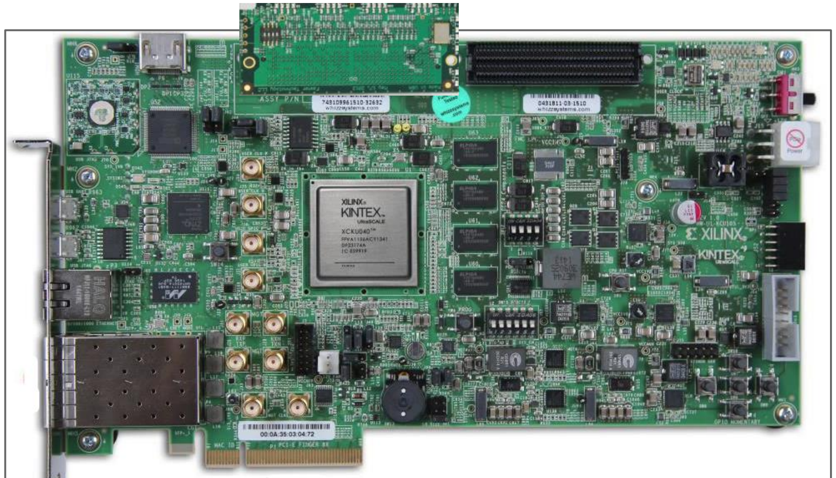
Latency

- The GTH latency~ 9 CLKs adding the 2
- FIFO CDC latency ~6 CLKs (crossing between clock domains)
- Total link latency add up to 23 CLKs (~3.5 BXs @250MHz).

Bit Error Rate Tests :

- Sending PRBS-31 data over an FMC copper loopback card.
- Run for more than 72 hours
- No errors resulting in a BER < $2x10^{-16}$

FMC XM107 Loopback Card



Xilinx KCU105 demo board

➞ IPBus integration work in progress (**K.Adamidis**)

- Slow control tested (resets, link status, error monitoring)
- Integrate with spy buffers to inject/read patterns

Ipbus integration over 16g asynchronous links

- The 16G links firmware is integrated with the Ipbus protocol to provide control (resets) and monitoring (link status indicators) over the links.

- IPbus Firmware was modified to operate over RJ-45 Ethernet cable and through the parts available on the demonstrator board:
  - A Marvell Alaska PHY device (M88E1111)
  - An RJ-45 Halo HFJ11-1G01E-L12RL connector

  The usage of the Gigabit Transceiver Marvell PHY also sets free one of the FPGAs SFP transceiver.

- Additional registers and memory blocks interaction can be added at any time for further testing, control and monitoring.

```
[kadamidi@psepc19 uhal_interface]$ python links_16g_control.py reset.all
Reseting reset.all Register
...
---------------------------
| Link Status:          0 |
| Link Down Latched:    0 |
| Lind Initialize Done: 0 |
---------------------------
[kadamidi@psepc19 uhal_interface]$
```

```
[kadamidi@psepc19 uhal_interface]$ python links_16g_control.py read_stats
---------------------------
| Link Status:          1 |
| Link Down Latched:    1 |
| Lind Initialize Done: 1 |
---------------------------
[kadamidi@psepc19 uhal_interface]$
```

```
[kadamidi@psepc19 uhal_interface]$ python links_16g_control.py reset.error
Reseting reset.error Register
...
---------------------------
| Link Status:          1 |
| Link Down Latched:    0 |
| Lind Initialize Done: 1 |
---------------------------
[kadamidi@psepc19 uhal_interface]$
```

Kosmas Adamidis

Hardware Summary

What is done

- ✓ Board Schematics completed
- ✓ PCB layout completed
- ✓ PCB fabrication completed

Scheduled

- ► Board assembly (ongoing)
- ► Testing (links, algorithms etc)

Firmware Summary

What is done

- ✓ 16Gb/s asynchronous links using 64/66b encoding
- ✓ Auto link alignment
- ✓ Slow control with IPBus over Ethernet

Scheduled

- ► Test more data bus width - link clock combinations
- ► Implement 16/25 Gbps links with GTYs