Phase-I FELIX Design Review

hardware development

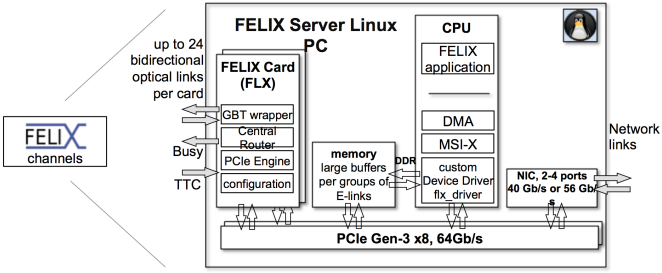
Kai Chen
On behalf of the FELIX group

FELIX Design Review
November 11, 2016

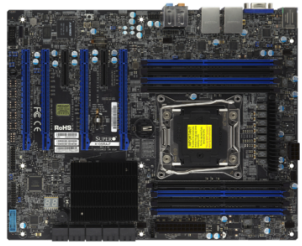
Outline

1. FELIX Hardware
2. Prototype design of the FELIX I/O module
3. Summary

FELIX Hardware

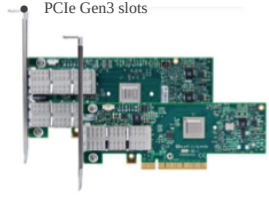


- Server Linux PC
- Up to two PCIe interface cards with Xilinx Ultrascale FPGA, depending on bandwidth needed (for two cards: using 2× PCIe slots Gen3 ×8 lanes, leaving enough lanes for the NIC(s))
- NIC, 40 or 100 Gb/s Ethernet interfacing or InfiniBand



SuperMicro X10SRA-F used for development

- Broadwell CPU, e.g. E5-1650V4, 3.6 GHz
- PCIe Gen3 slots

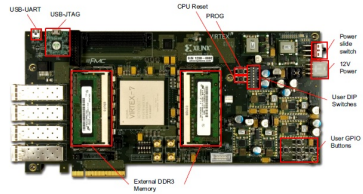


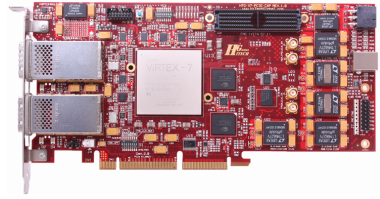
Mellanox ConnectX-3 VPI

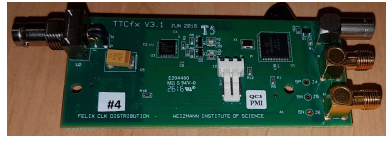
- 2× FDR/QDR Infiniband
- 2× 10/40 GbE

FLX-709 & FLX-710

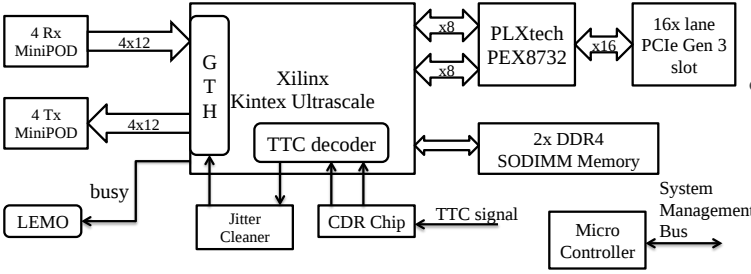
- FLX-709: Xilinx VC-709 evaluation board
  - 4 channels
  - on-board jitter cleaner doesn't meet requirements
    - SI5324 doesn't support 0-delay mode
- FLX-710: Hitech Global HTG-710 evaluation board
  - 24 channels
  - only support using on-board oscillator
- TTCfx card is designed to interface TTC system and clean the clock
  - FLX-709: TTCfx outputs clean clock to transceivers: via SMA cable
  - FLX-710: the hardware must be changed
- These boards can be used to develop the FELIX functions when custom boards are not yet available
  - FLX-709: targets detector and trigger system test setups
  - FLX-710: development without interfacing TTC system (mostly for software development nowadays)







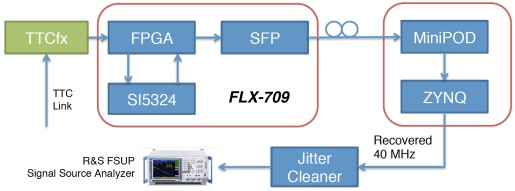
Design of custom PCIe card



- A PCIe card was designed for LAr LTDB (LAr Trigger Digitizer Board) test setup at BNL

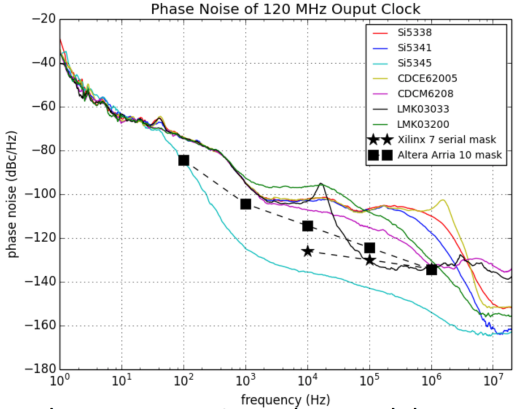
- Basic functions:
  - PCIe Gen 3 $\times$16 lane
  - 48 bidirectional optical links up to 14 Gb/s
  - 2$\times$ DDR4 SODIMM connectors support: capacity up to 16 GB & 2.1 GT/s
  - with circuits to interface TTC system; with on-board jitter cleaner
  - FPGA resources (logic cells) are about twice of the FLX-709 or FLX-710
- Meets the FELIX requirements. A new function is added for FELIX:
  - micro-controller to support FPGA reprogramming, and firmware update
  - important for detector operation & maintenance
- It is the baseline choice of Phase-I FELIX prototype

Selection of the clock jitter cleaner



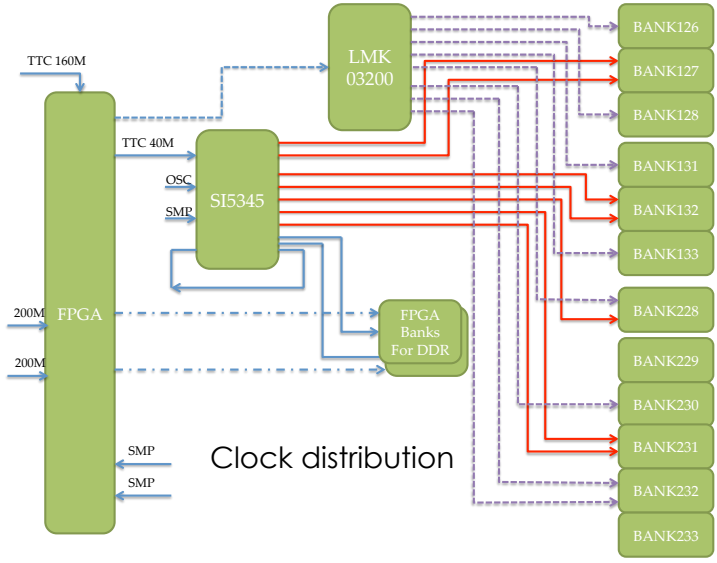
Device	SI5338	SI5345	SI5341
Jitter (ps)	8.58	0.09	6.39
Device	CDCM6208	LMK03200	LMK03033
Jitter (ps)	2.06	5.91	2.74
Device	CDCE62005		
Jitter (ps)	8.61		

The jitter from 10 kHz to 1 MHz



- This survey was originally done to choose a clock device for a FrontEnd board. It also provided input for the selection of jitter cleaner for FELIX.
- SI5345 showed the best performance.
  - it supports 0-delay mode.
  - 10 outputs.
  - meets requirement for transceivers in both of 7 series FPGA and Ultrascale FPGA.

Clock distribution for the FLX-711 V1P5

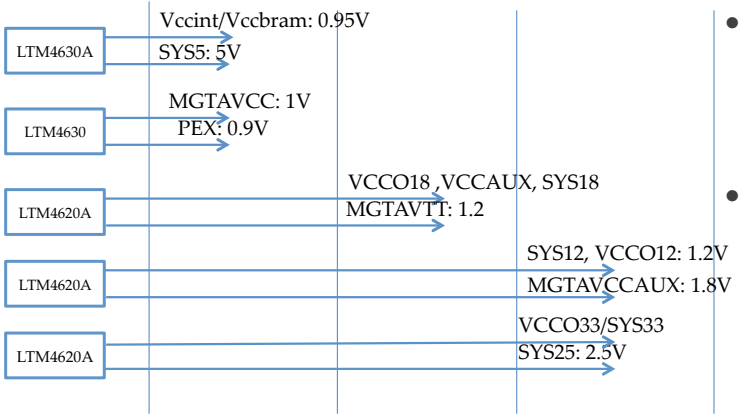


- LMK03200 is a backup for SI5345. Both of them support 0-delay mode.
- Each bank can use the reference clock from both of Si5345 & LMK03200.

Power supply for the FLX-711 V1P5

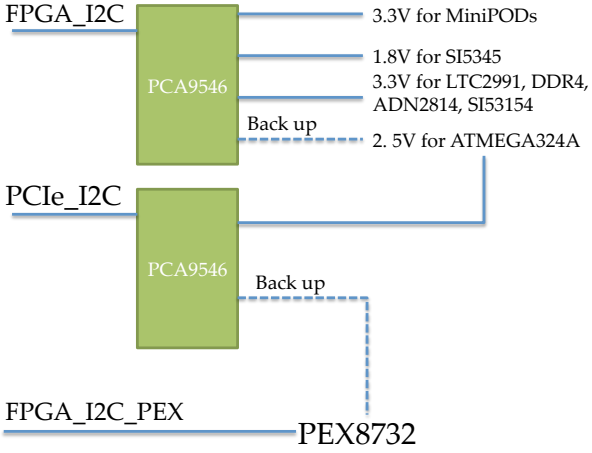
Three stages of power on sequence.

LTM4630A: 18/36A, in: 4.5-15, out: 0.6-5.3
LTM4630: 18/36A, in: 4.5-15, out: 0.6-1.8
LTM4620A: 13/26A, in: 4.5-16, out: 0.6-5.3



- Will use LTM4630A for all the 5× DC-DC power modules.
  - same price.
  - pin compatible.
  - new generation.
- Worst case power dissipation:
  - FPGA: about 37 W.
  - others: about 27 W.

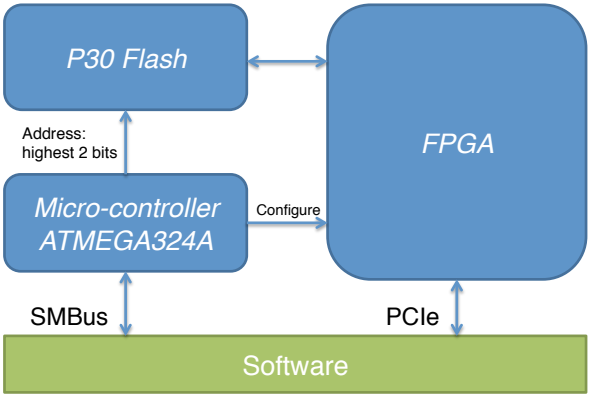
The use of I2C switch



Dev	Addr
PCA9546	1110000
MiniPODs	01100XX & 01011XX
SI53154	1101011
ADN2814	1000000
SI5345	11010XX
LTC2991	100100X
TCA6408	0100000
ATMEGA	XXXXXXX
PEX8732	0111000

- The use of I2C switches eases the design of firmware and software (compatibility with flx-i2c). One firmware module can handle all I2C slaves.
- Reduce the use of level translator.

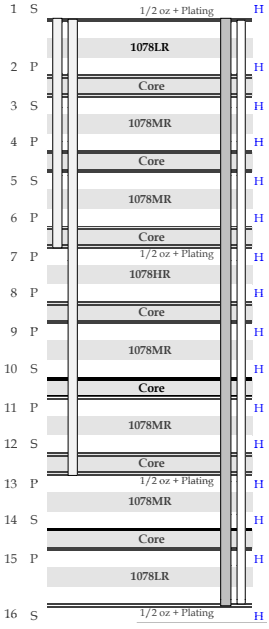
The use of micro-controller



- The flash can store 4 different bit files. It is selected by software via the micro-controller. A golden version can be saved in one of them.
- The software can reprogram the FPGA via micro-controller.
- The FPGA firmware can receive bit file from software via PCIe interface, and update the chosen flash partition.

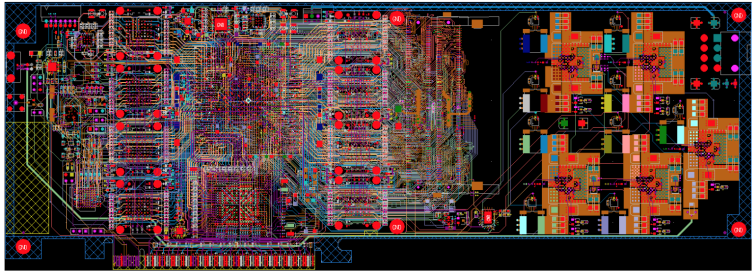
- *This design is originally from C-RORC (ATLAS RobinNP) board.*
- The program running in the micro-controller and script to control the FPGA programming are modified from the design by Heiko Engel of ALICE group.
- The flash programming via PCIe is being tested.

The PCB stackup

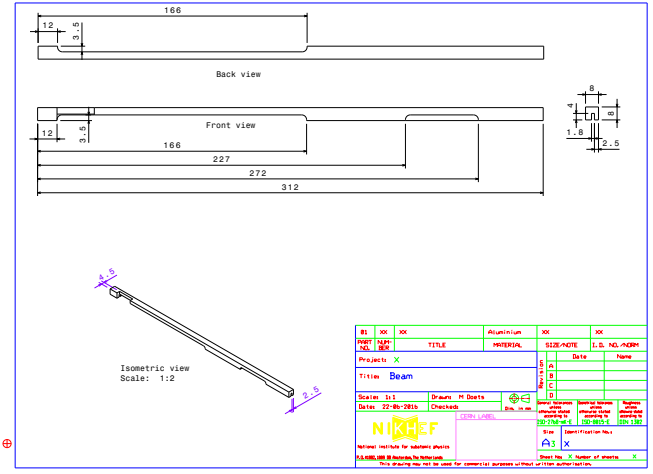


- The layout is complicated:
  - board thickness (1.57 mm) limits the number of layers.
  - board height requirement causes the traces to be very dense.
  - special impedance requirement for DDR4.

- Two kinds of blind vias are used. 3 sequential laminations are needed when producing the PCB.
  - one (layer 1-6) is for the MiniPODs.
  - one (layer 1-12) is for DDR4 traces.



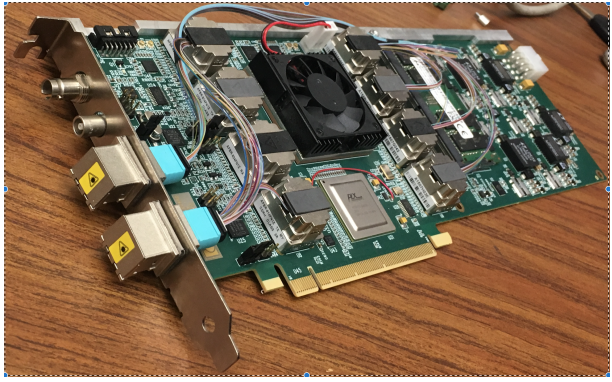
The mechanical design for the FLX-711



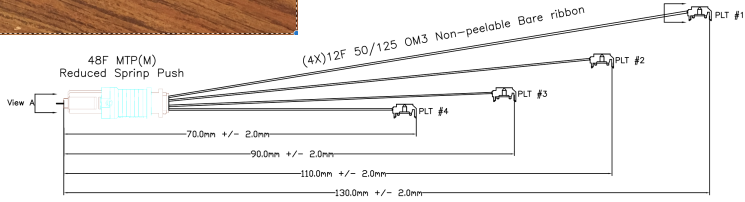
- Form factor of the card:
  - thickness: 1.57mm (62 mils).
  - height: 111.15 mm (4.376 inches)
  - length: 312.00 mm (12.283 inches)

- The stiffener bar for this long board.
  - designed by Nikhef
  - machined at BNL.
  - will be glued on the board.

The front bracket and fiber



- The front panel:
  - TTC fiber
  - LEMO connecter for BUSY signal.
  - two MTP couplers, in each:
    - 24 Tx channels
    - 24 Rx channels



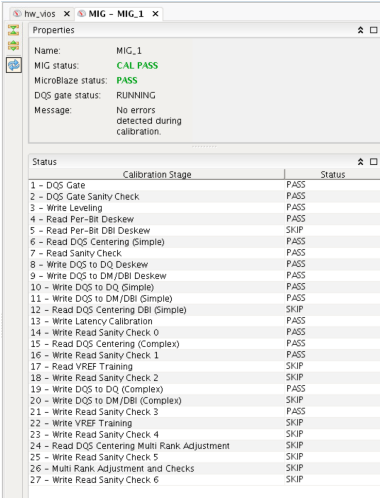
- Two harnesses are ordered, each has 4× 12-channel fibers with different lengths.

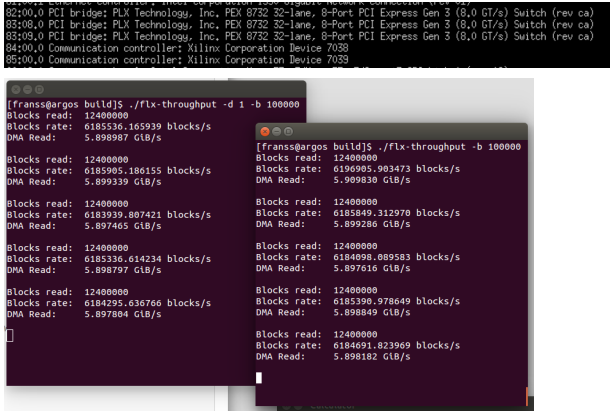
Minor changes from V1P0 to V1P5

- Some improvements are made to the first version board, after the testing of V1P0.
  - Mapping of the 48 transceivers:
    - make sure the transmitter and receiver in each transceiver has same sequence number in the Tx and Rx MiniPOD
    - make sure the 4 transceivers in each quad are connected to the same sub-group (channel 1-4, 5-8 or 9-12) in one MiniPOD.
  - Simplify the clock distribution design, SI5345 is used to replace the SI5338.
  - Power sensing of the DC-DC modules are used, to provide accurate VCCINT, MGTAVCC, MGTAVTT to the FPGA.
  - I2C switches are used.
- Bug fixes
  - Current capacity of the 1.8V power for the PCIe switch PEX8732.
  - Use dedicated clock pins for the DDR system clock.
  - Switch is added between flash and FPGA, since the special FPGA bank 0 is unusable by firmware. These pins will be connected to other bank, when we update the bit file in flash.
  - PCIe lane bit order in each quad.

Test results of FLX-711 prototype

Speed & throughput

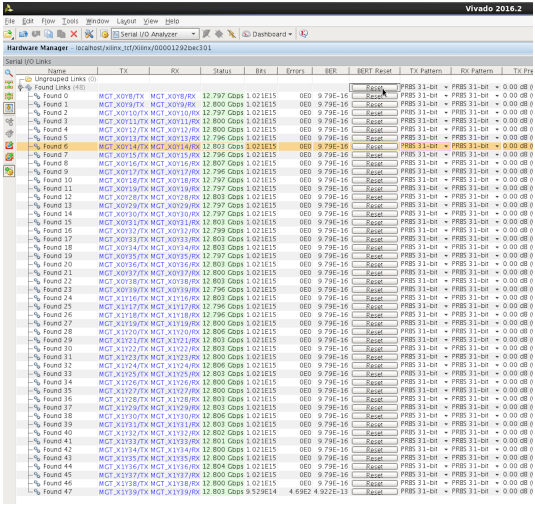


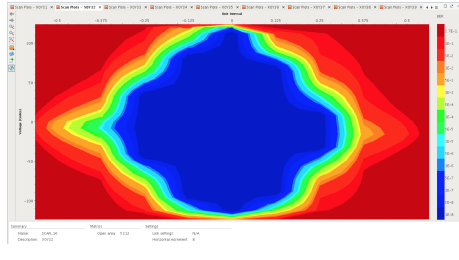


- The two DDR4 modules work well at a speed of 2.11 GT/s.
  - calibration is OK.
  - passed the 100 times of write and read checking.

- Two Xilinx PCIe endpoints are found by *lspci*.
- Reading of the two PCIe endpoints is done in parallel. In this example the firmware sends simple counter data.
- Total throughput is about 101.7 Gb/s.

IBERT test for all transceivers



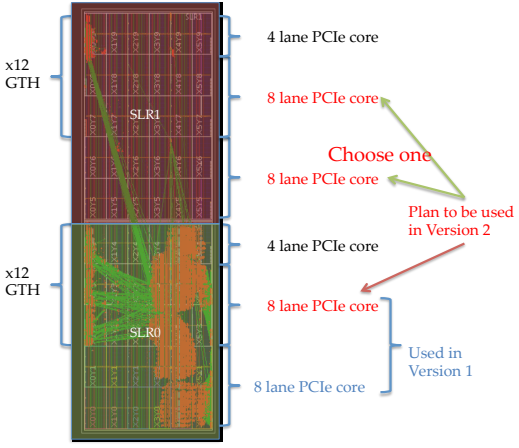


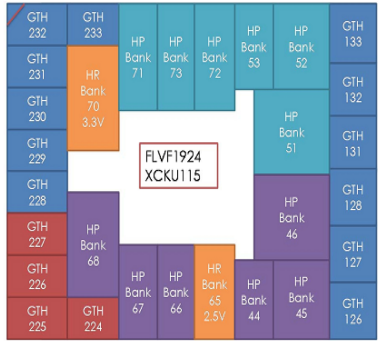
- IBERT testing is done for all the 48 links at 12.8 Gb/s.
  - BER < 1E-15.
  - the last link: RX_P pin on FPGA side is open due to assembly issue.
  - local clock & LMK03200 are used. TTC clock will be used when SI5345 configuration is ready.
- The typical eye diagram is shown above: open area is 5312.

Other testing

- Firmware version control:
  - the software can configure the FPGA to load firmware from 1 of the 4 bit files in Flash, via the communication with micro-controller.
- Power dissipation:
  - current for 48× 12.8 Gb/s IBERT project: VCCINT needs about 9A, MGTAVCC is 10A, MGTAVTT is 5A.
  - project with 4× GBT links, and 2× PCIe cores: VCCINT is 3A, MGTAVCC is 3A, MGTAVTT is 1A, 0.9V for PEX8732 is about 5A.
  - *Summary: the power consumption is very close to the analysis done before board design. The whole board will dissipate <64W in the worst case.*
- Cooling:
  - for the 48 channel 12.8 Gb/s IBERT project, the FPGA inside temperature is about 63.7 degree.
  - for the IBERT project with 4.8 Gb/s. The temperature inside FPGA is about 53 degree, the outside temperature is about 38 degree.
  - the FPGA will use a fansink, we don't expect MiniPOD (<2W each) will need a heat sink, the air flow in 2U server should be sufficient.
  - *One day's running with 12.8 Gb/s links shows no evident problem is related to overheating.*
  - the flx-tools will support monitoring temperature of FPGA, MiniPODs and PEX8732.

Plan towards the pre-production design





- Now: bank 224-227 are used for PCIe. This creates congestion in the placement of the logic around that area. The data crossing SLR (Super Logic Region) boundary will increase timing violation.

- Next version: bank 226/227, and 229/230 or 231/232. A complete study will be done to compare these options. If all will work well, the one easier for PCB routing will be chosen.

Other changes

- Remove the DDR4 modules:
  - the PCB routing will be much easier.
  - no blind via will be used, to minimize sequential lamination.
  - board will be shorter. Space will be available for the integration with TTC PON, or white rabbit modules.

Summary

- FELIX prototype development is progressing well.
  - the v1p0 board has been tested extensively and used in the FELIX integration test.
  - the v1p5 board testing is ongoing.
  - good progress after the issue with power chips assembly was resolved.
  - the DDR4, PCIe, transceivers and flash programming have been tested.
  - more boards will be assembled & tested in the coming month, then distributed to FELIX development institutes.
- FELIX pre-production design will be launched soon.
  - major improvements have been identified.
  - new features will be incorporated based on prototype V1P5 test results.
  - board is expected to be available in second half of 2017.