Front**E**nd **LI**nk e**X**change

FELIX: a PCIe based high-throughput approach for interfacing front-end and trigger electronics in the ATLAS upgrade framework

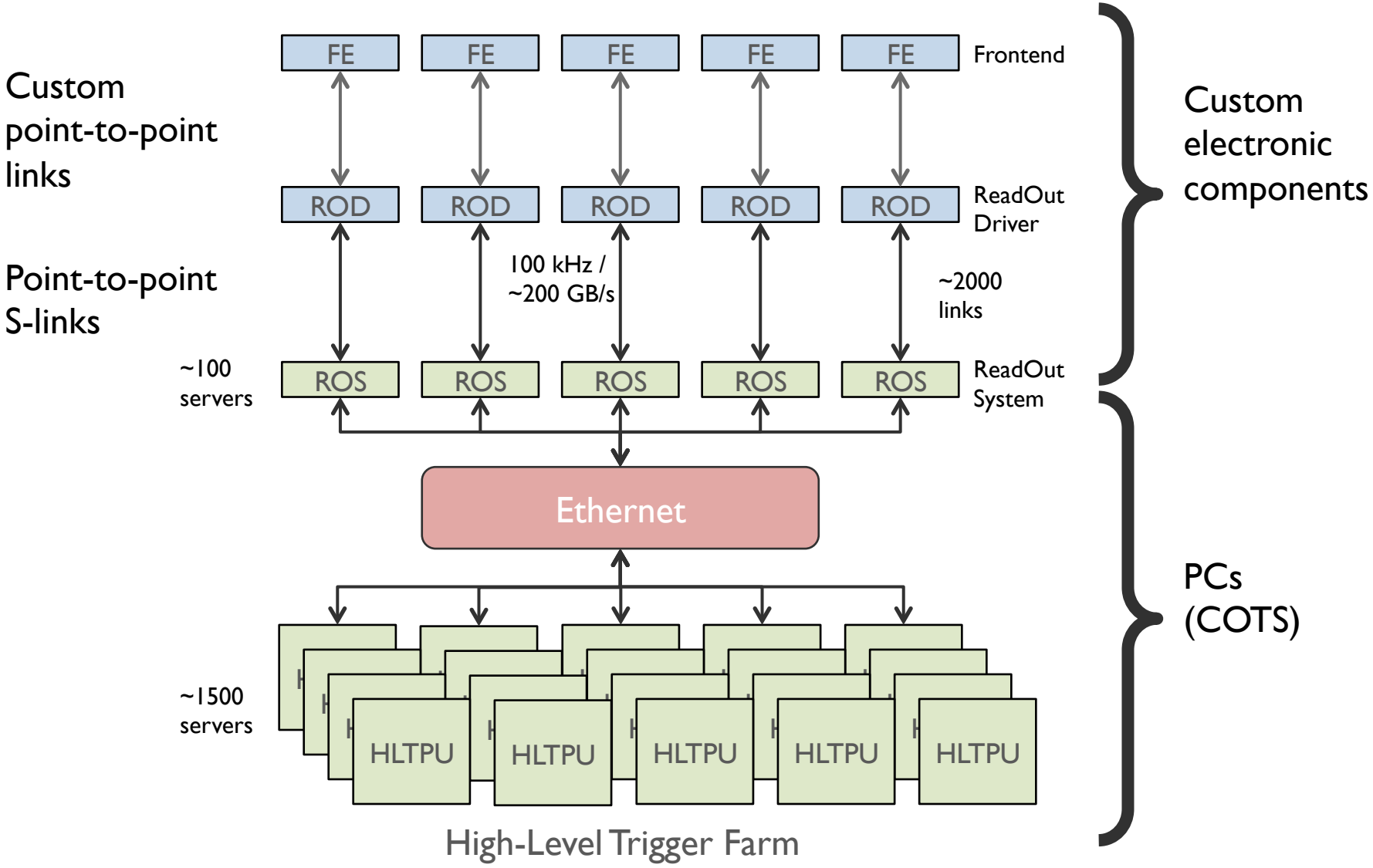
Kai Chen
Brookhaven National Laboratory

On behalf of the ATLAS Collaboration

Outline

- Introduction to the FrontEnd LInk eXchange (FELIX) system
- FELIX hardware
- FELIX firmware & software
- Integration test with Front-Ends
- Status & summary

ATLAS DAQ Today



Today

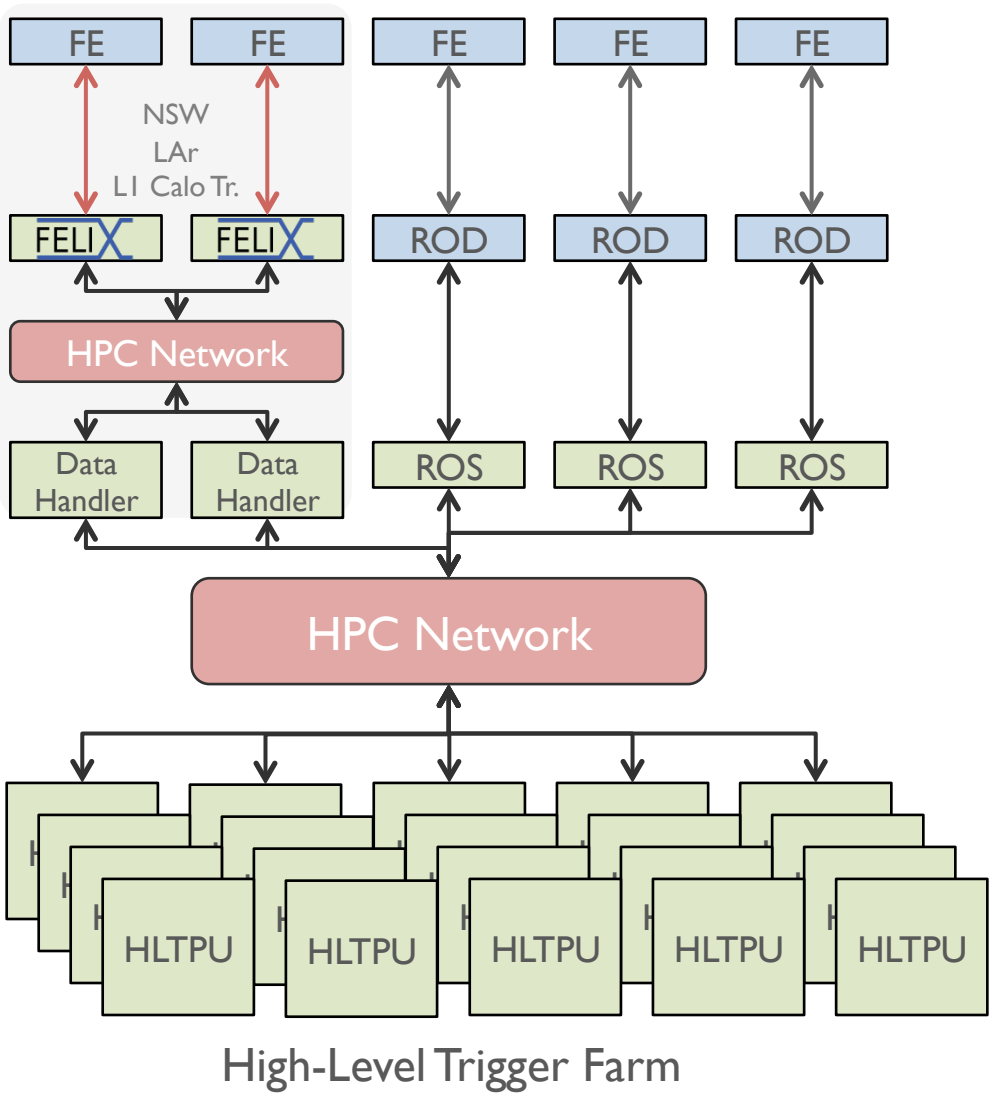
Run 2 2015 — Run 3 2017 2019 — Run 4 2021 2023 2025

Upgrade for Phase-I

Versatile Link, GBT (GigaBit Transceiver)

PCs

40 Gb Ethernet, Infiniband



Custom electronic components including FELIX cards

PCs (COTS)

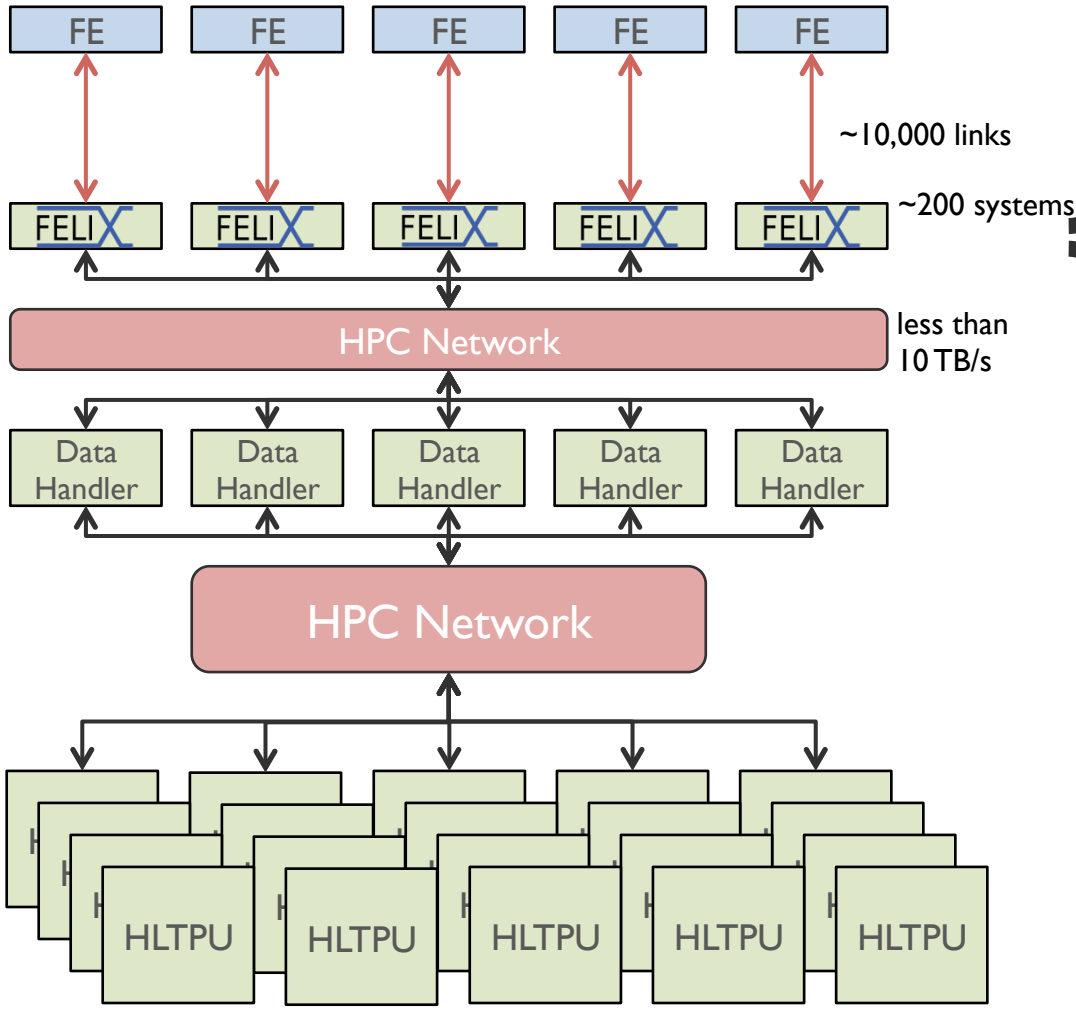
2018

Run 2 2015 2017 | Run 3 2019 2021 | Run 4 2023 2025

Upgrade for HL-LHC

Versatile Link,
GBT,
LpGBT (Low power GBT)

COTS network technology



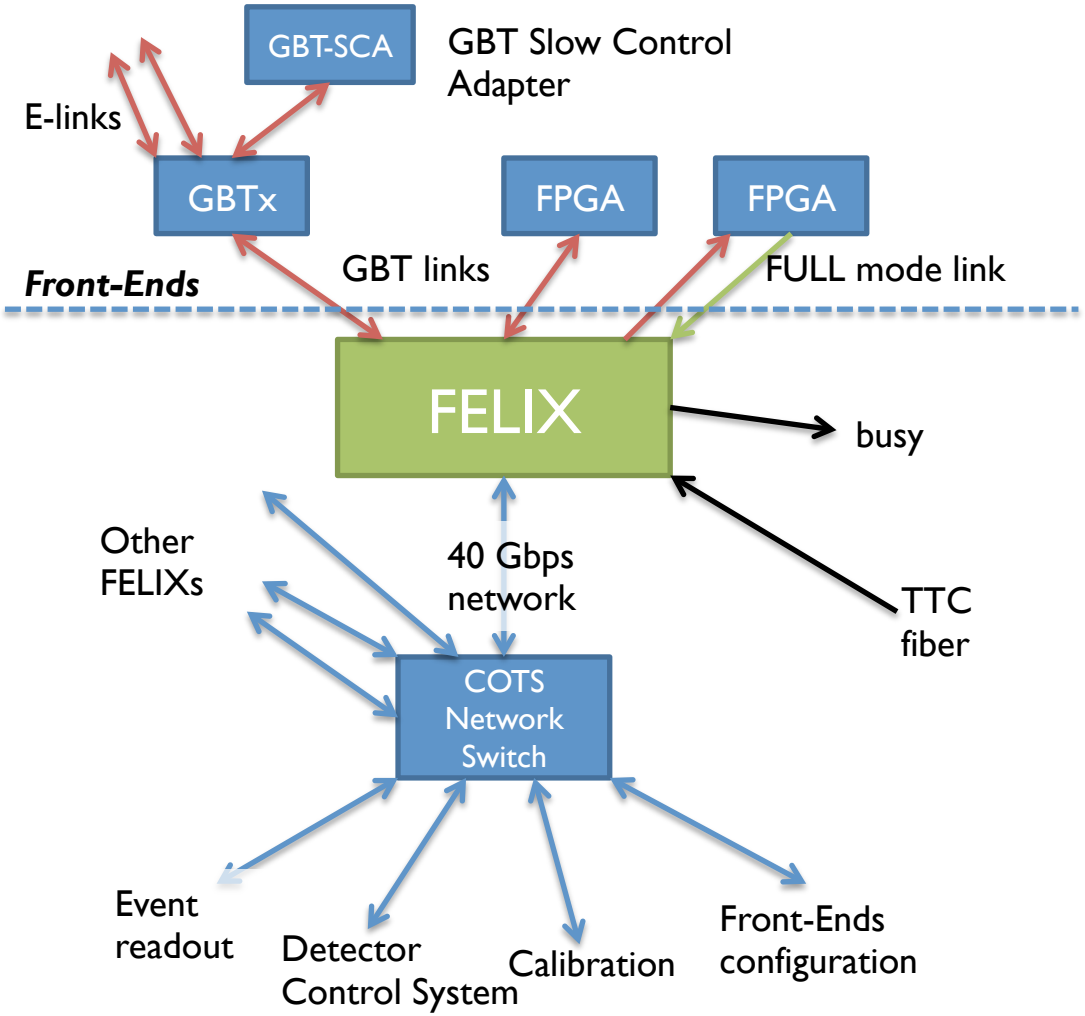
Custom electronic components including FELIX cards

PCs (COTS)

2023

Run 2 2015 | 2017 | Run 3 2019 | 2021 | Run 4 2023 | 2025

FELIX functionality



Normal GBT mode: **3.2** Gbps payload, with FEC (forward error correction)
GBT Wide-bus mode: **4.48** Gbps payload
FULL mode: current plan is **9.6** Gbps link speed in 8B/10B

- Scalable architecture
- Routing of event data, detector control, configuration, calibration, monitoring
- Connect the ATLAS detector Front-Ends to the DAQ system, for both the to and from FE directions
- E-links configuration configurable
- Detector independent
- TTC (Timing, Trigger and Control) distribution is integrated

FELIX server PC components

TTCfx

→ Custom FMC with TTC input , busy output
→ ver1: ADN2814 + CDCE62005
→ ver2: ADN2814 + SI5338
→ Ver3: ADN2814 + SI5345

FLX-709: Xilinx VC-709

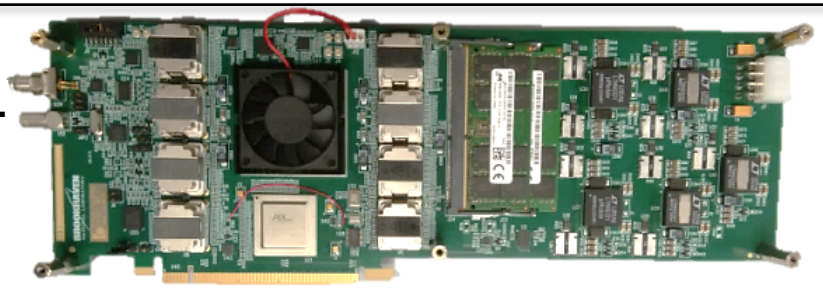
→ Subset of full FELIX, intended for FE development support
→ Virtex-7 X690T
→ 4 SPF+ connectors
→ PCIe Gen3 x8

or

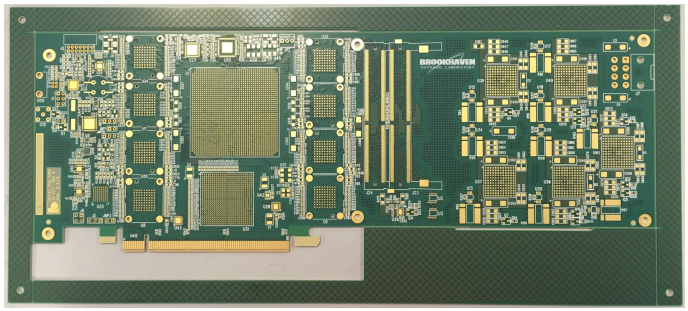
FLX-710: HiTech Global HTG-710

→ FELIX development
→ Virtex-7 X690T
→ 2x12 bidirectional CXP connectors
→ PCIe Gen3 x8

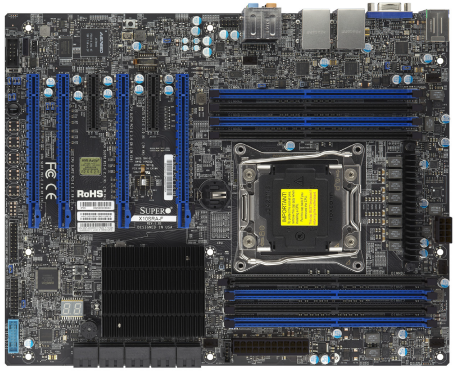
FLX-711 from BNL



→ FELIX Phase-I prototype
→ TTC input ADN2814 + SI5345
→ Xilinx Kintex **Ultrascale** XCKU115
→ 48 duplex optical links (based on MiniPODs)
→ PCIe Gen3 x16

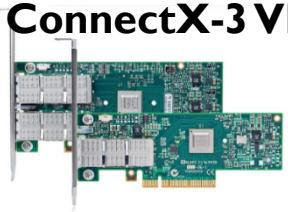


SuperMicro X10SRA-F used for development



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

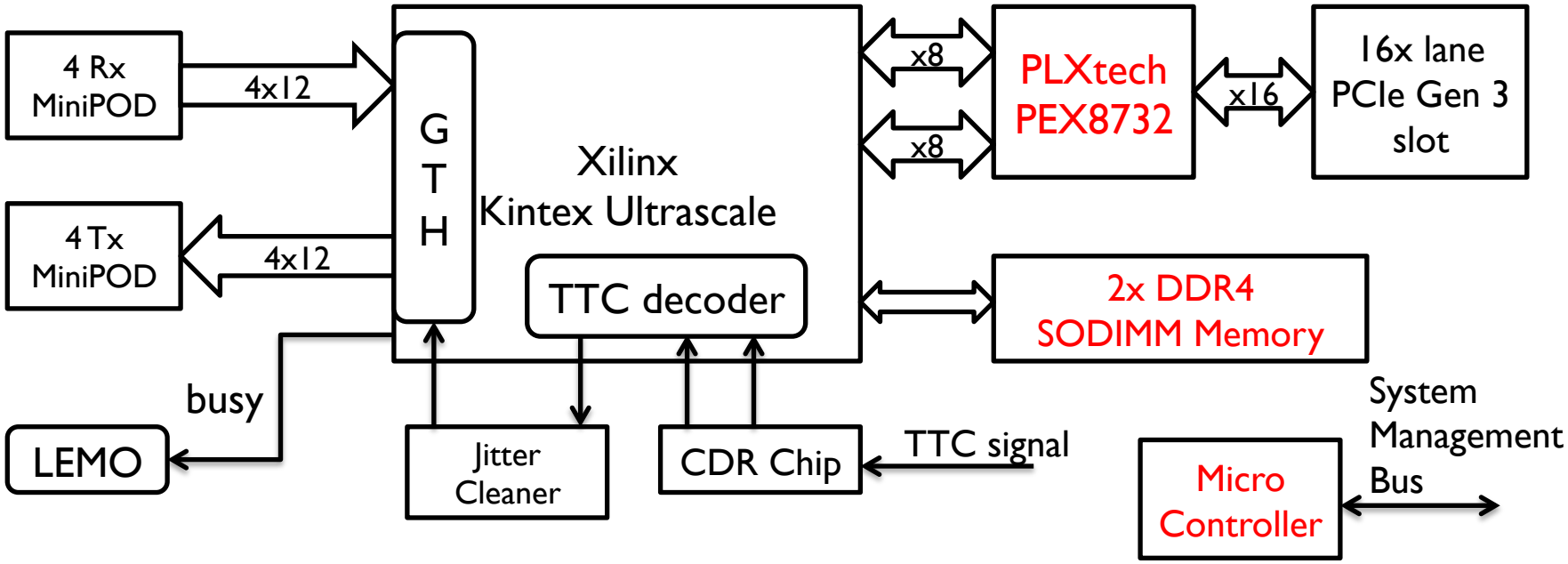
Mellanox ConnectX-3 VPI



→ 2x FDR/QDR Infiniband
→ 2x 10/40 GbE

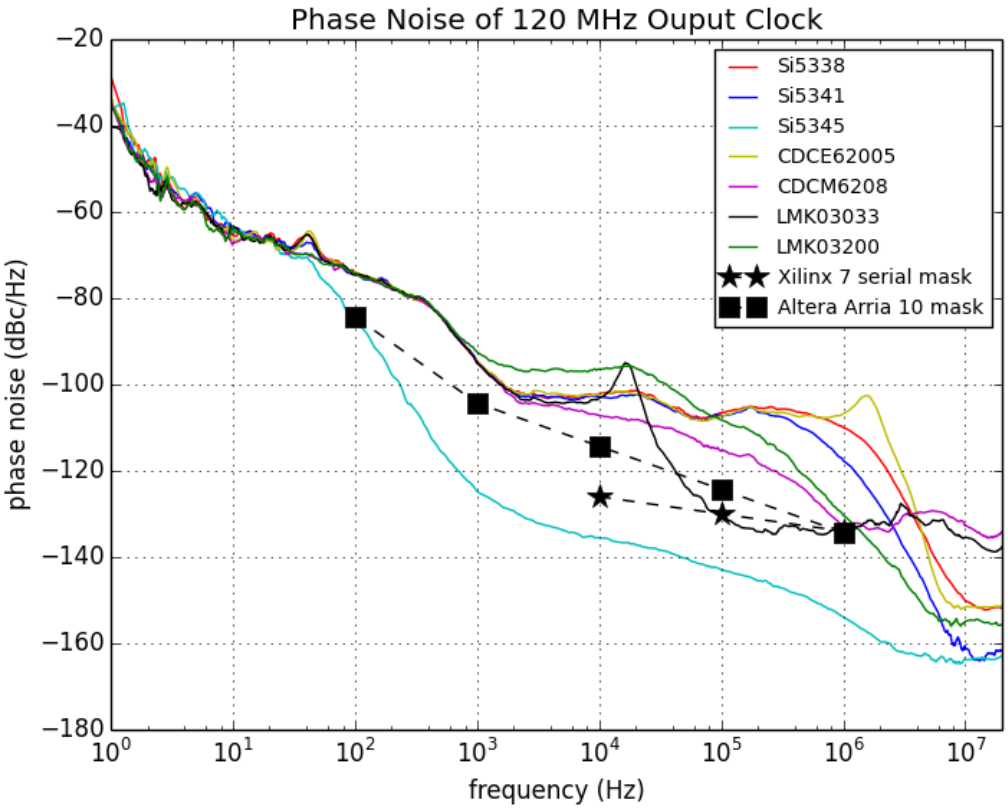
Features of BNL-711

FELIX base line hardware platform: PCIe FPGA board gen3 x16, "BNL-711"



- Developed at BNL also as the DAQ platform for the LTDB (Liquid Argon Trigger Digitizer Board) production test platform
- PLXtech PEX8732 to handle PCIe Gen3 x16 lanes (max 128 Gbps) interface to host
- 48-ch MiniPOD TX & RX, up to 14Gb/s per link
- 2x SODIMM DDR4 interfaces (not used in FELIX)
- Integrated TTC interface, busy output, and on-board jitter cleaner
- Micro-Controller (Atmega 324A) for FPGA firmware update and version control

Status of BNL-711

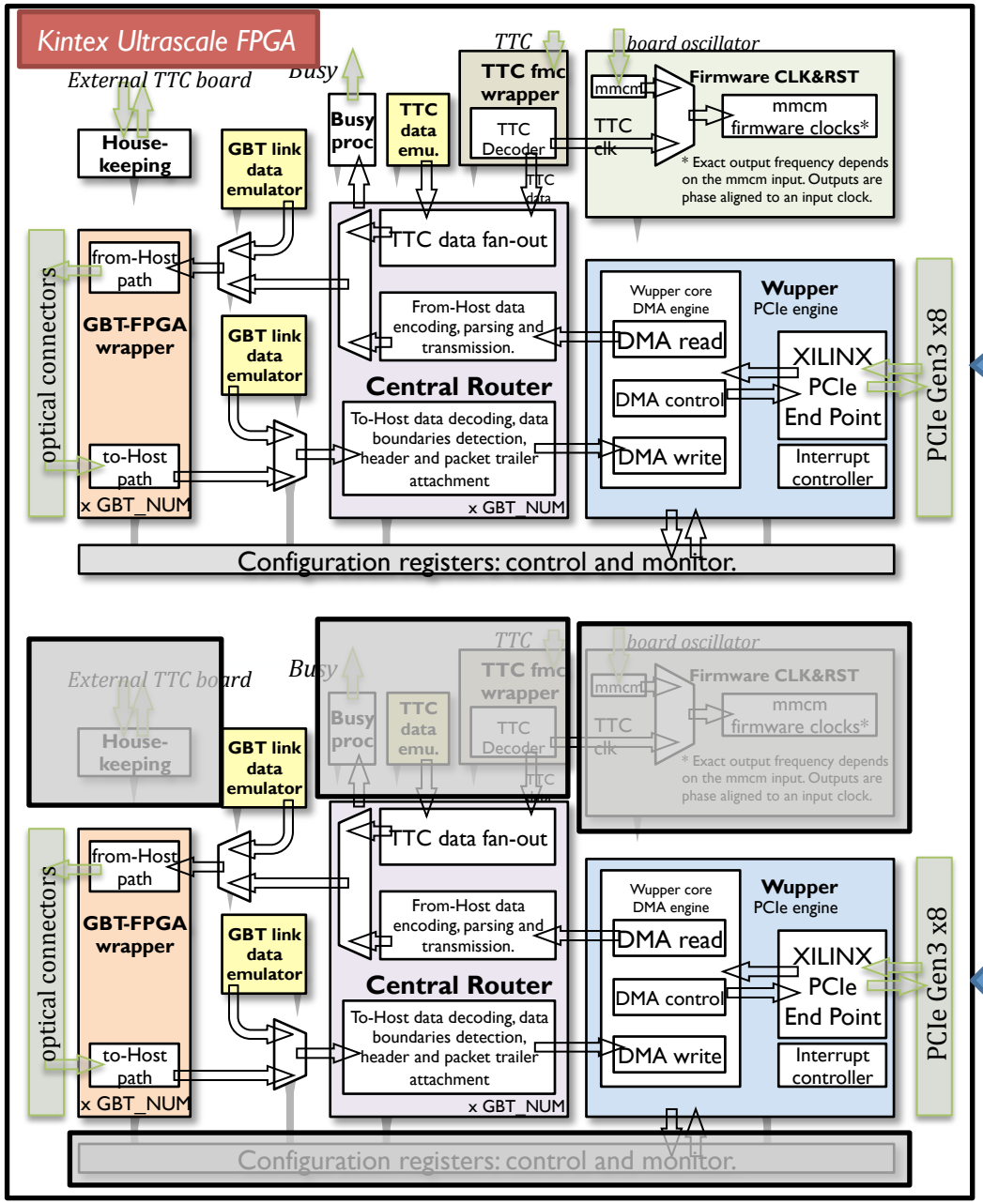


- All hardware features are verified.
  - PCIe throughput: 101.7 Gbps.
  - Optical links can operate at 12.8 Gbps.
  - The FPGA can be (re)programmed with different bitfiles in the FLASH (4 images per memory).
- FELIX firmware is successfully migrated to BNL-711
- A second version has been fabricated.
  - Jitter cleaner with better performance (**Si5345**).
  - More accurate FPGA voltage control.

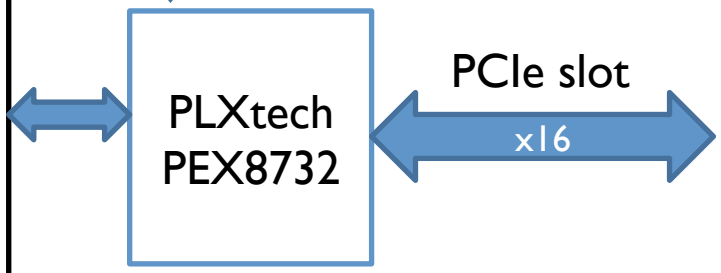
Jitter (10 kHz to 1 MHz)

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

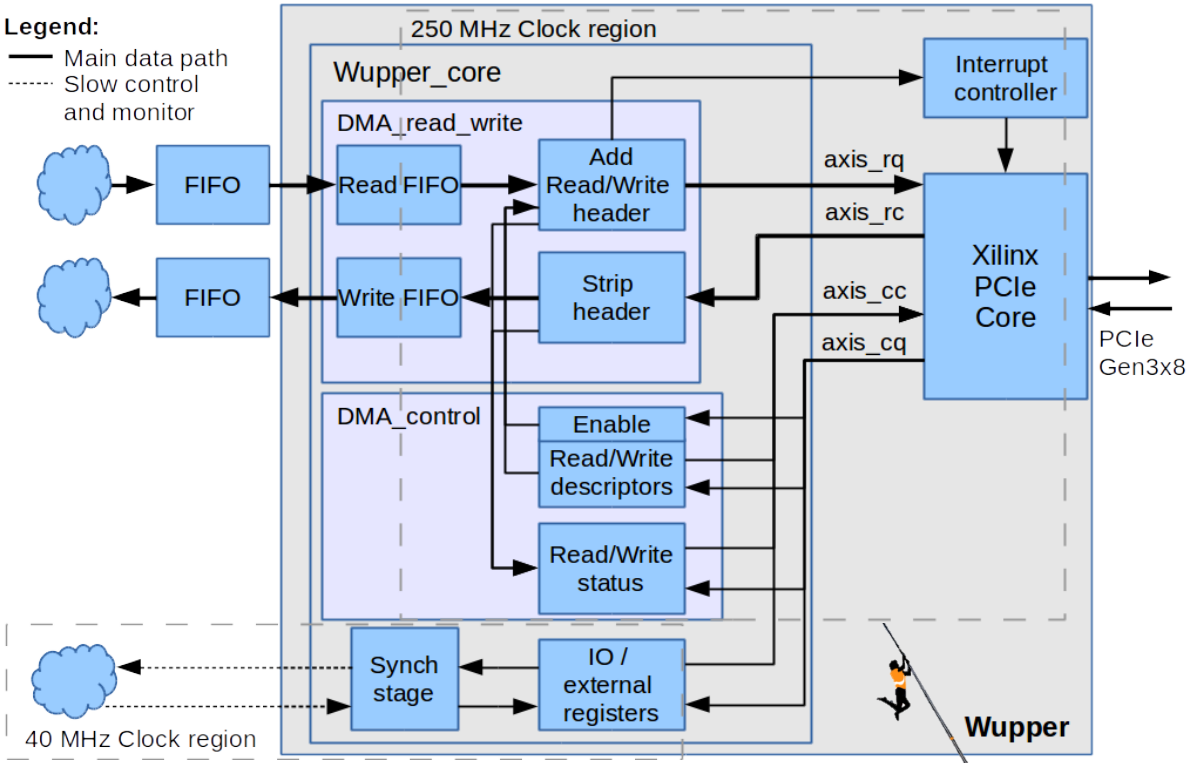
FELIX firmware block diagram



- For BNL-711: the PC will see **two** independent Gen3 x8 lanes EndPoints, as Xilinx PCIe devices.
- 2 sets of an identical firmware block are instantiated in the top level design.
- Shared facilities in grey (include clock resources, the housekeeping module, and the registers controls and monitors).



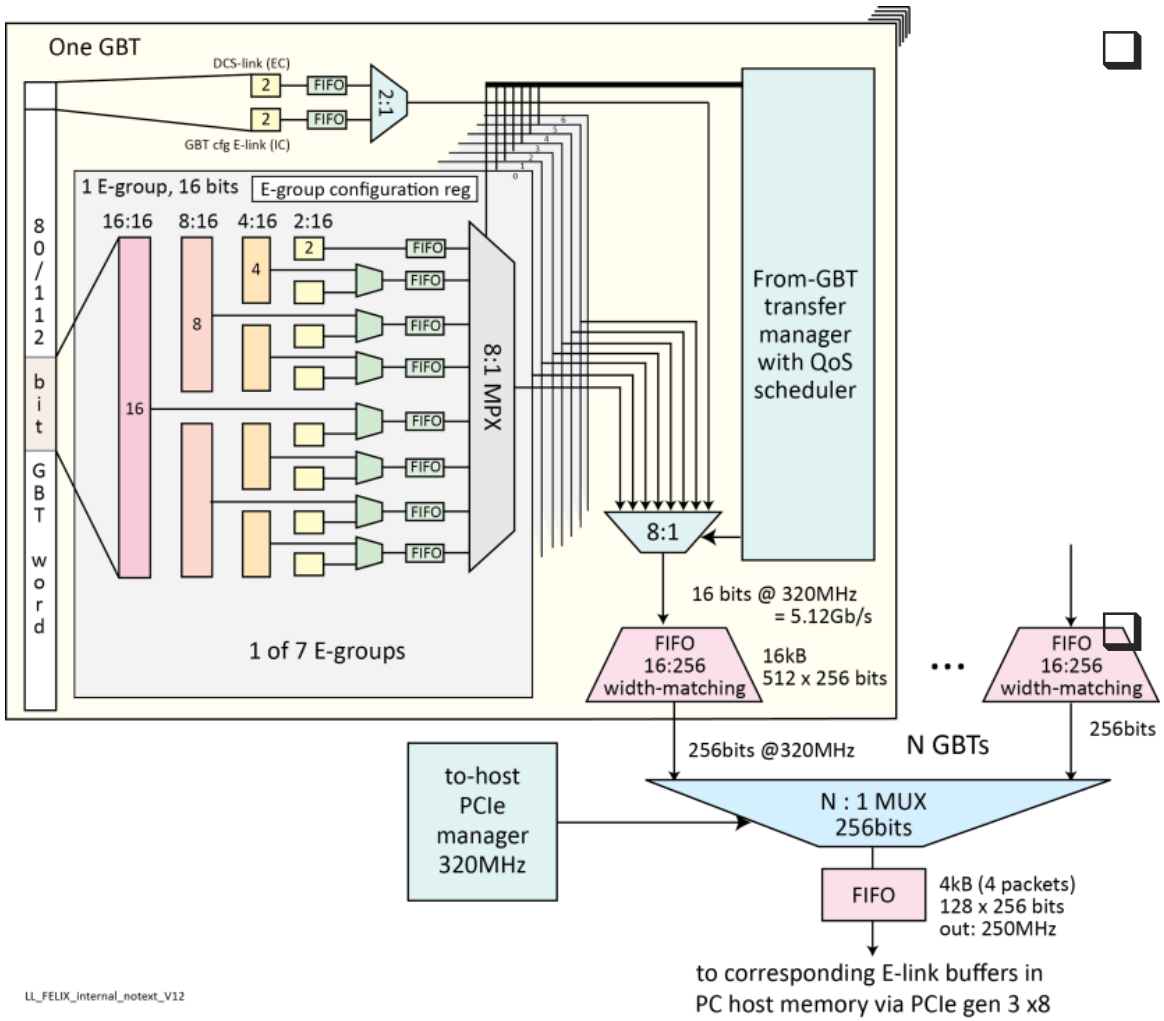
Wupper: PCIe engine for FELIX



- PCIe Engine with DMA interface to the Xilinx Virtex-7 (Kintex Ultrascale) PCIe Gen3 Integrated Block for PCI Express
- Xilinx AXI (ARM AMBA) Stream Interface (UG761)
- MSI-X compatible interrupt controller
- Applications access the engine via simple FIFOs
- Register map for programmed I/O synchronized to a lower clock speed

- Developed for use in FELIX
- Published as Open Source (LGPL) on OpenCores
  http://opencores.org/project,virtex7_pcie_dma
- Extended functionality to support and endless DMA
- Extensive stress testing of the core
- Core maturing to maintenance only phase
- Positive feedback from the community

Central router: internal data multiplexing



Shown: 1 of 7 E-groups of a GBT; 1 of "N" GBTs

GBT mode

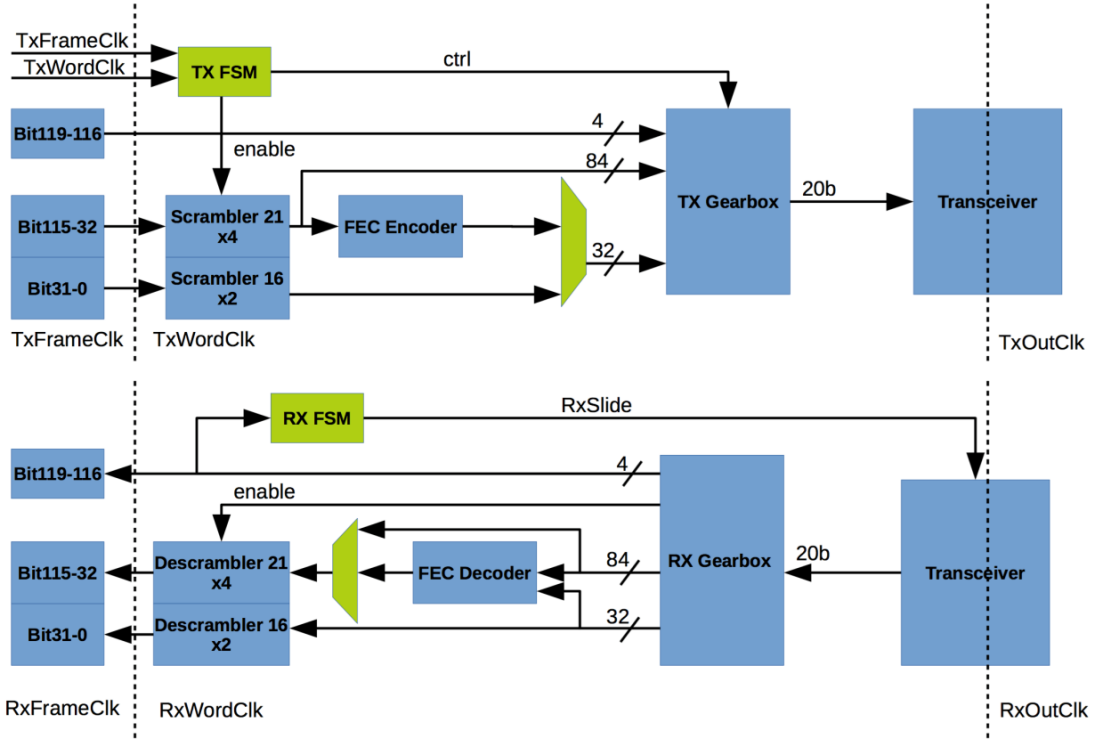
- Handles data streams:
- Routes TTC information
- Dedicated manager for E-Links in GBT channels
- Communicate with GBTx & GBT-SCA
- Main manager toward PCIe Engine

FULL mode

- Line rate: 9.6 Gb/s
- Maximum user payload: 7.68 Gb/s: 8B/10B encoding
- Packets unit: 32-bit
- Option to include a **stream id** for transmitting different logical data streams on same physical link
- Support for BUSY-ON and BUSY-OFF

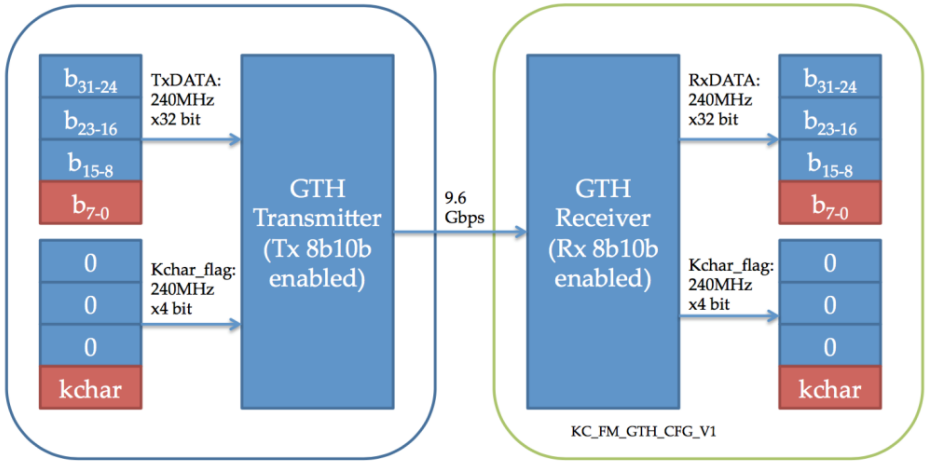
Transceiver wrapper for FELIX

- Supports 4.8 Gb/s **GBT mode**, and 9.6 Gb/s **FULL mode** defined for FELIX.
- FELIX **GBT Wrapper** is based on CERN GBT-FPGA, with some improvements:
  - Separated GBT firmware from transceiver block.
  - Run-time choice of **GBT mode** : Normal (FEC) mode or Wide-Bus mode.
  - **Lower fixed latency** (Tx: 27.8~32 ns; Rx: FEC mode 56.4ns; Rx: Wide mode 43.9 ns).
    - The GBT encoding/decoding are in the 240 MHz domain.
  - Some blocks like *Rx side frame alignment*, *Tx side time domain crossing* are redesigned.
  - The single channel example design for KC705/VC709 can be found at https://github.com/simpway/GBT_KC705

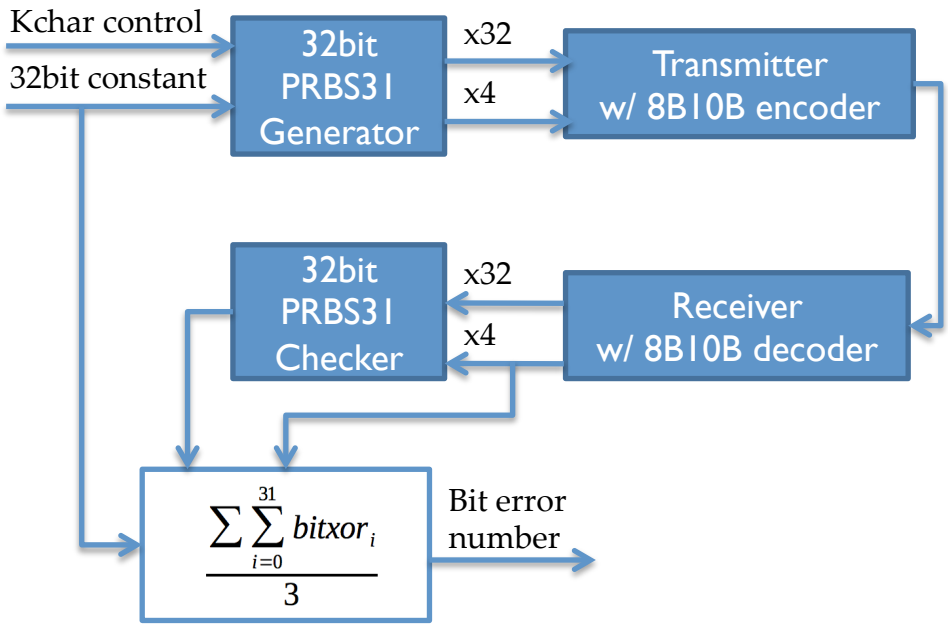


Transceiver wrapper for FELIX

- Supports 4.8 Gb/s **GBT mode**, and 9.6 Gb/s **FULL mode** defined for FELIX.
- FELIX **FULL mode**:
  - The transceiver is configured as 32 bit x 240 MHz.
  - The Least Significant Byte is used for the 8B/10B alignment.
  - An example design with **32-bit PRBS31** generator and checker is provided.
    - **5 control symbols** are used for the flow control, busy control.
    - No error occurred for ~72 hours run. **BER < 1E-15**.
    - Has been used in the integration test with **gFEX** (Global Feature Extractor) board.



The transceiver configuration for **FULL mode**

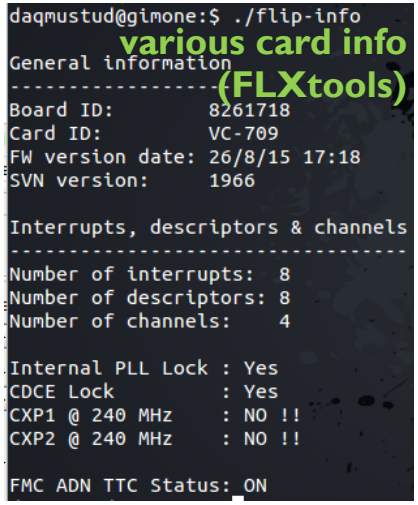


The example for PRBS testing

FELIX software: control, monitor

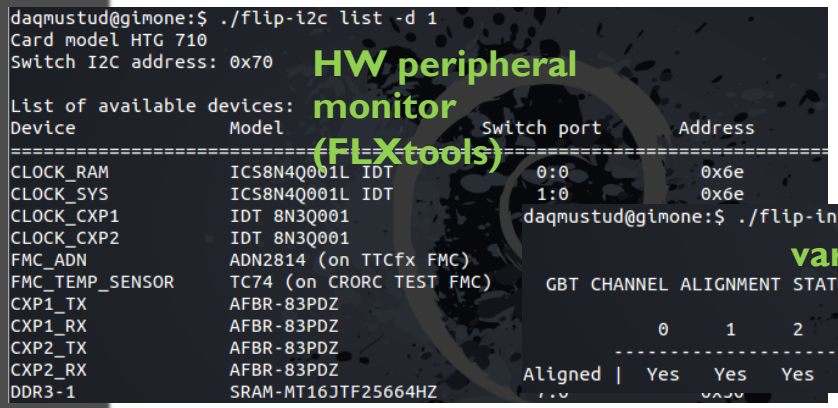
FELIX low-level tools: *FLXtools*

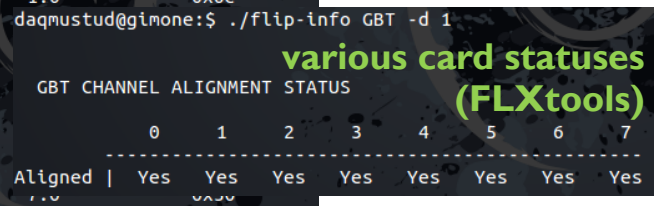
FELIX firmware configuration through register map	→ Configure the FLX card via the internal registers access by the PC.
FELIX 'housekeeping' functions	→ Control FLX external devices (e.g. clock chips) via I2C and SPI.
System test and debugging	→ Performance testing, data flow control and verification.
Complete user toolset	→ Collection of tools to control, monitor, configure different FLX cards.




```
$ flx-dump-blocks -n 100 -f output.blocks
$ flx-config
$ flx-dma-test
$ flx-eeprom
$ flx-i2c
$ flx-info
$ flx-init
$ flx-irq-test
$ flx-spi
$ flx-reset
$ flx-throughput
```

FLXtools set





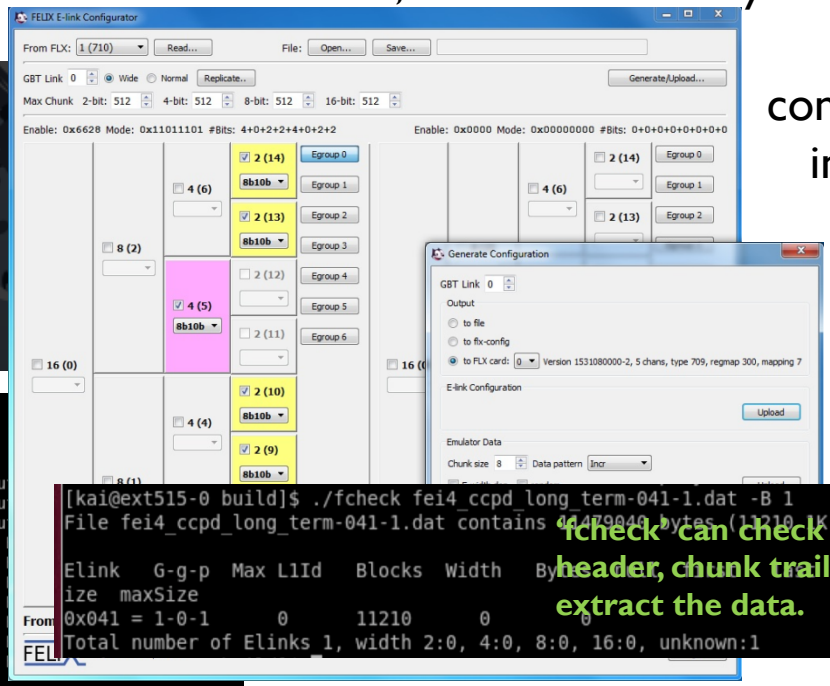
FELIX software: support

FELIX support software (*WupperCodeGen, fel, fdaq, fcheck, FELIX E-link configurator)*

Automatic generation of a register map	→ Automation software (Jinja based) to handle register map consistently between firmware (HDL), software (C headers), and documentation (Latex) and beyond (ATLAS configuration database, ...).
Firmware development tools	→ Configuration of the internal data emulators to verify the functionality 'off-line'.
Firmware test and debugging	→ Communication with FLX card, data read and analysis.

```
daqmustud@gimone:$ ./fel -u 2048 -f 64 -t 8 -s
FelixCardBuffer 1

BufSize 4194304
2 8 0.000068
   0 a0  1 cd ab aa  0 89  0  0 bb aa  4  0  1
  16  4  5  6  7  8  9  a  b  c  d  e  f 10 11
  32 14 15 16 17 18 19 1a 1b 1c 1d 1e 1f 20 21
  48 24 25 26 27 28 29 2a 2b 2c 2d 2e 2f 30 31
```

loopback via GBT links of data from internal generator ('fel' tool output)



FELIX configuration interface

```
[kai@ext515-0 build]$ ./fdaq -t 10 adc31_test-3.dat
Opened FLX-card 0, firmw 1608052337-3193 (cmem buffersize=1073741824)
**NB**: FanOut-Select registers are locked!
**START** using DMA #0
-> 1 sec, Rates: recv 323.7 MB/s, file 316.6 MB/s; Total: recvd 323 MB, file 316 MB; Bu
-> 2 sec, Rates: recv 320.1 MB/s, file 324.4 MB/s; Total: recvd 643 MB, file 641 MB; Bu
-> 3 sec, Rates: recv 318.4 MB/s, file 96 MB/s; Total: recvd 96 MB, file 956 MB; Bu
-> 4 sec, Rates: recv 320.2 MB/s, file 320.0 MB/s; Total: recvd 1282 MB, file 1276 MB;
-> 5 sec, Rates: recv 319.9 MB/s, file 321.3 MB/s; Total: recvd 1602 MB, file 1597 MB;
-> 6 sec, Rates: recv 320.2 MB/s, file 320.0 MB/s; Total: recvd 1922 MB, file 1917 MB;
-> 7 sec, Rates: recv 320.1 MB/s, file 318.4 MB/s; Total: recvd 2242 MB, file 2236 MB;
-> 8 sec, Rates: recv 317.8 MB/s, file 320.0 MB/s; Total: recvd 2560 MB, file 2556 MB;
-> 9 sec, Rates: recv 317.0 MB/s, file 320.6 MB/s; Total: recvd 2877 MB, file 2876 MB;
->10 sec, Rates: recv 320.0 MB/s, file 320.4 MB/s; Total: recvd 3197 MB, file 3197 MB;
**STOP**
-> Totals: Recvd 3204 MB, File 3197 MB (last 1045 MB)
Exiting..
```

'fdaq' is used to stream data to disk continuously.

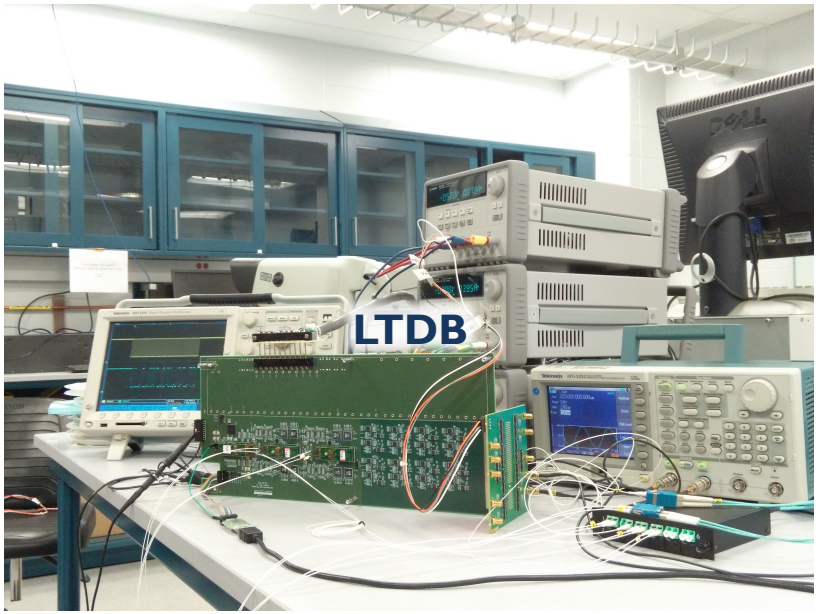
```
[kai@ext515-0 build]$ ./fcheck fei4_ccpd_long_term-041-1.dat -B 1
File fei4_ccpd_long_term-041-1.dat contains 44479040 bytes (11210 1K blocks +0)

Elink   G-g-p  Max L1Id   Blocks  Width   Byte
ize  maxSize
0x041 = 1-0-1      0       11210       0       0
Total number of Elinks 1, width 2:0, 4:0, 8:0, 16:0, unknown:1
```

'fcheck' can check the block header, chunk trailer, and extract the data.

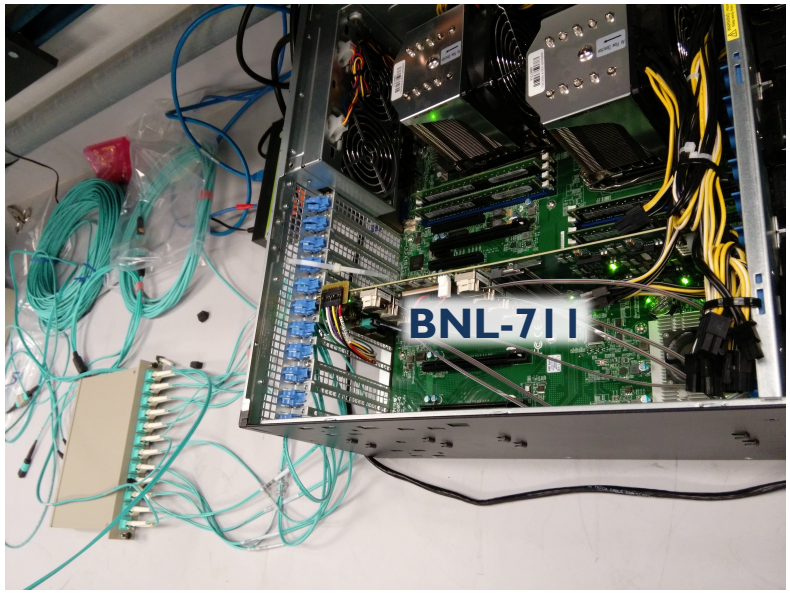
Integration test with front-ends

- ATLAS sub-detectors connected to FELIX
  - Liquid Argon Calorimeter
    - **LTDB** (LAr Trigger Digitizer Board)
    - **LDPB** (LAr Digital Processing Blade)
  - Level-1 calorimeter trigger
    - **gFEX** (Global Feature Extractor)
    - **ROD**, **Hub** for **eFEX** (Electron Feature Extractor) and **jFEX** (Jet Feature Extractor)
    - **TREX** (Tile Rear Extension) modules
  - New small wheel of the muon spectrometer
    - **sTGC** (Small-strip Thin Gap Chamber) and **MicroMegas** (Micro Mesh Gaseous Structure) detector for muon tracking
  - Tile Calorimeter
    - Test system for Phase-II readout
  - CaRIBOu (Control and Readout ITk Board)
    - Test system for Phase-II ITk HV-CMOS pixel sensor R&D

Integration test with front-ends



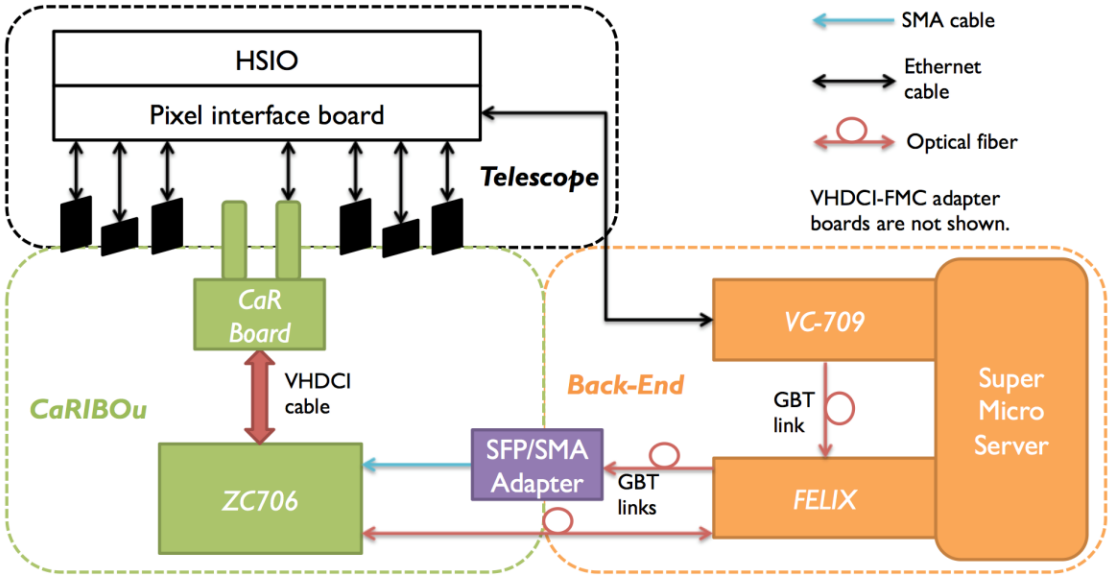
64-ch LTDB prototype

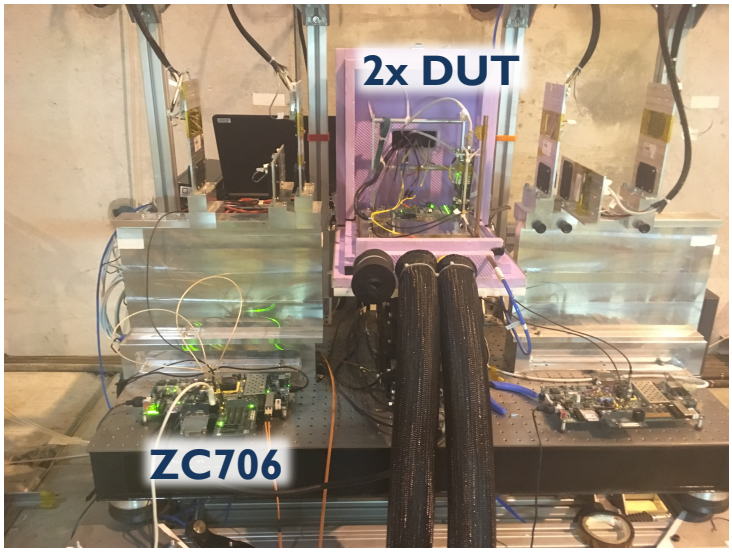


Back-End for LTDB test

- 64-ch LTDB (Liquid Argon Calorimeter Trigger Digitizer Board) Prototype
  - With GBTx and GBT-SCA on board
  - **Software based** tool is used to control the GBTx and GBT-SCA
  - FELIX (BNL-711 or FLX-709) systems are used
    - Perform control and monitoring
    - Clock and **BCR** (Bunch Counter Reset) command distribution.
  - Successful testing was accomplished in April 2016.

Integration test with front-ends

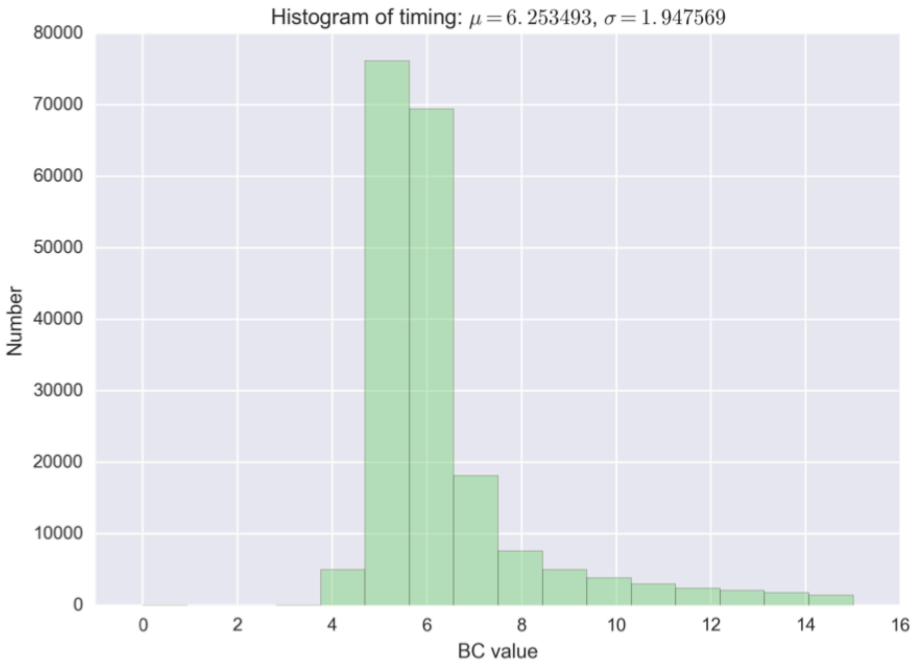


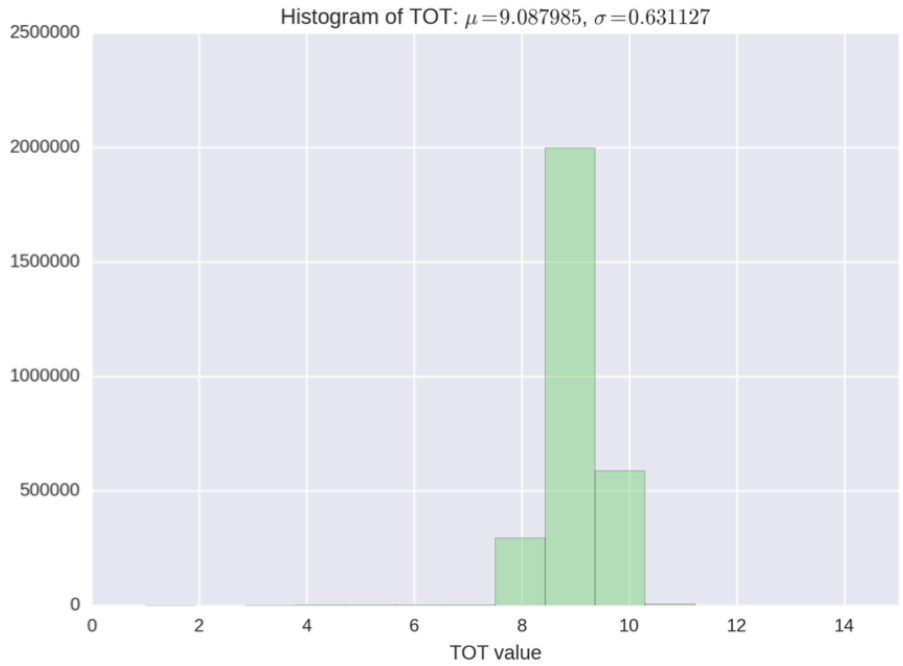


Telescope & CaRIBOu with 2 DUT (Device Under Test)

- CaRIBOu readout system for HV-CMOS sensor
  - Data transmission between CaRIBOu and FELIX is in normal GBT mode.
  - FELIX is used to
    - control and monitor the setup
    - distribute clock and TTC commands.
    - calibrate the sensor, and readout ASIC FE-I4 by software.
  - ***fdaq***: to stream data to disk continuously; ***fcheck***: to check and extract FE-I4 data.
  - The testbeam for the system integration was carried out successfully at the end of August 2016 @ CERN.

Results of CaRIBOu + FELIX integration





Timing distribution compared to trigger from telescope

Time-over-Threshold distribution of the events

- Very stable continuous data taking (up to 13 hours in one night) without any glitch
- Calibration can be done in ~5 minutes, while FE-I4B calibration takes 2 minutes 50 seconds only.
- These results are consistent with current telescope readout system (http://arxiv.org/abs/1603.07798).

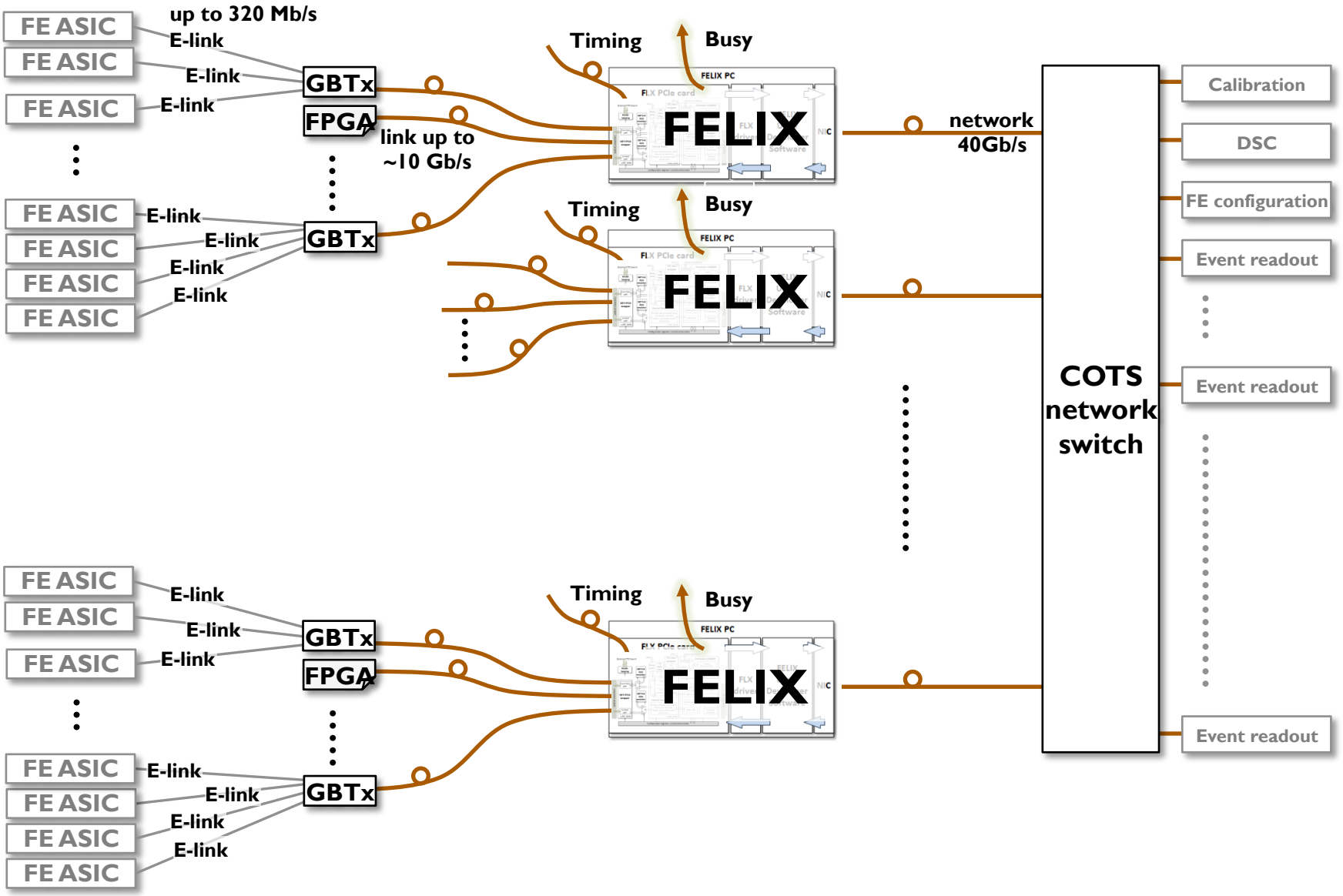
Summary

- FELIX is a **router** between custom serial links and a commodity network, separates data transport from data processing.
- In **LHC Run-3 (2020-2022)** FELIX will be used by some detectors and trigger systems to interface the data acquisition, detector TTC systems. In LHC Run-4 this is planned for all ATLAS detectors.
- Status:
  - FLX-709 (Xilinx VC-709) and its software reached a development status sufficient to be distributed to ATLAS Sub-Detectors Front End developers.
  - Initial tests with the BNL FLX-711 successfully accomplished.
  - A second revision with minor bug fixes is produced and will be tested soon.
- Ongoing efforts
  - Increase overall system reliability.
  - Increase the number of input channels supported.
  - Extend the system testing and integration: firmware to control of GBTx ASIC via its IC port, GBT-SCA ASIC via EC port.
  - A C++ API, an OPC server and client are in progress to fully support the GBT-SCA ASIC.
- **Final Design Review** planned for **November 2016**.

FELIX

Thanks!

FELIX data flow overview



FELIX software: data path

FELIX application:

Read blocks from file descriptors	→ DMA transfers and reads 1KByte blocks of data, encoded for the transfer over PCIe. Every block contains accumulated data from one E-link.
Block decoding	→ Decode into variable sized packets for transmission over network.
Statistics	→ Count processed blocks, transfer rates, etc.
Extract Meta Information	→ Meta-information, for example event ID, is extracted and matched against a routing table.
Route	→ Route.
Load Balancing, Error Handling	→ Distribute load among multiple systems, handle automatic failover in case of system failures.
Send to network	→ Asynchronous message service 'netio' for data exchange with network hosts.