



FELIX

A high-throughput network approach for interfacing to front end electronics for ATLAS upgrades

Jörn Schumacher

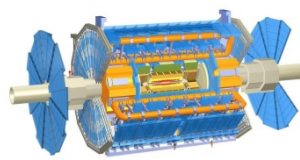
CERN / University of Paderborn, Germany

jorn.schumacher@cern.ch

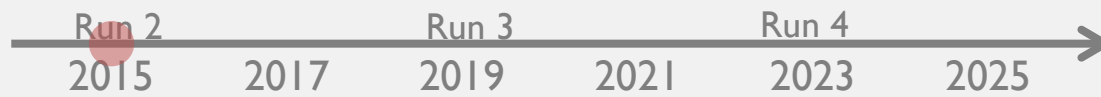
On behalf of the ATLAS FELIX Developer Team

ATLAS DAQ:

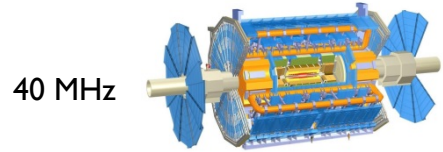
40 MHz



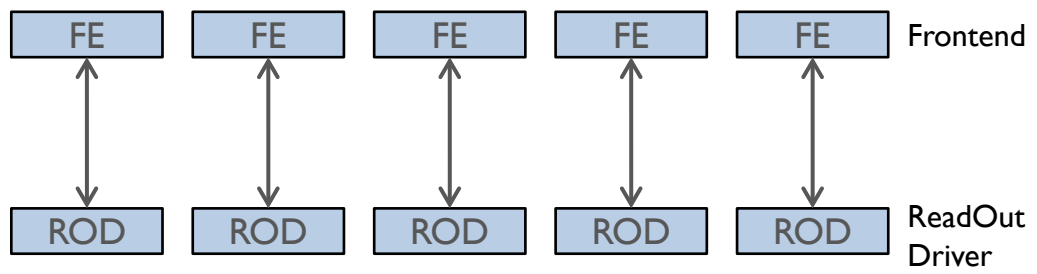
Today



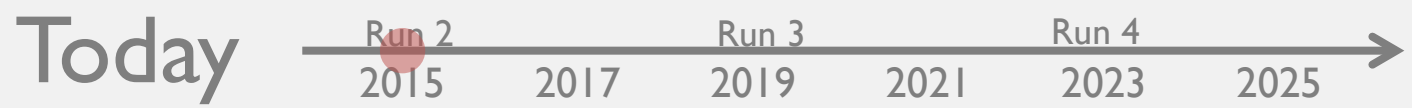
ATLAS DAQ:



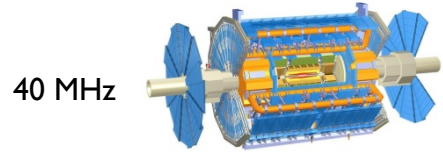
Custom
point-to-point
links



Custom
electronic
components

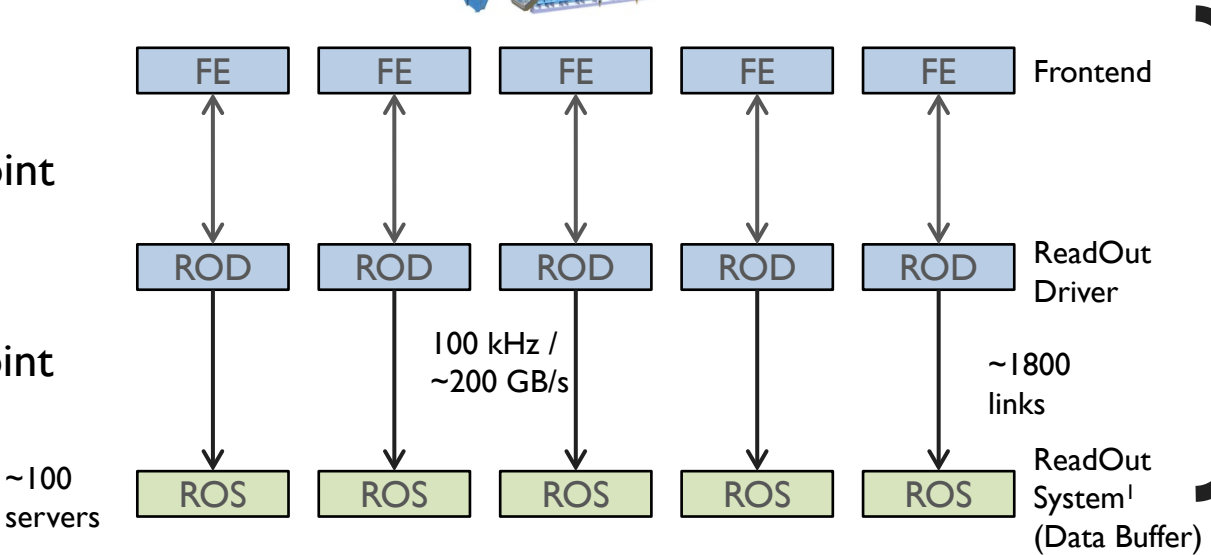


ATLAS DAQ:



Custom point-to-point links

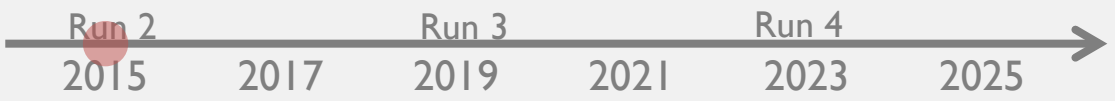
Point-to-point S-links



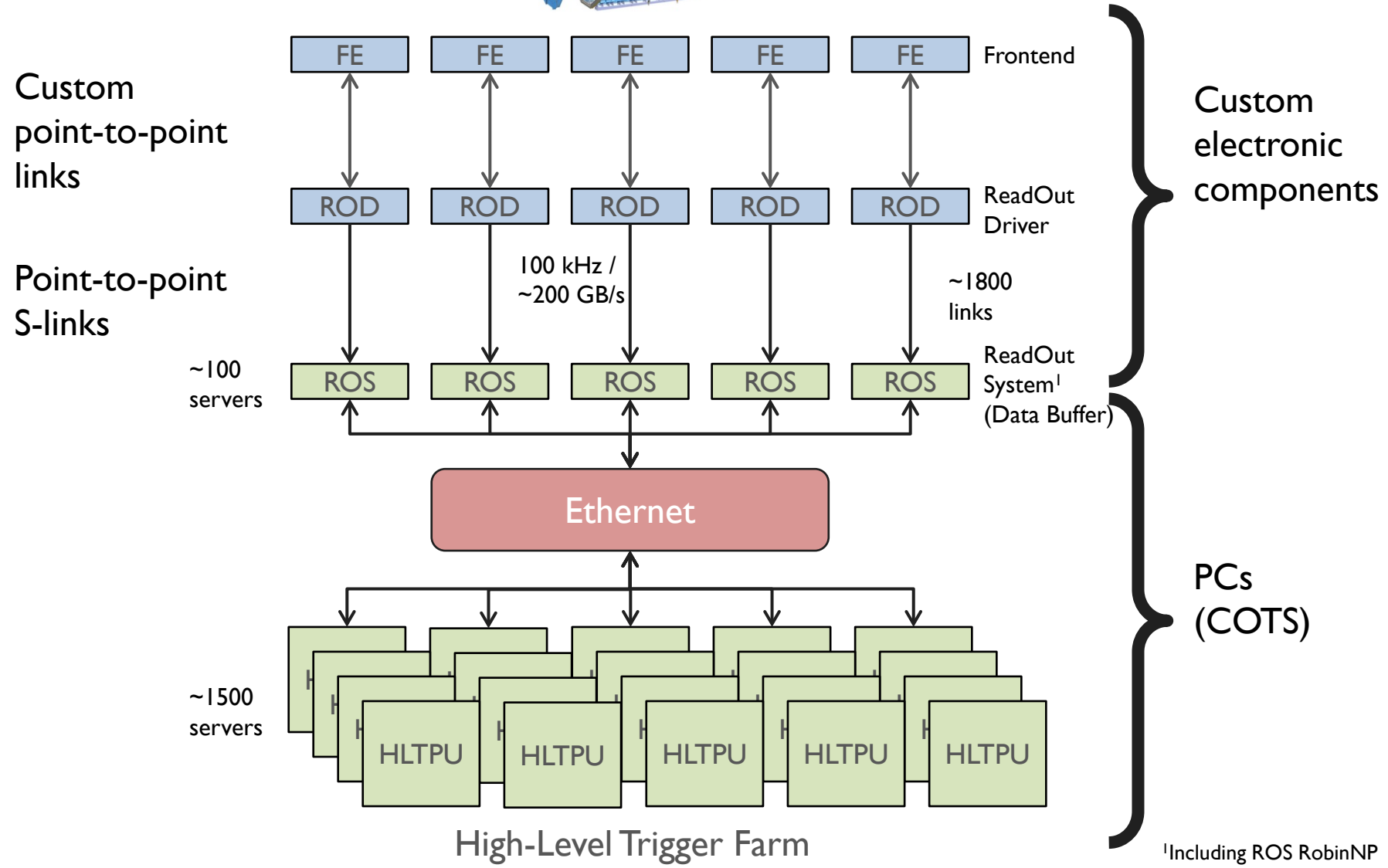
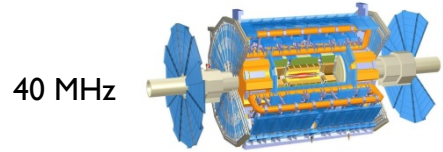
Custom electronic components

¹Including ROS RobinNP

Today

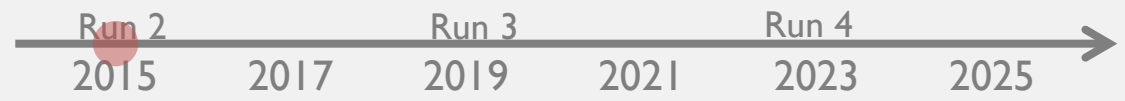


ATLAS DAQ:

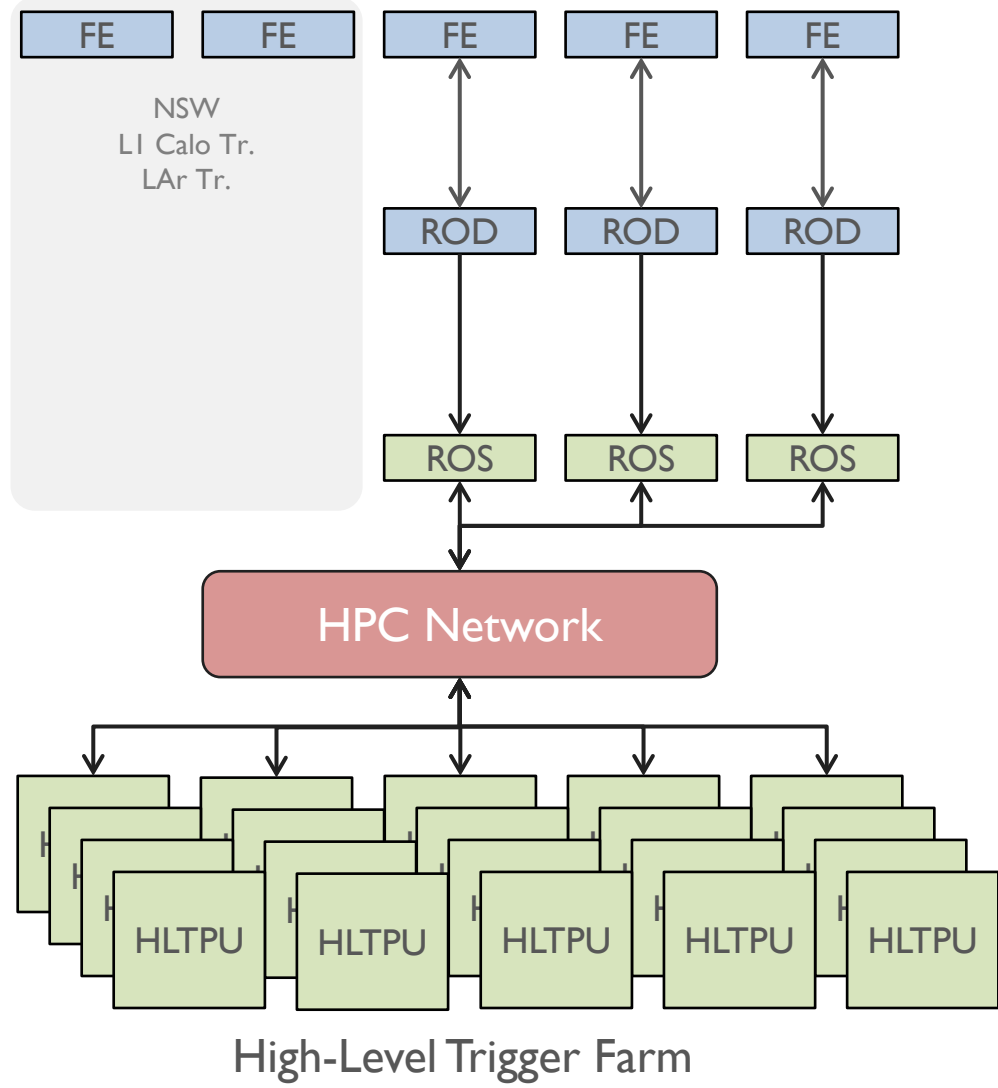
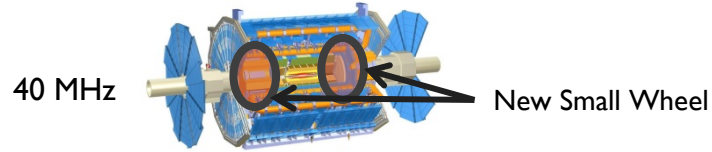


¹Including ROS RobinNP

Today



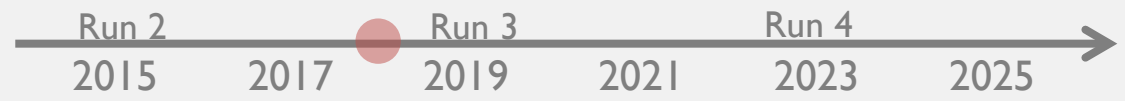
ATLAS DAQ:



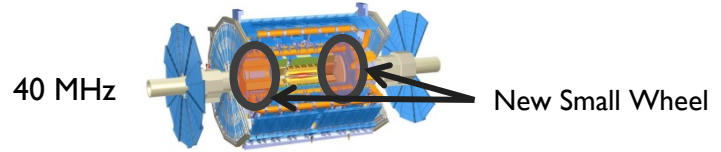
Custom electronic components

PCs (COTS)

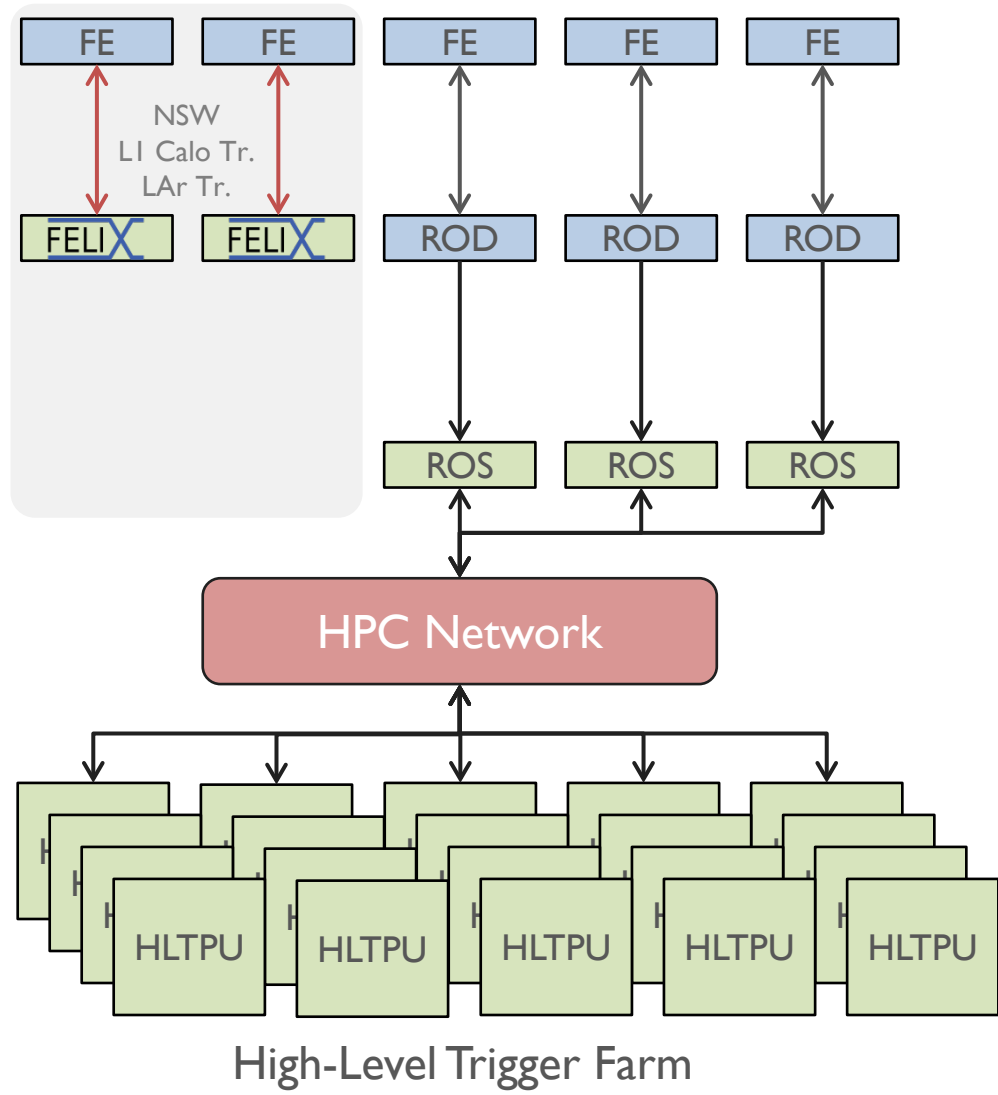
2018



ATLAS DAQ:



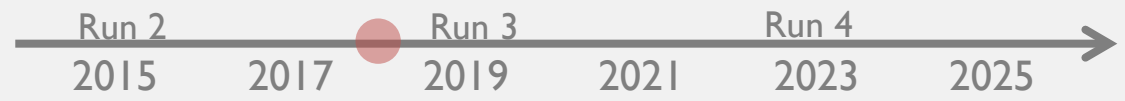
Versatile Link,
GBT
PCs



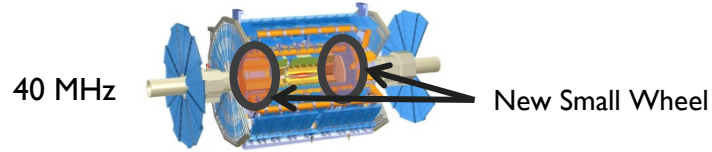
Custom electronic components

PCs (COTS)

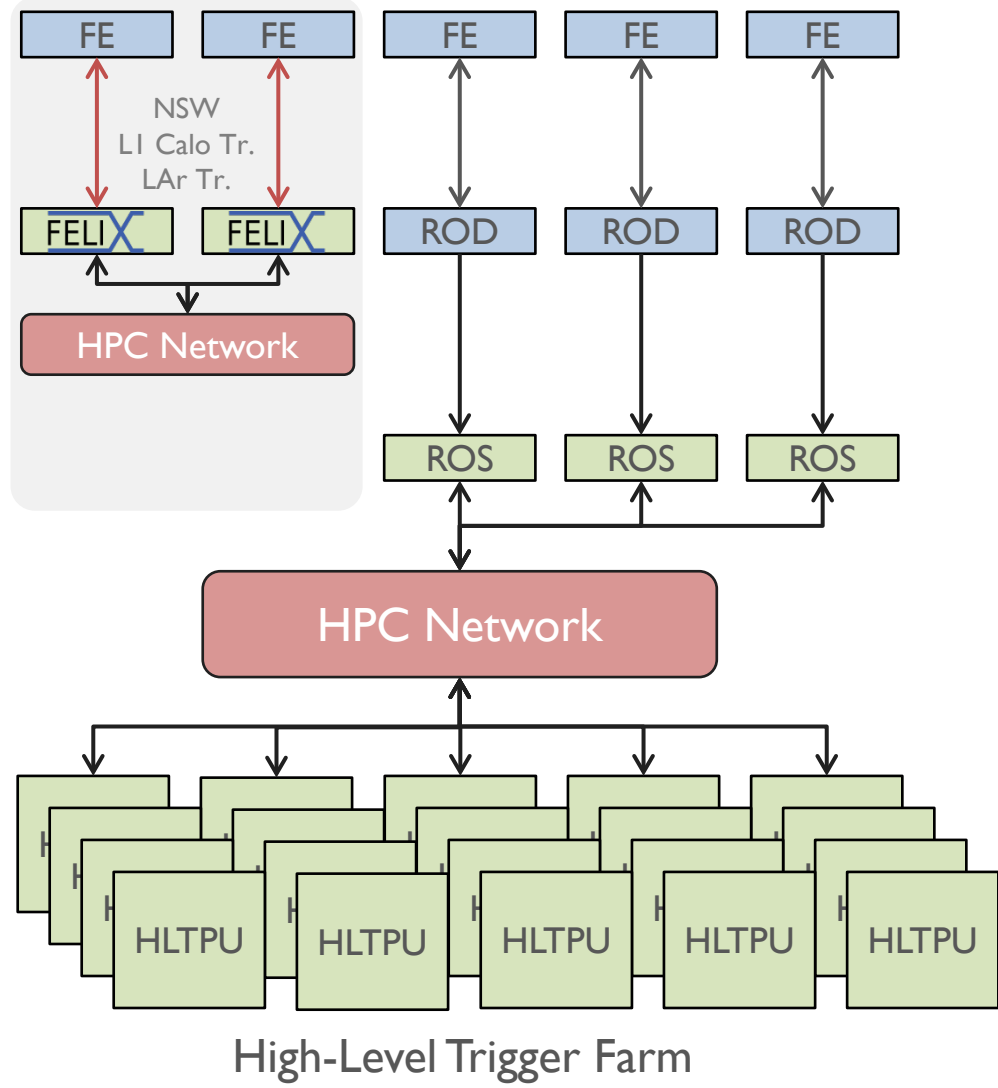
2018



ATLAS DAQ:



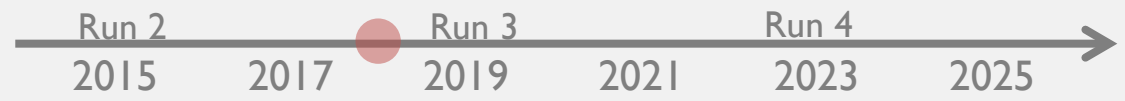
Versatile Link,
GBT
PCs
40 Gb Ethernet,
Infiniband



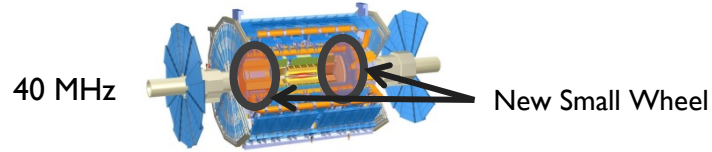
Custom electronic components

PCs (COTS)

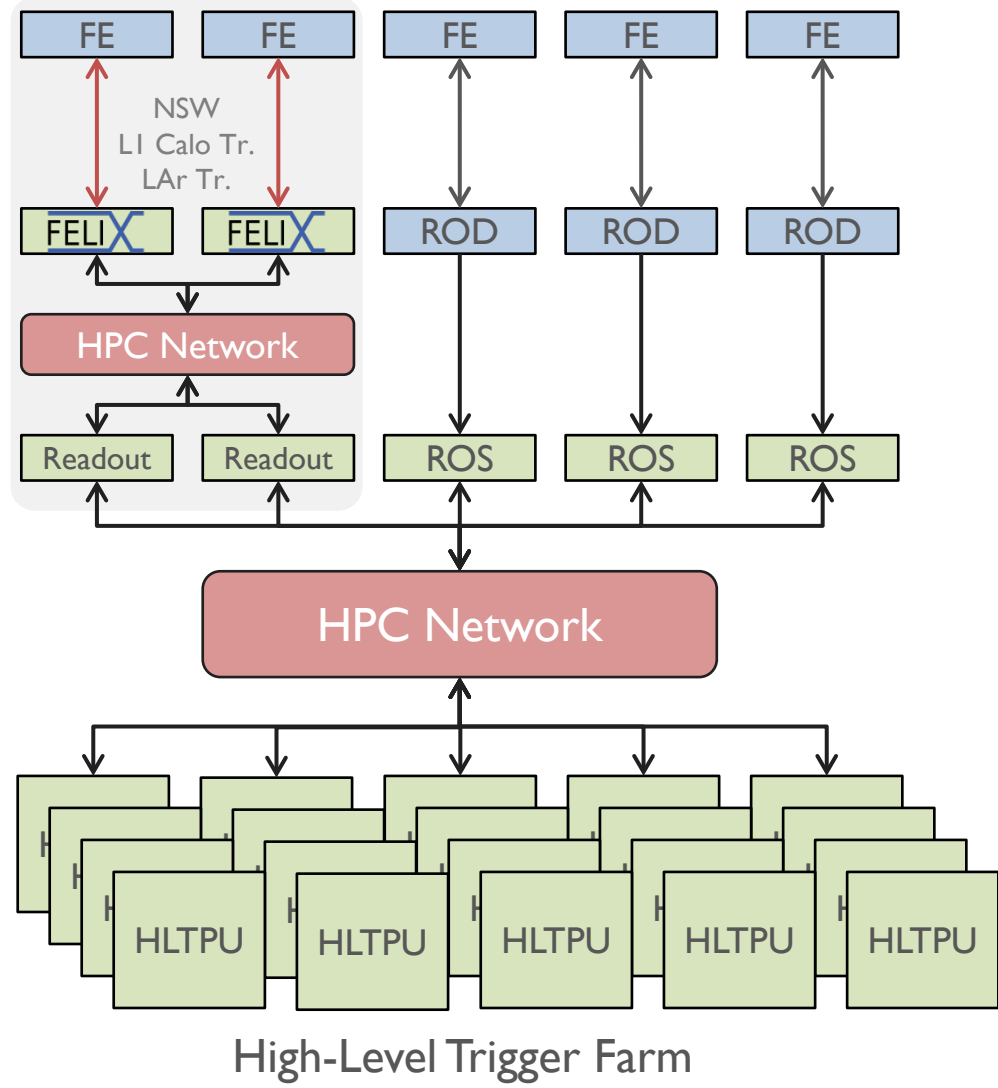
2018



ATLAS DAQ:



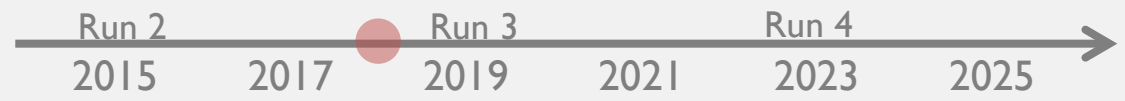
Versatile Link,
GBT
PCs
40 Gb Ethernet,
Infiniband



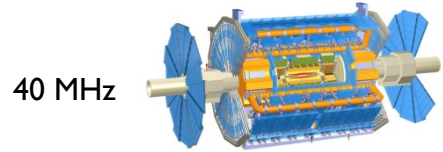
Custom electronic components

PCs (COTS)

2018

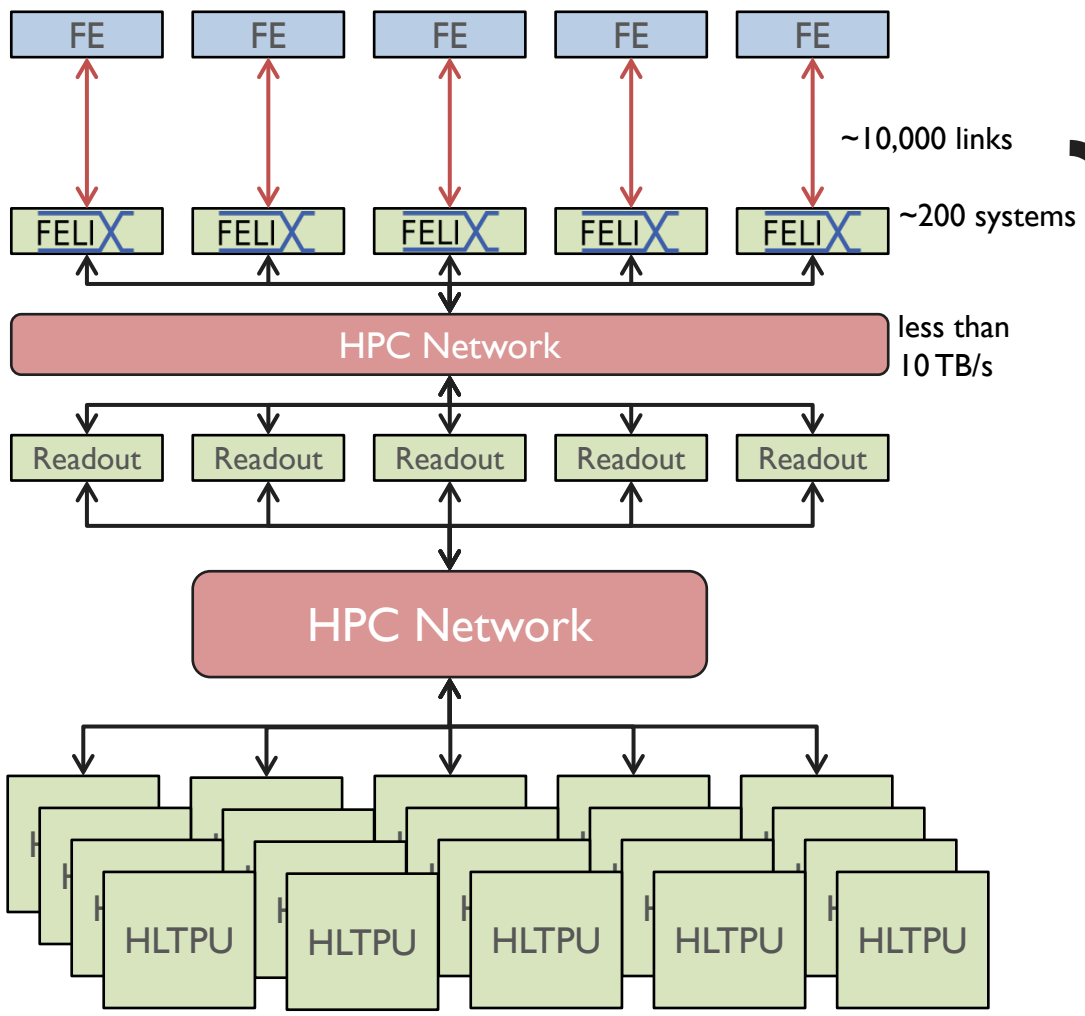


ATLAS DAQ:



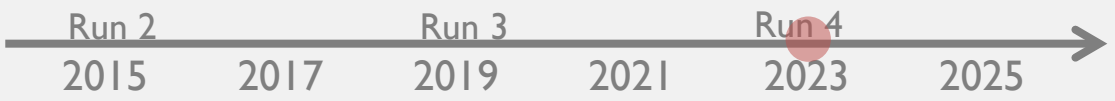
Versatile Link,
GBT, LpGBT

COTS network
technology

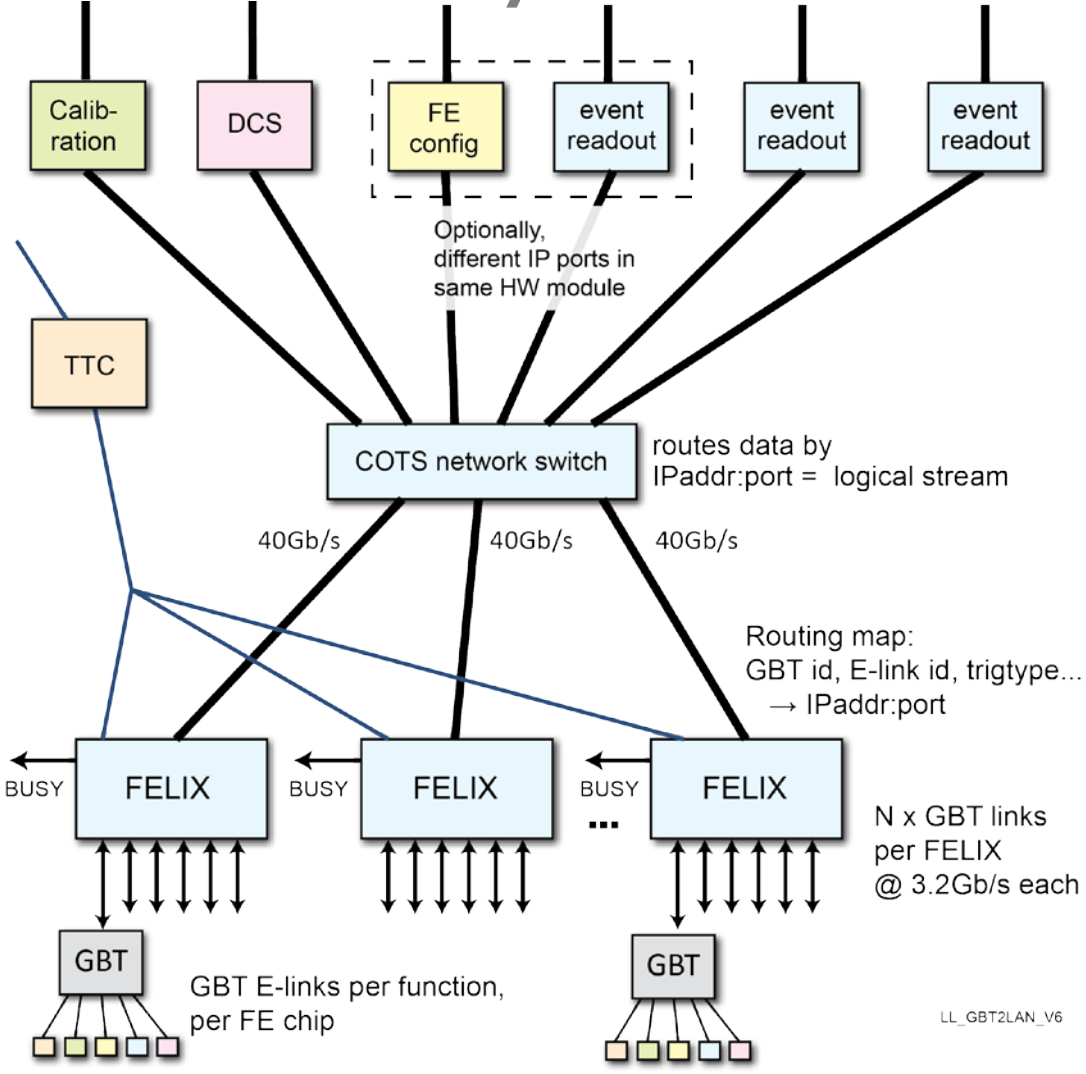


PCs
(COTS)

2023



Functionality



E-link: variable-width logical link on top GBT. Can be used to logically separate different streams on a single link.

Scalable architecture

Routing of multiple traffic types: physics events, detector control, configuration, calibration, monitoring

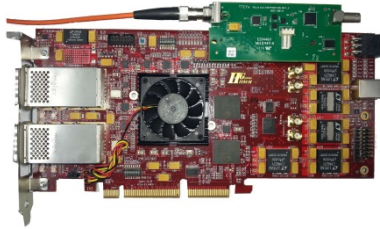
Industry standard links: data processors/handlers can be SW in PCs - Less custom electronics, more COTS components

Reconfigurable data path, multi-cast, cloning, QoS

Automatic failover and load balancing

TTC integration, LHC clock distribution

Development Platform



HiTech Global PCIe development

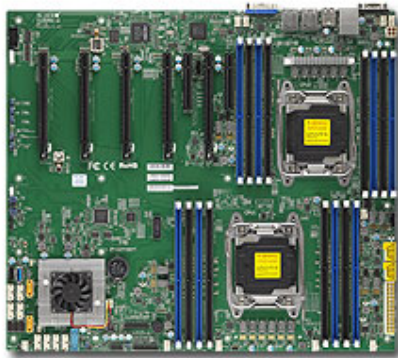
Xilinx Virtex-7

PCIe Gen-2/3 x8

24 bi-directional links

<http://hitechglobal.com/Boards/PCIE-CXP.htm>

With custom TTCfx FMC



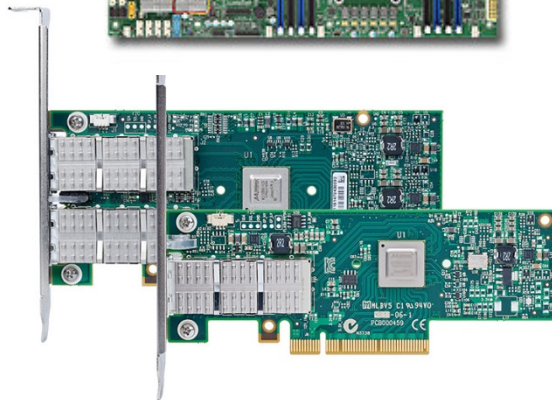
SuperMicro X10DRG-Q

2x Haswell CPU, up to 10 cores

6x PCIe Gen-3 slots

64 GB DDR4 Memory

<http://supermicro.com/products/motherboard/Xeon/C600/X10DRG-Q.cfm>

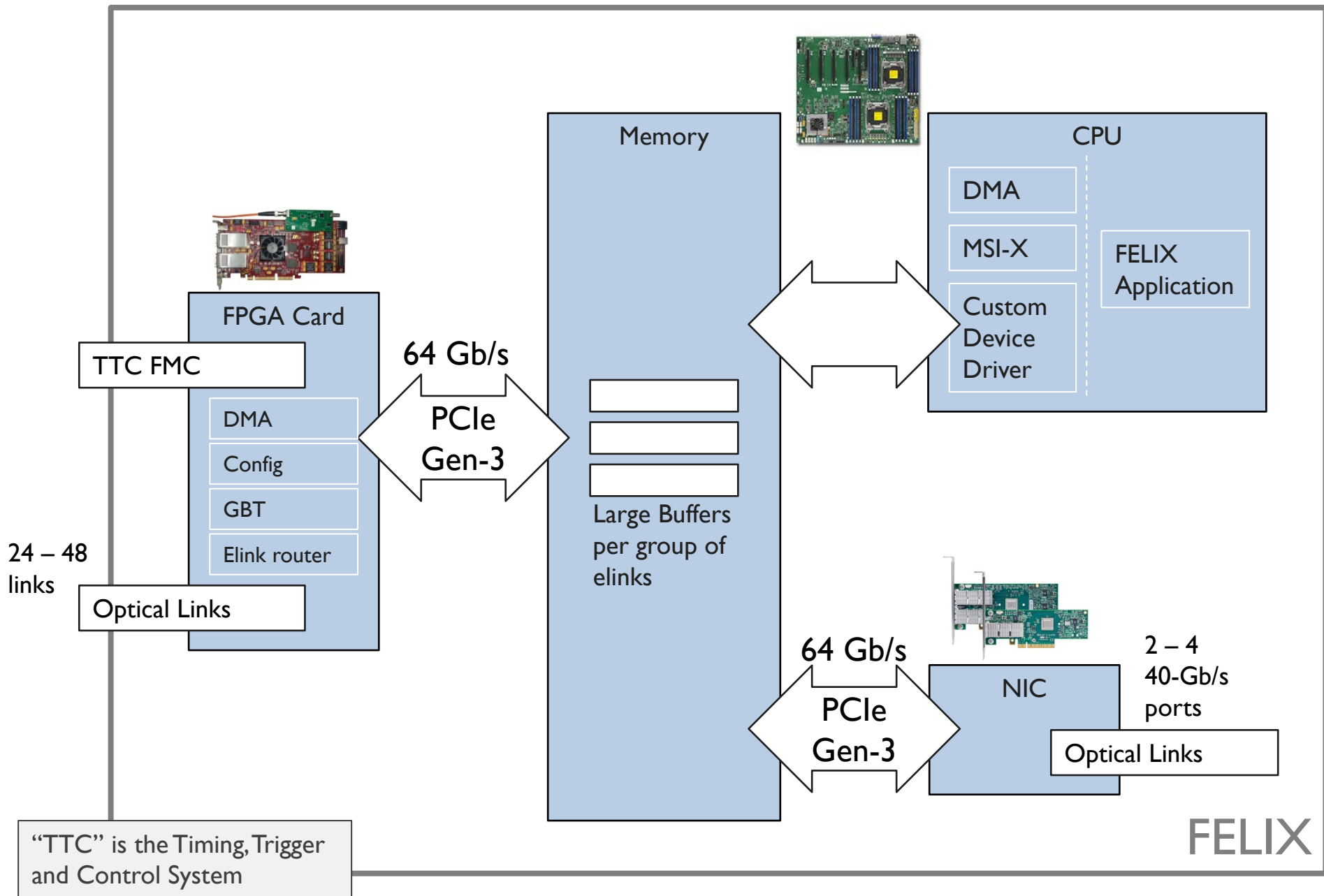


Mellanox ConnectX-3 VPI

FDR/QDR Infiniband

2x10/40 GbE

http://www.mellanox.com/page/products_dyn?product_family=119&mtag=connectx_3_vpi

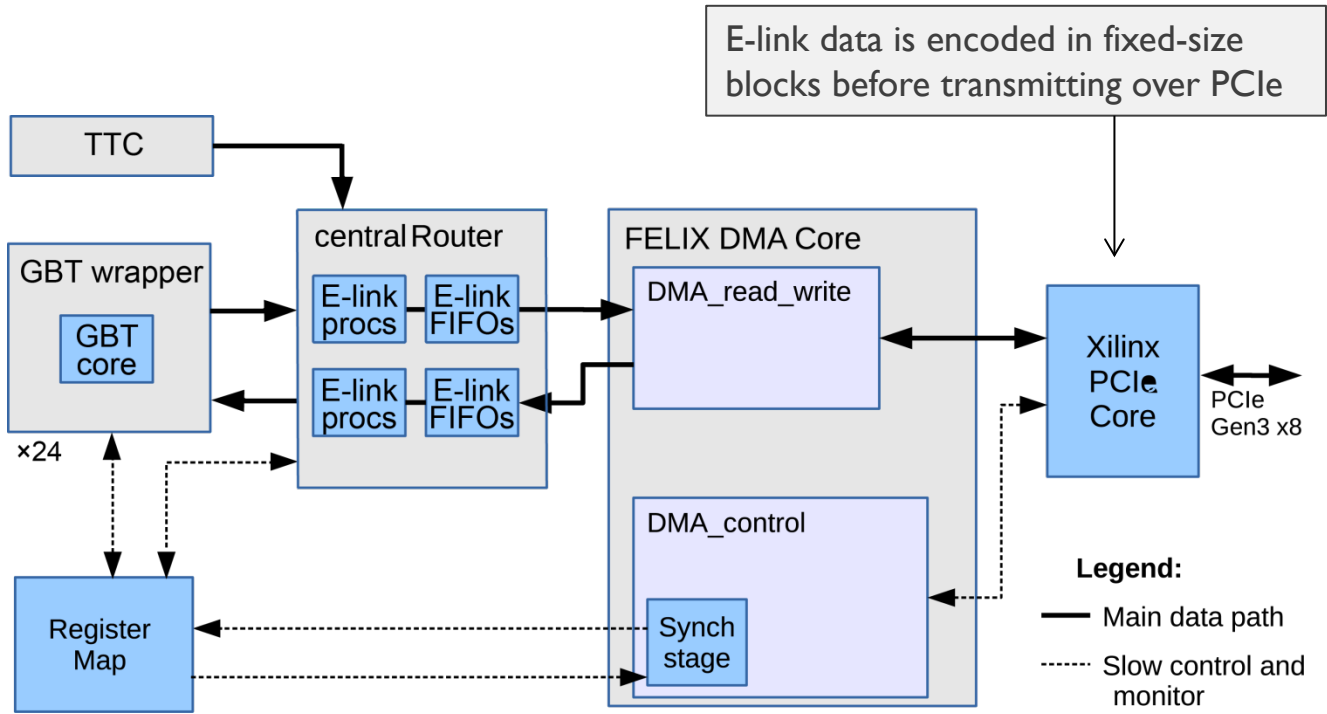


“TTC” is the Timing, Trigger and Control System

FELIX

FELIX Demonstrator System

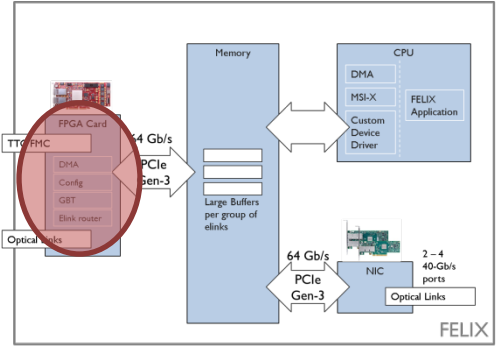
Demonstrator Firmware Design



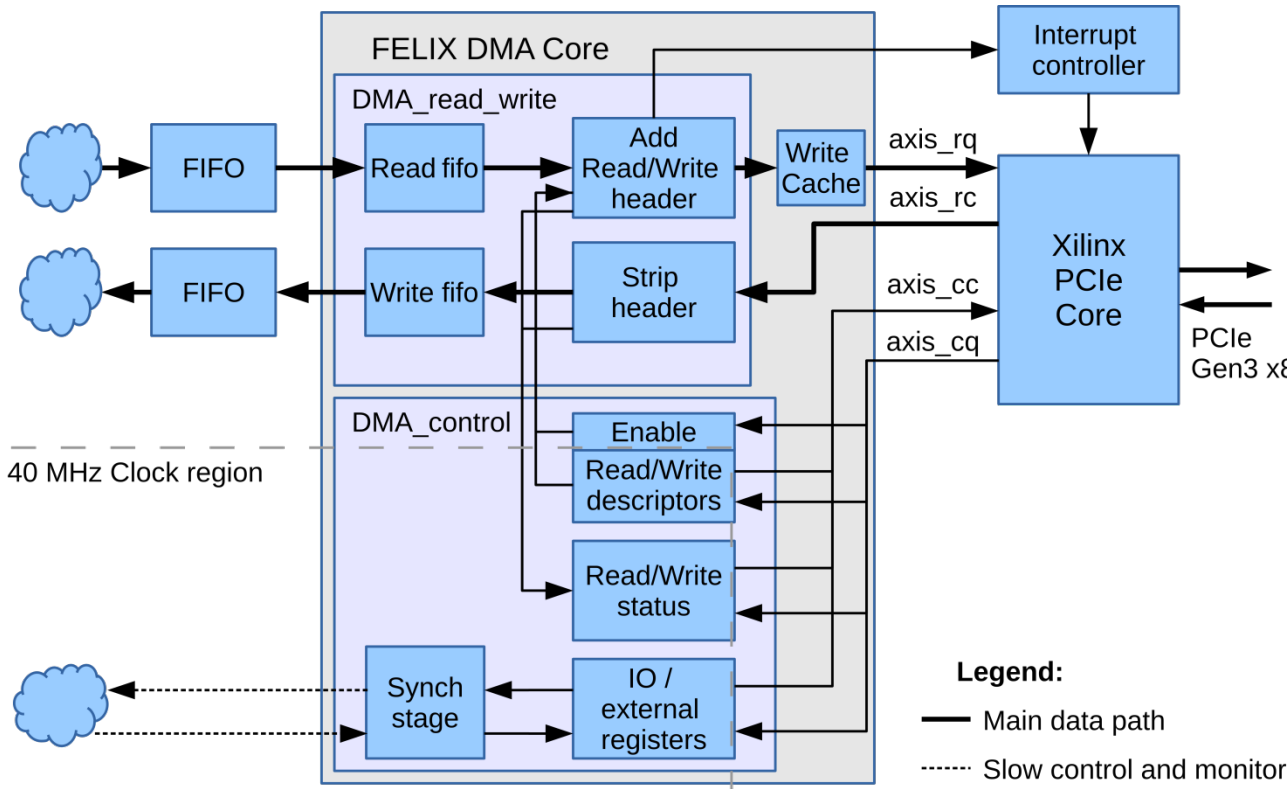
Different GBT modes (normal, wide, ...) supported

Connection to legacy TTC system via FMC

Internal data generator for testing



FELIX DMA Core



DMA interface to the Xilinx Virtex-7 PCIe Gen3 hard block

MSI-X compatible interrupt controller

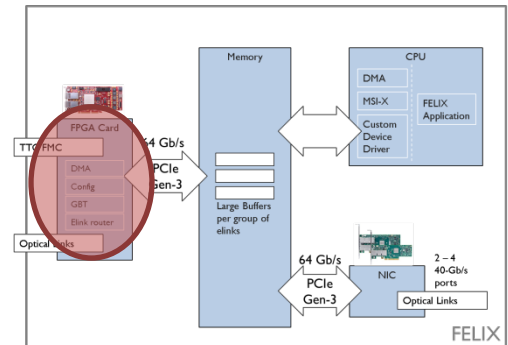
Developed at Nikhef for use in FELIX

Legend:

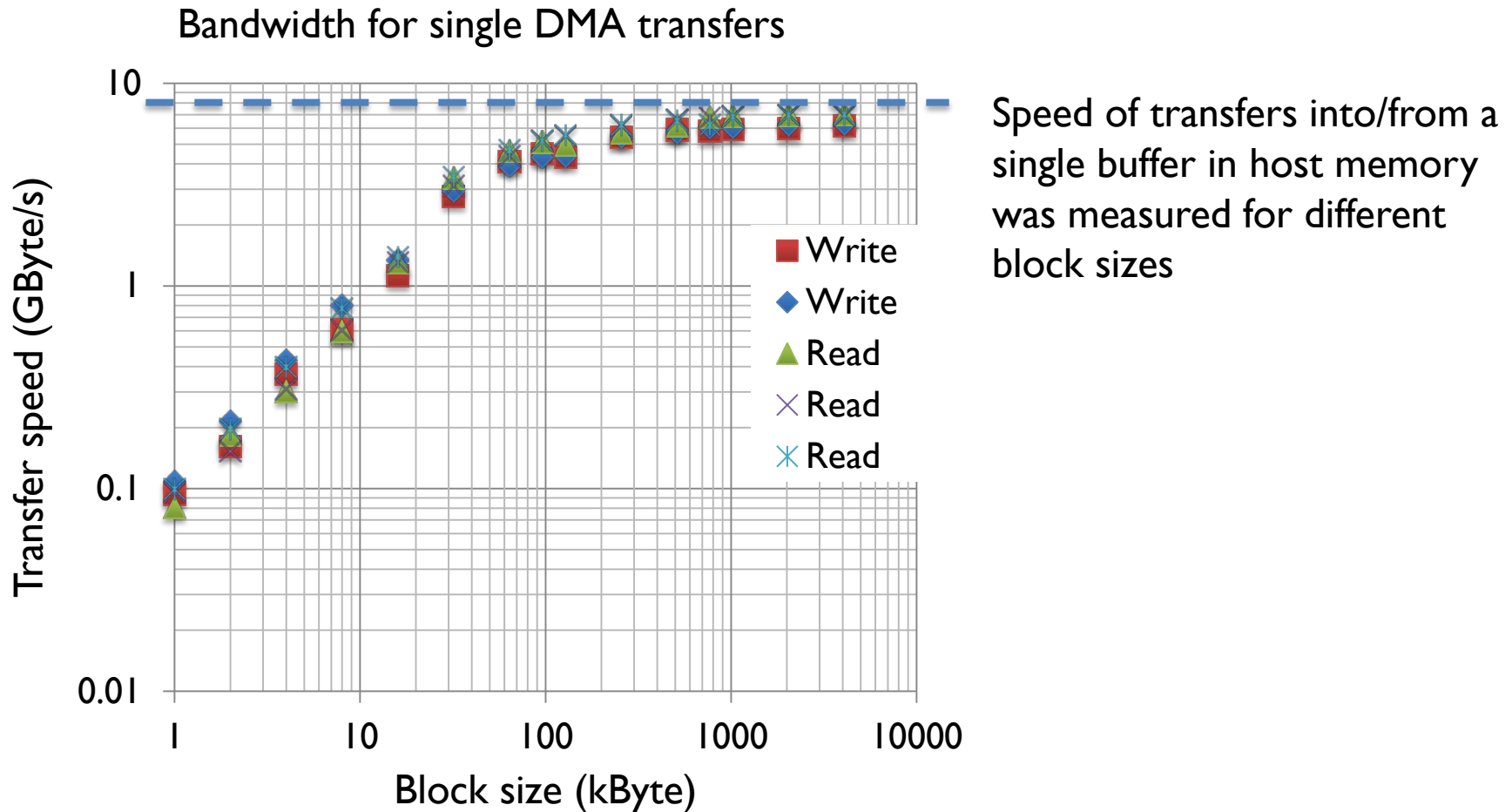
- Main data path
- Slow control and monitor

Published as OpenSource (LGPL) on OpenCores

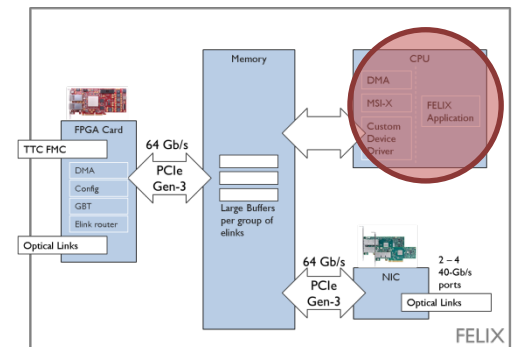
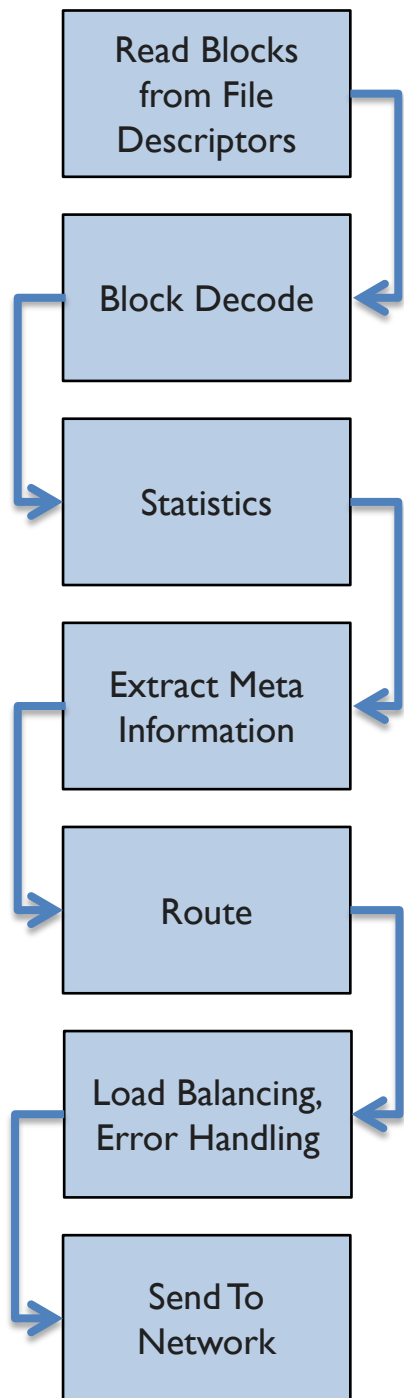
http://opencores.org/project,virtex7_pcie_dma



DMA Core Benchmarks



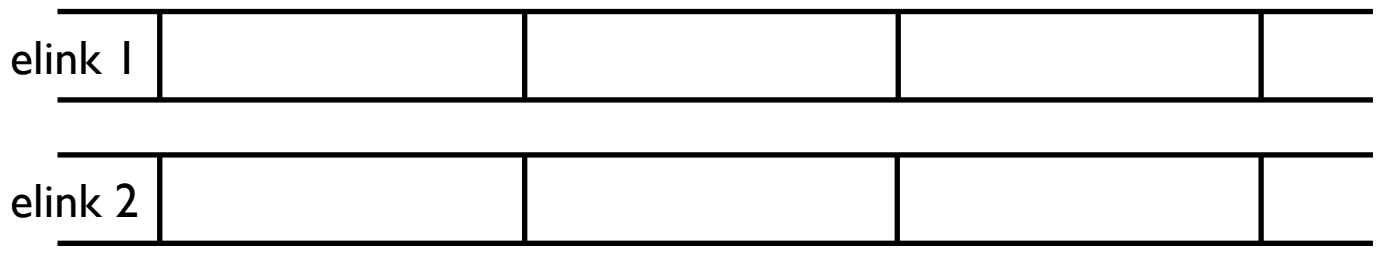
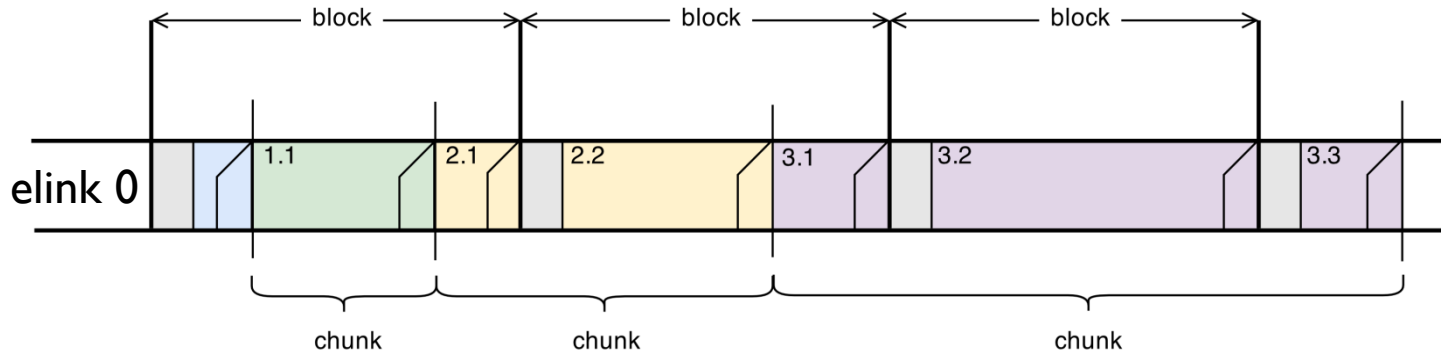
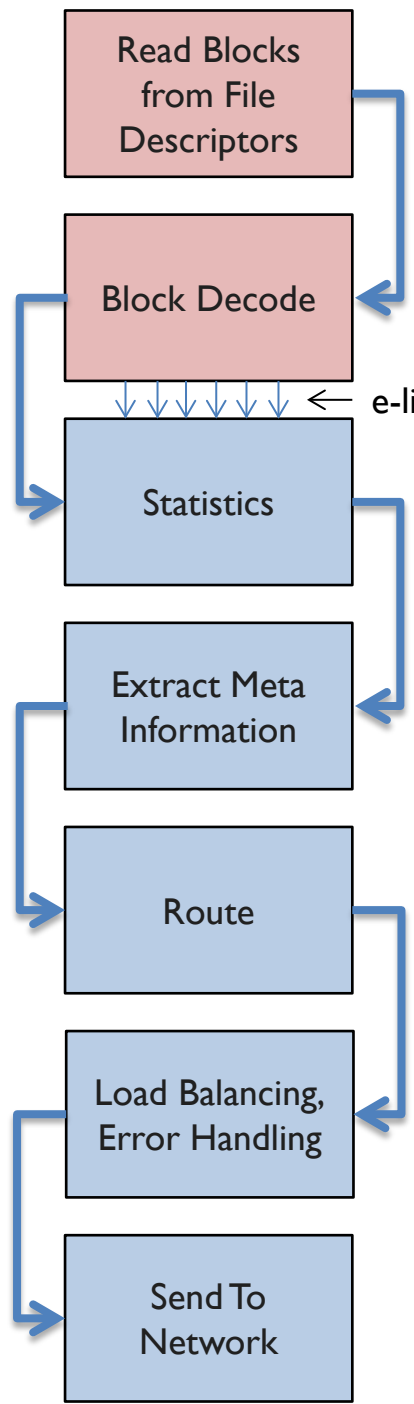
CPU Data Processing Pipeline



CPU Data Processing Pipeline

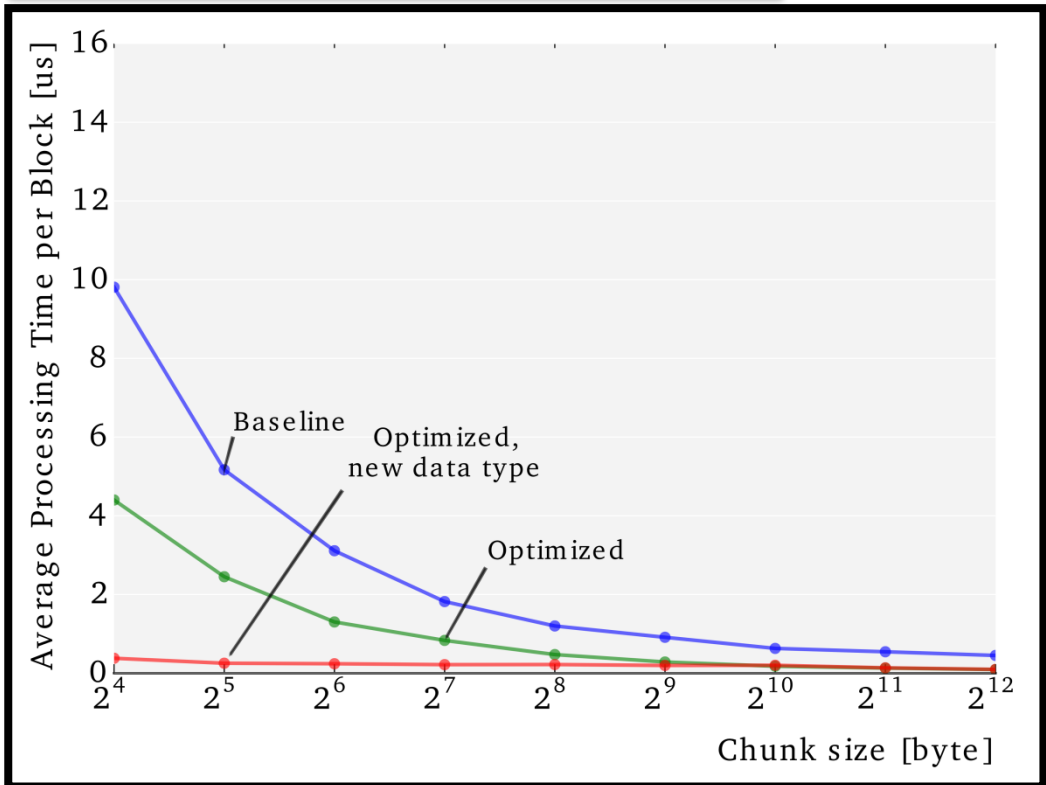
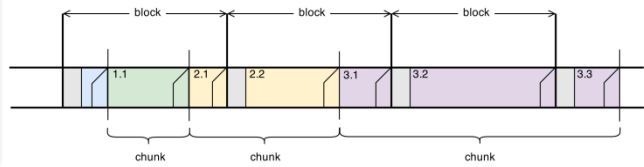
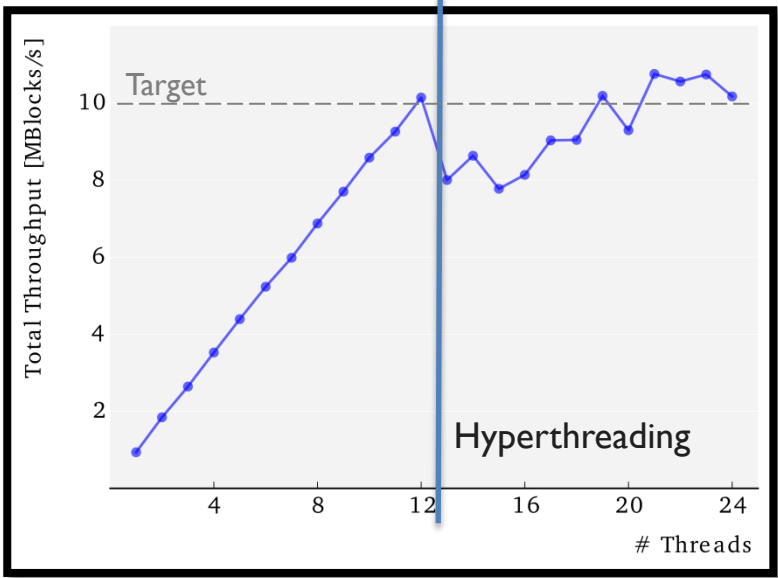
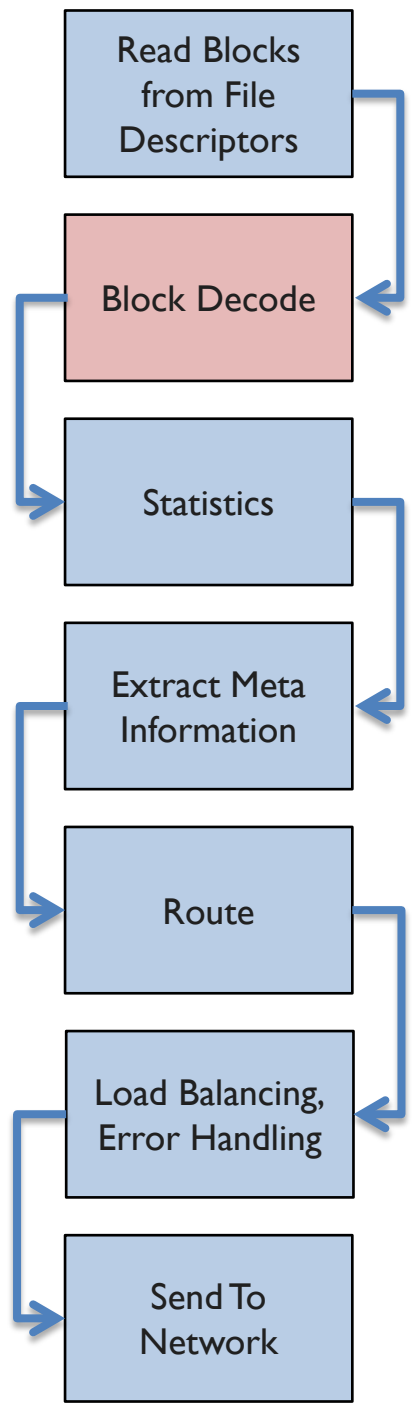
Program DMA transfers and read fixed blocks of data that have been encoded for the transfer over PCIe

Decode into variable sized chunks for transmission over network



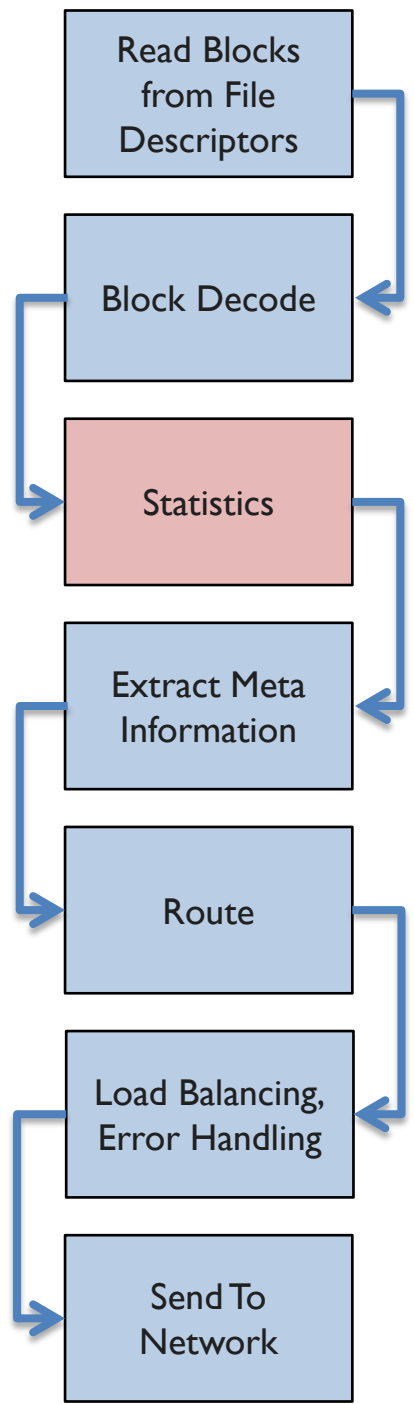
...

CPU Data Processing Pipeline



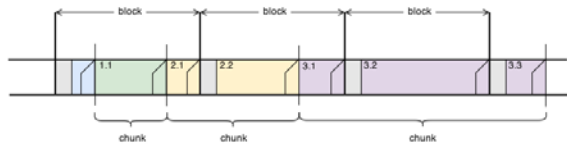
- Optimizations targeting memory throughput
- More efficient data layout
- Better use of STL containers

CPU Data Processing Pipeline

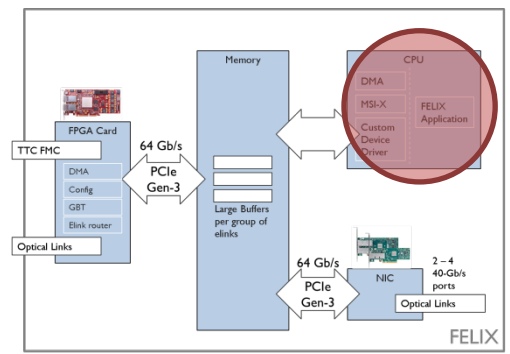


Program DMA transfers and read fixed blocks of data that have been encoded for the transfer over PCIe

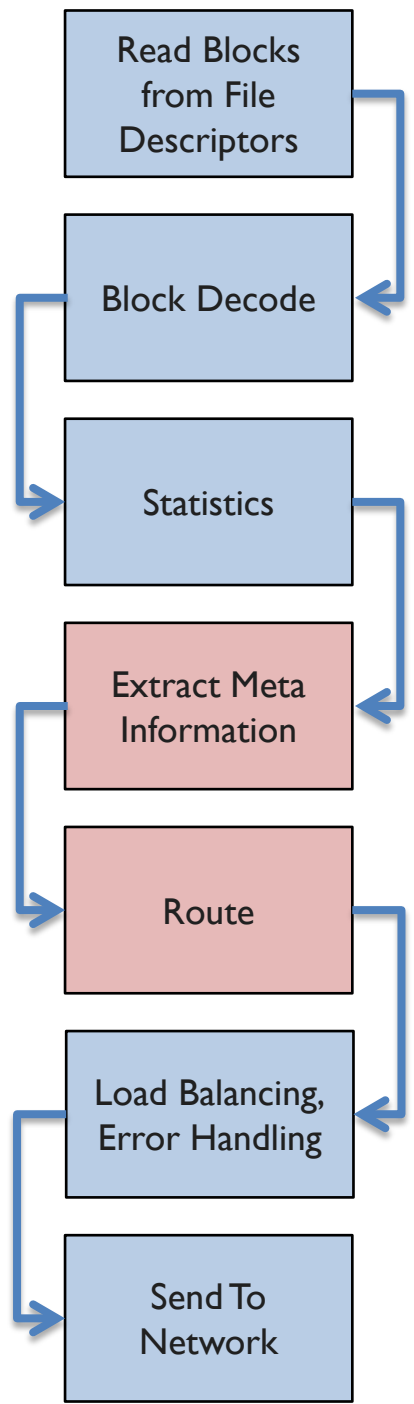
Decode into variable sized chunks for transmission over network



Count processed blocks, transfer rates, etc.

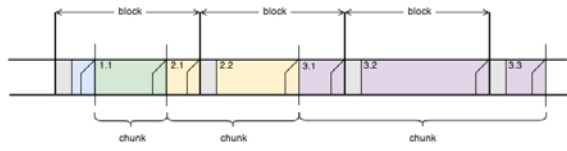


CPU Data Processing Pipeline



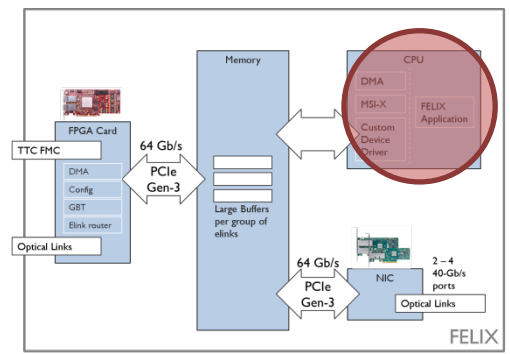
Program DMA transfers and read fixed blocks of data that have been encoded for the transfer over PCIe

Decode into variable sized chunks for transmission over network

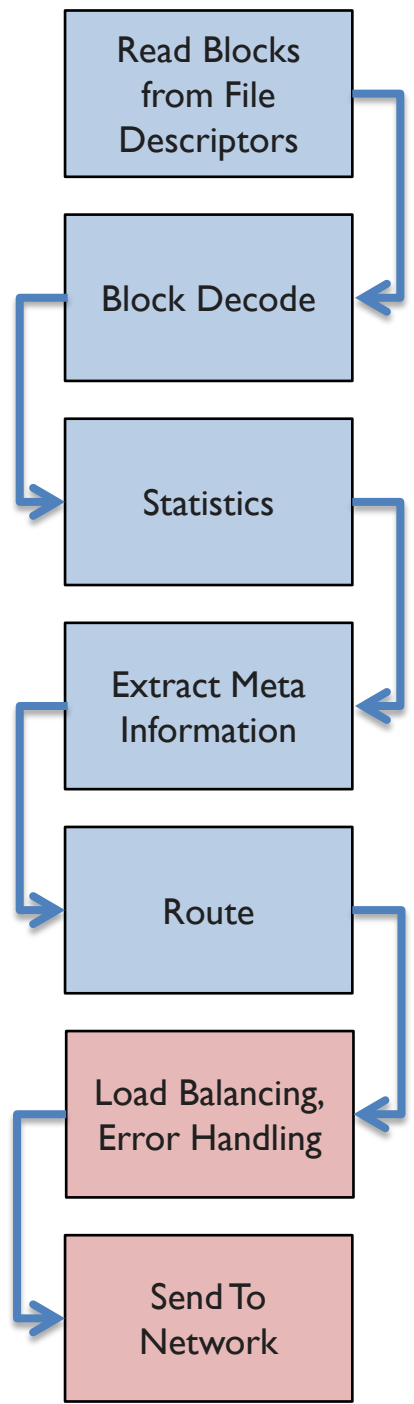


Count processed blocks, transfer rates, etc.

Metainformation, for example event ID, is extracted and matched against a routing table

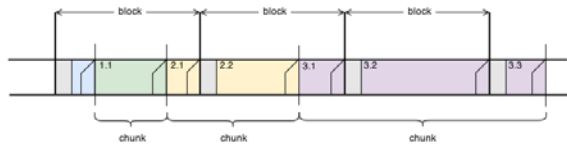


CPU Data Processing Pipeline



Program DMA transfers and read fixed blocks of data that have been encoded for the transfer over PCIe

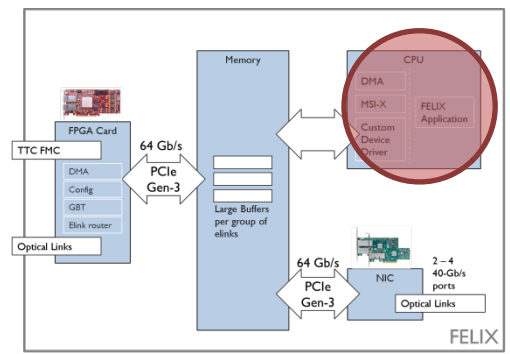
Decode into variable sized chunks for transmission over network



Count processed blocks, transfer rates, etc.

Metainformation, for example event ID, is extracted and matched against a routing table

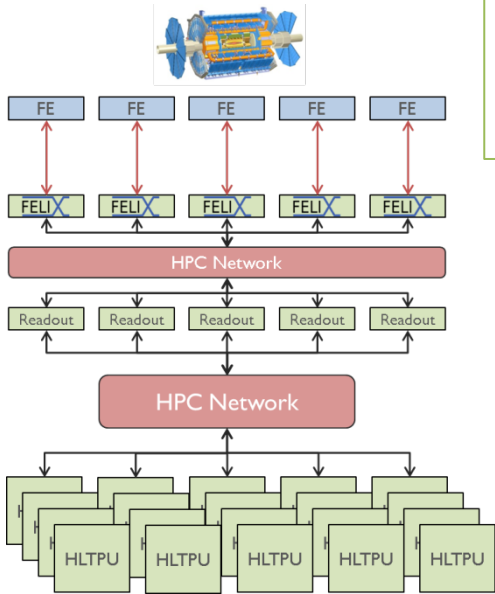
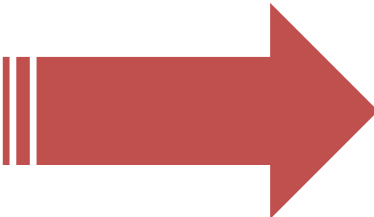
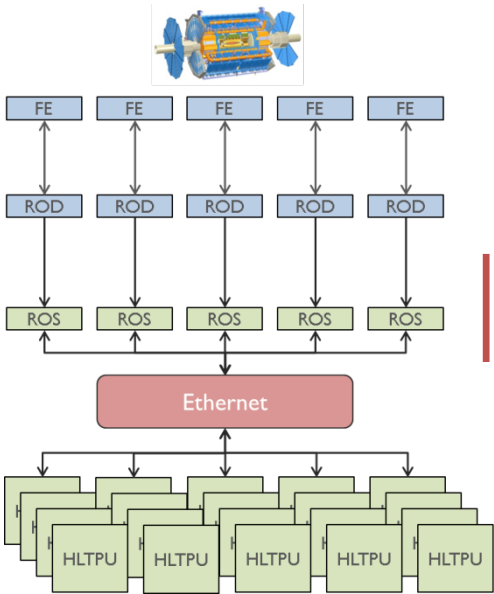
Distribute load among multiple systems
Handle automatic failover in case of system failures



Summary

FELIX

- general-purpose data-routing device
- connects ATLAS detector frontends to the ATLAS DAQ system
- load balancing, automatic failover, routing based on type of data (physics events, control, calibration, ...)



Technology demonstrator currently under development

BACKUP

Examples for Functions

Forward

*all incoming data of e-link 1
to
pc-daq-00.cern.ch*

Send a complete e-link to a destination

Forward

*Packets with `data[5] == 0xAB`
to
10.0.0.1:1234*

Send only certain packets, marked by a flag in the data stream, to a certain destination

Forward

*a random 10% sample of e-link 1
to
my-monitoring-server*

Send a certain percentage of incoming data to a monitoring system

Forward

*all incoming data of e-link 2
to any of
{pc-daq-00, pc-daq-01, pc-daq-02}*

Load balancing: Distribute load among multiple systems

GBT Firmware

