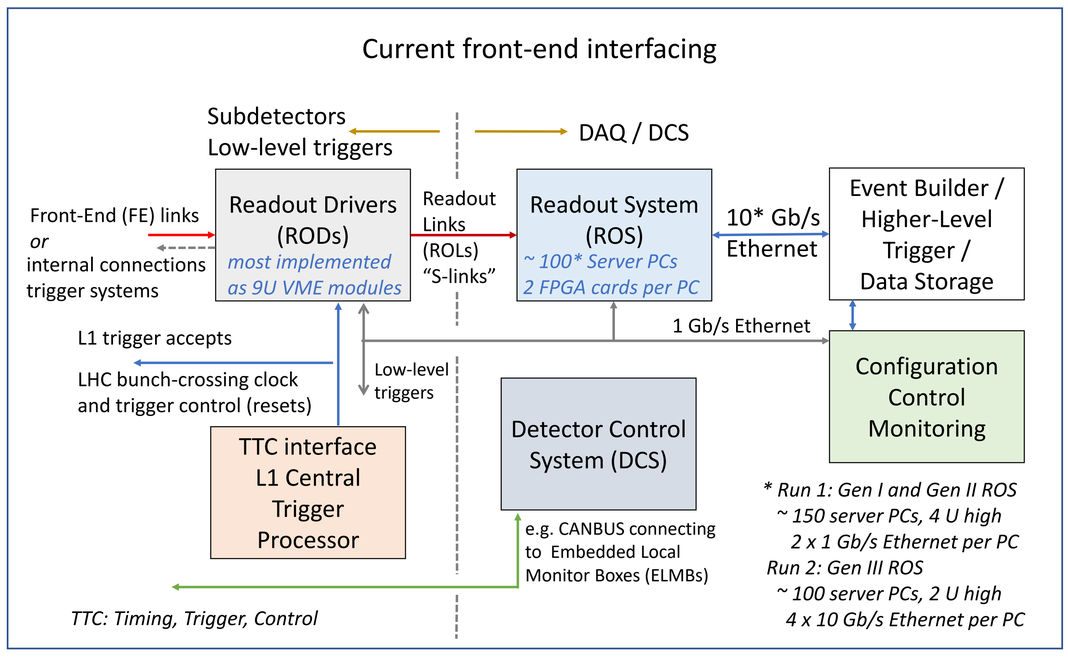
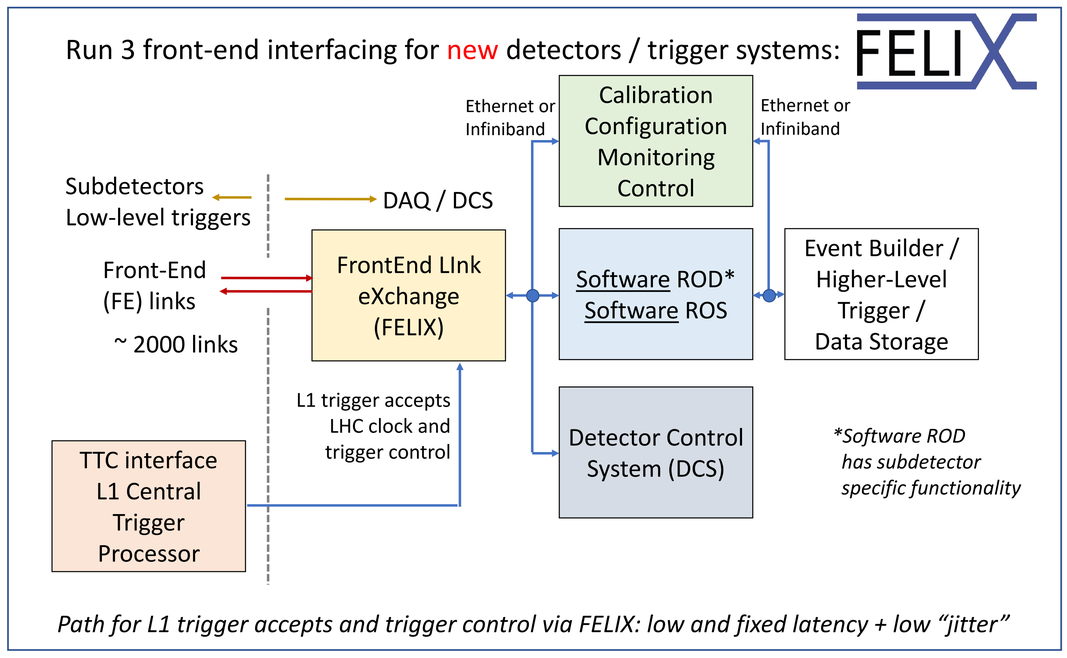
FELIX: the new detector interface for the ATLAS experiment

Overview

Interfacing to new subdetectors and trigger systems to be installed in the ATLAS experiment during the next long shutdown of the LHC, starting at the end of next year, will be done by means of the **FELIX** (Front-End LInk eXchange) system, instead of the current interfacing with ReadOut Drivers (RODs) and via dedicated links of the Timing, Trigger and Control (TTC) system and of the Detector Control System. FELIX connects to subdetectors and trigger systems via custom optical links, to **general purpose commodity networking** and to the TTC system.



FELIX is transparent for the data it receives, with the exception of routing information included in the data, and acts as a **heterogeneous stateless switch**. Data processing functionality of the RODs, typically currently implemented in firmware, is moved to **"software RODs"** running on server PCs connected to the general purpose network. The LHC bunch-crossing clock, trigger accepts and counter resets as provided by the TTC system are forwarded with low and reproducable latency and with low jitter to front-end electronics. Relevant trigger information is also sent to the software RODs.



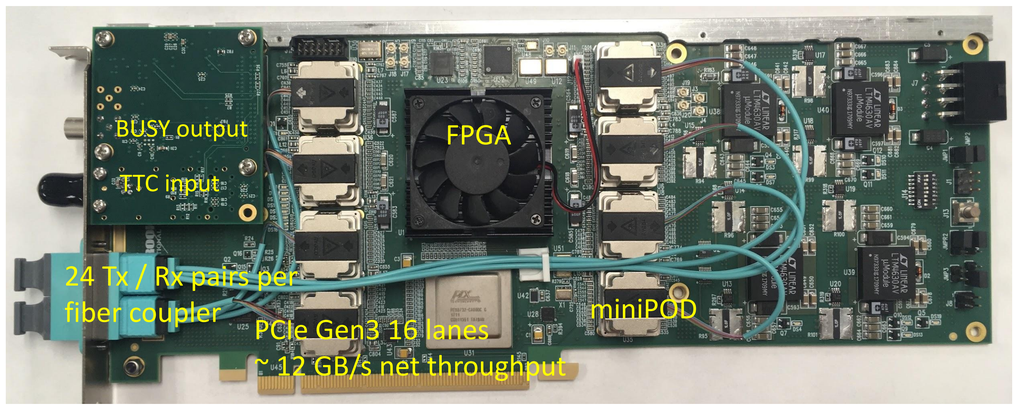
Protocol on Front-End (FE) links for FELIX

Downlinks, towards detector or trigger system, use the standard protocol of the GBTx ASIC, developed by CERN. Uplinks, towards FELIX, use GBT or a custom protocol, "FULL mode". The downlinks provide fixed and reproducible latency. The standard GBT protocol, with a bandwidth of 4.8 Gb/s, provides a configurable aggregation of independent 80, 160 and 320 Mb/s serial links, known as E-links.

Event data transferred across E-links is 8B/10B coded with control symbols indicating event fragment boundaries, HDLC or no coding is possible for other data. **"FULL mode"** links are 9.6 Gb/s serial links used for transferring a single 8B/10B coded event data stream, control symbols again indicate event fragment boundaries. XON-XOFF signaling is implemented for "FULL mode" links with the help of control symbols.

Hardware

FELIX will be built from **FPGA PCIe cards**, server PCs and Network Interface Cards (NICs). The baseline choice of the FPGA card, known as the FLX-712, has a 16-lane PCIe Gen3 interface, a **Xilinx Kintex Ultrascale KU115 FPGA**, 8 miniPODs interfacing to 48 bi-directional links, a TTC interface and a LEMO connector for a "BUSY" signal output. The server PCs run Linux and have a rack mountable chassis.



The FLX-712 FELIX FPGA PCIe card. The on-board FLASH memory stores up to 4 different FPGA configurations

Either one FLX-712 or two FLX-712s will be installed per PC, each connecting to either 12, 24 or 48 links, depending on the type of front-end electronics. **Commercially available Xilinx VC709 boards,** equipped with a small mezzanine, the TTCfx, for interfacing to the TTC system, are used for **testing and development.** A VC709 interfaces to four bi-directional links and has an 8-lane PCIe Gen3 interface.

	# FE Links	# PCs	# of FLX-712 per PC	# of links per FLX-712	Link protocol	Remarks
LAr calorimeter on-detector trigger boards	*620*	*7*	*2*	*48*	*GBT*	*No event data, no throughput requirements*
Calorimeter off-detector trigger electronics uplinks / downlinks	*295 / 71*	*25*	*1*	*12*	*FULL mode / GBT*	*Throughput PCIe i/f: ~throughput of 12 9.6 Gb/s links, 2 x 100GbE*
New Small Wheel muon detectors	*1152*	*30*	*2*	*24*	*GBT*	*2 x 25 GbE*
BIS78 muon detectors	*16*	*1*	*1*	*16*	*GBT*	*2 x 25 GbE*

-> 63 PCs, 100 FPGA boards

Data Transfers

The **data received** via the FE links is decoded and buffered in the FPGA per E-link or per FULL mode link and is transferred under DMA control via the PCIe interface in **blocks with a fixed size of 1 kByte**. Each block has a 32-bit header with a 16-bit fixed pattern and with a link identifier and a sequence number. The blocks are transferred into a contiguous area, functioning as a circular buffer, in the main memory of the PC. The DMA runs continuously, thereby eliminating DMA setup overheads and achieving high throughput (**about 12 GB/s for the 16-lane interface of the FLX-712**). Transfers will be halted if data that is not yet processed by the software would be overwritten, but this should not occur in practice. Event fragments or other types of data arriving via the FE links are referred to as **"chunks"** and can have an arbitrary size. Trailers are therefore appended to the data with information on chunk length etc. as specified in the table.

For forwarding data packets via an FE link the FelixCore application initiates for each packet a DMA transfer after appending it, in a contiguous memory area, to a header with a link identifier, the data is encoded by the FPGA.

Sub-chunk trailer read as u_int_16_t (u_short) (16-bit word) from memory

Payloads of the blocks	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	*Chunks are always aligned on a 16-bit boundary. A padding byte is added at the end if the number of bytes is odd*
are organized as chunks	Type			T	E	R	sub-chunk length in bytes										
that may extend over	0	0	0	<- signals fill pattern caused by timeout, all other bits are also 0													
several blocks. Each	0	0	1	<- first part of a chunk consisting of more than one part													
chunk has a 16-bit trailer,	0	1	0	<- last part of a chunk consisting of more than one part													
if it extends over more	0	1	1	<- chunk consisting of one part													
than one block then	1	0	0	<- not first or last part of a chunk consisting of more than one part													
each part of the chunk	1	0	1	<- E-link timeout													
("sub-chunks") has	1	1	0	<- reserved													
its own 16-bit trailer	1	1	1	<- out of band													

T: Truncation flag, 1 if chunk was truncated E: Error flag, 1 if an error occurred R: Reserved

Software

The FelixCore application, the multi-threaded application running on the server PCs of the FELIX system, controls the FPGA cards via PCIe registers. These are accessible via memory mapped I/O, set up with the help of a dedicated driver. The contiguous memory needed for data transferred under DMA control is allocated by another driver.

FelixCore reconstructs complete chunks from the data in the circular buffers and forwards these via the network, using **NetIO**, a network protocol agnostic software layer, to any destination that registered with FelixCore for receiving the data. FelixCore also forwards data, received from the network by means of NetIO to the FPGAs.

Towards deployment

Firmware and software development are in an **advanced state**, FLX-712 production this autumn. **Installation and commissioning early in 2019**.

Outlook

Phase-II upgrade (2024-2026): all FE interfacing will be done with FELIX.

Information

http://cern.ch/atlas-project-felix

Use outside ATLAS

protoDUNE-SP

http://cenf-dune-proto.web.cern.ch/duneexp

Serguei Kolos[1], Jos Vermeulen[2], Nikolina Ilic[3]
on behalf of the ATLAS TDAQ Collaboration. CHEP 2018, Sofia
[1]University of California, Irvine, USA , [2]Nikhef, Amsterdam Netherlands, [3]Radboud University, Nijmegen, Netherlands