FELIX Phase II

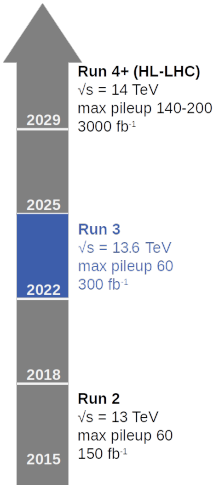
The ATLAS readout system for LHC Run 4

Frans Schreuder f.schreuder@nikhef.nl

On behalf of the ATLAS TDAQ collaboration

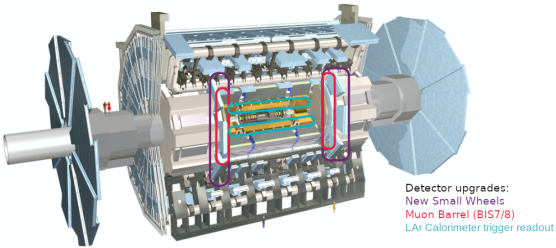
September 4-8, 2023

ATLAS Detector - Upgrades for Run 3



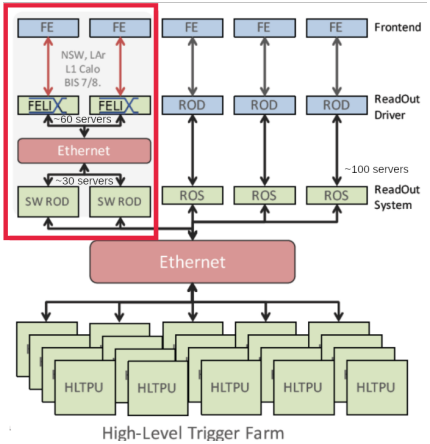
In Run 3:

- Similar conditions to Run 2, higher energy: 13.6 TeV
- New detector and trigger system to improve background rejection
- New readout system: FELIX and Software ROD for new systems



ATLAS Trigger and DAQ in Run 3

New readout chain for upgraded systems



Front-End Link eXchange (FELIX)

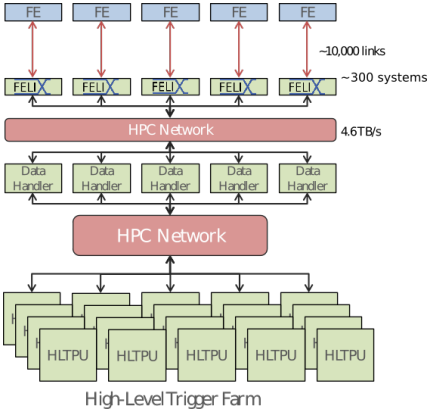
- Custom FPGA based PCIe cards installed in commercial servers
- Readout, trigger, clock distribution, Slow Control, BUSY
- 100 cards, 60 host servers
- Removes one layer of custom electronics

Software Readout Driver (SW ROD)

- Builds and aggregates events, detector-specific data processing
- 30 servers

ATLAS Trigger and DAQ in Run 4

Phase II Upgrade, 2029-2032



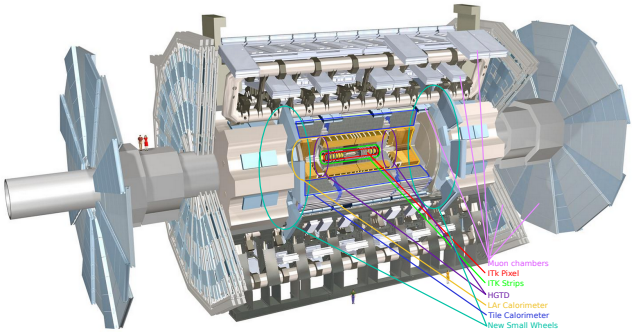
Front-End Link eXchange (FELIX)

- Implemented for all subdetector systems
- x10 low level trigger rate (1 MHz)
- x3 interactions per bunch crossing (200)
- x20 readout data rate (4.6 TB/s)

Data Handlers

- Evolution of SW ROD with similar functionality
- Interfaces to the high level trigger farm

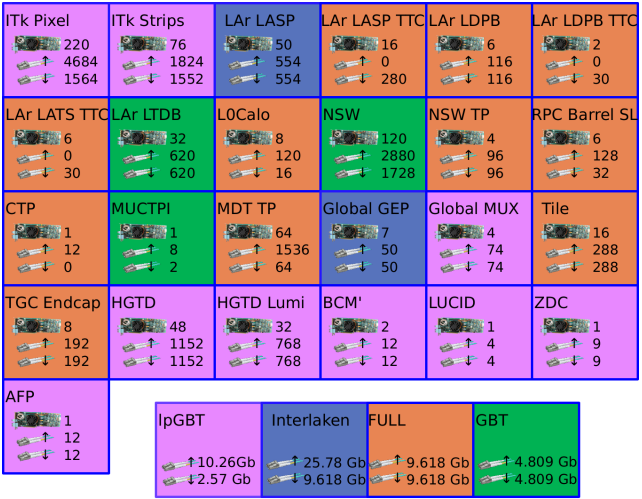
Atlas Phase II upgrade - LHC Run 4



- Upgraded Trigger and Data Acquisition
- New Inner Tracker
  - ITk Pixel
  - ITk Strips
- High Granularity Timing Detector (HGTD)
- New Muon chambers
  - Inner barrel region, new RPCs, sMDTs and TGCs
- Electronics upgrades, including LAr, Tile and Muon
- Additional small upgrades

[1] T. Affolder, ATLAS Phase II Upgrade

ATLAS Subdetectors - Link protocols

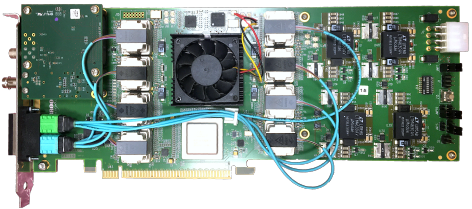


- Four link protocols, various new protocols on top
- lpGBT and GBT
  - Radiation hard ASIC for on-detector electronics
  - Logical links (E-Links)
    - Aurora, 8b10b, 6b8b, HDLC, TTC [1], Endeavour (custom protocol for ITk)
- FULL (8b10b) and Interlaken
  - FPGA $\rightarrow$ FELIX communication
  - Distribute Timing, trigger and control over 8b10b link

[1] TTC: Trigger, timing and control

FELIX Phase I system (Run 3)

- Used extensively for Phase II development
  - FELIX firmware and software development
  - Subdetector integration



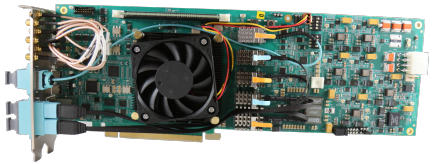
The FLX-712 card

- FPGA: Xilinx Kintex UltraScale XCKU115
- 4 or 8 MiniPODs to support 24 or 48 bidirectional optical links
- 16-lane PCIe Gen3 interface (120 Gb/s)
- Interface for Trigger/Timing/Control (TTC) and BUSY

Server

- Intel Xeon E5-1660 v4 @ 3.2GHz
- 32 GB DDR4 2667 MHz memory
- Mellanox Connect-X 25/100 GbE

FELIX Phase II system (Run 4)



- Design of an alternative card with AMD Versal Premium VP1552 in progress: FLX-155
  - PCIe Gen5x16 interface (482 Gb/s)
    - 2x 8 lanes bifurcated
  - Up to 48 bidirectional optical links

The FLX-182 card

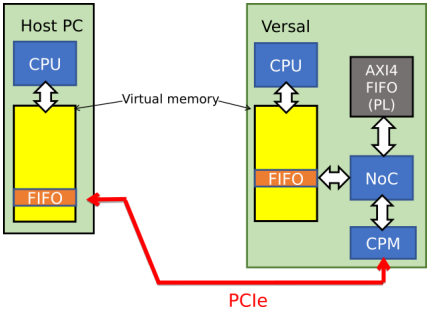
- FPGA: AMD Versal Prime VM1802
- 4 Firefly transceivers to support 24 bidirectional optical links
  - Up to 25 Gb/s per link
- 1 Firefly for LTI/TTC interface
  - New protocol for Timing, Trigger and Control
  - 100Gb/s Ethernet or White Rabbit are optional
- 16-lane PCIe Gen4 interface (240 Gb/s)
  - 2x 8 lanes bifurcated

Server hardware for testing

- AMD Epyc 9004 CPU (Genoa)
- 96 GB DDR5
- Dual 200 Gb/s Ethernet on PCIe Gen5

The FLX-182 card

Accessing Linux running on the SoC



- FELIX dataflow is handled by the Versal programmable logic (PL)
- The Versal Processing system, running PetaLinux, handles tasks such as
  - Internal temperature and voltage monitoring
  - Updating the FLX-182's flash memory
  - A built-in self test
- PetaLinux can be accessed through a direct 1 GbE connection via the bracket, or using a virtual network connection over PCIe [1]
  - A virtual network device is exposed on the FLX-182's PCIe device to the host
  - In the programmable logic, this network connects through AXI4 to the processing system
  - A virtual network device is created in Petalinux
  - This allows control using SSH, HTTPS, etc.

[1] E. Zhivun, FELIX Versal example drivers

The FLX-182 card

Built-in self test

- A Built-in self test (BIST) for the FLX-182 card was developed
- The BIST runs on the Versal processing system
- Capable of testing and monitoring the card, also for production tests
  - Monitoring all I2C peripherals on the board
  - Test fiber and PCIe links using fiber- and PCIe loopback (Eye scan)
  - Generate a test report which can be automatically published in a database

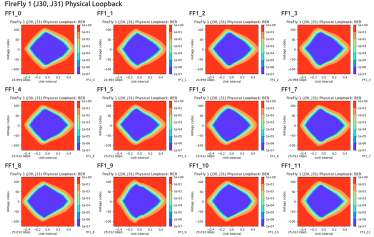
Temperature monitor

Input	Value
max_max	43.1 C
min	32.3 C
min_min	34.0 C
temp	35.9 C

Voltage monitor

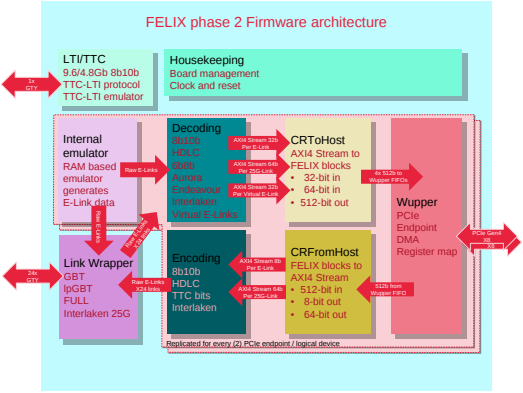
Input	Value
gty_avcc_103	0.876 V
gty_avcc_104	0.875 V
gty_avcc_105	0.876 V
gty_avcc_106	0.876 V
gty_avcc_200	0.879 V
gty_avcc_201	0.879 V

FireFly 1 (J30, J31) Physical Loopback



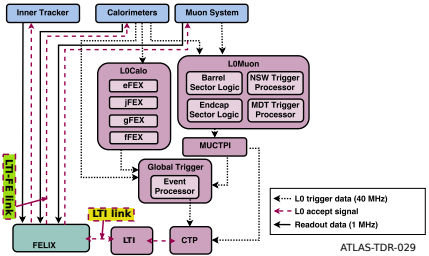
Firmware upgrade for Phase II / Run 4

- PCIe DMA Core (Wupper)
  - Adds support for Gen4 and Gen5
  - Two PCIe endpoints per card
  - 5 ToHost, 2 FromHost DMA channels allowing multithreading and separate streams
- More modular design, separation between routing and encoding / decoding
  - AXI4 stream to communicate between blocks
- Encoding / Decoding / Link wrapper
  - Support for lpGBT and 25 Gb/s Interlaken
  - lpGBT E-Links support Aurora, 6b8b, 8b10b, HDLC, Endeavour
- New interface for trigger / timing control



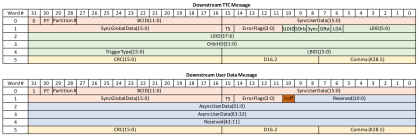
[1] F. Schreuder, FELIX Phase2 firmware specs

FELIX Timing, Trigger and Control



LTI-FE data format

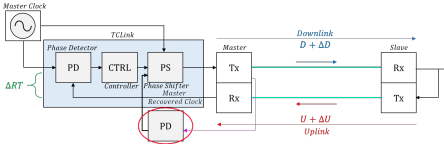
- 9.61896 Gb/s 8b10b



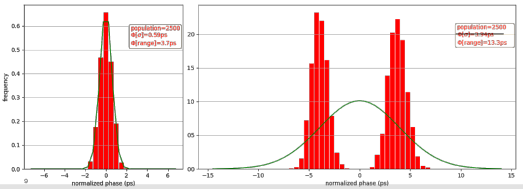
- LTI distributes the LHC clock, along with trigger and control information to FELIX
  - Clock is compensated for long term phase drift with TCLink
- FELIX redistributes the clock to Frontend Electronics
  - Using GBT, lpGBT or new LTI-FE format
  - Clock phase is compensated with TCLink
- An LTI format data emulator based on an off the shelf Zynq Ultrascale+ board was developed for lab set-ups.

Timing Compensated Link - TCLink

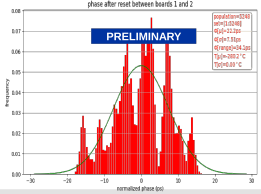
- Several subdetectors, especially new High Granularity Timing Detector (HGTD), require clock distribution with < 5 ps precision
- TCLink was designed to improve long term phase stability[1]
- Startup phase determinism on Virtex Ultrascale + and Versal Prime worse than KU with default TCLink design[2]
- Improving startup phase determinism on Versal Prime
  - Additional phase detector (DDMTD) to measure TXOUTCLK to ensure a deterministic startup phase
  - Initial system works, must be tailored for Versal



GTHe4 (ZCU102), Zynq US+ GTYe4 (VCU118) Virtex US+



GTYe5 (VMK180) Versal Prime



TX Phase detector

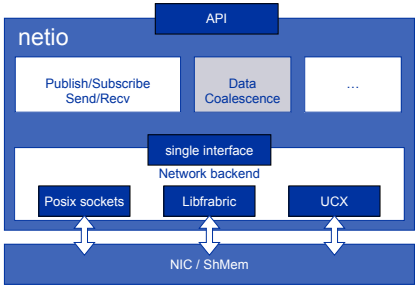
Aiming to achieve GTHe4-like startup determinism for Versal Prime

[1] TCLink reference note
[2] Results from F. Bonini (Cern, L1CT)

FELIX Software evolution for Phase II

Revised Netio-next design

- Netio-next provides a protocol-agnostic API for network I/O



- Phase I experience suggests 3 possible improvements
  - Versatility, modularity, maintainability
    - Unchanged user interface except for feature requests
    - Netio-next implements API to abstract underlying network
    - Port C to C++
  - DCS [1] and calibration dataflow
    - Dedicated DCS buffers and threads
    - TCP/IP instead of RDMA
  - Readout dataflow
    - Felix-star performance towards 1 MHz
    - Up to 10 DMA threads per PCIe card
    - Optionally align data by L0ID to move work load from data handler to FELIX

[1] DCS: Detector Control System

FELIX in the wild

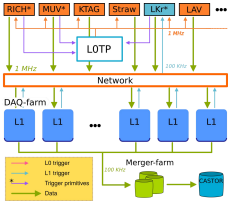
More information in backup slides

ProtoDUNE SP

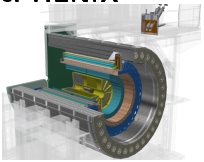
- 2 MHz readout via FELIX
- 15360 channels, 55 GB/s throughput
- Data taking started in 2018

NA62

- Kaon physics experiment
- At CERN SPS
- Readout with FELIX



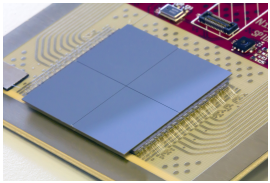
sPHENIX



- Located at RHIC, BNL
- 3 subdetectors readout with FELIX
- Streaming and triggered readout

SPIDR4

- Readout of Timepix4 pixel sensor
- Modified version of FELIX firmware and software



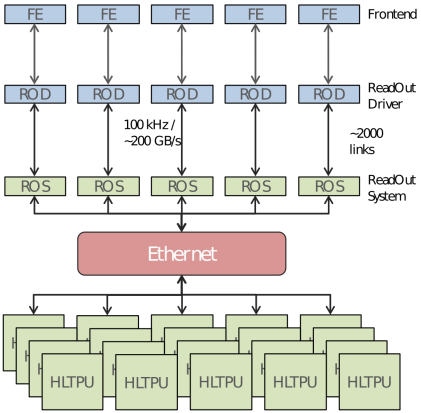
Summary

- The ATLAS Phase II upgrade is planned to be ready for LHC Run4 in 2029
- As part of the DAQ upgrade FELIX upgraded
  - New hardware design following increased bandwidth
  - Firmware redesigned to be more modular and added protocols to support all subdetectors and low level trigger systems
  - Software being redesigned for performance, modularity and different dataflows
- FELIX becoming more widely used in various forms outside ATLAS

Backup Slides

ATLAS DAQ Architecture in Run 2

Before FELIX, 2015-2018

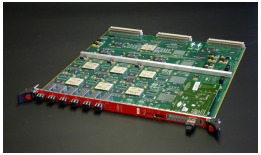


Readout Driver (ROD)

- Several types of links from FrontEnd to ROD
- VME boards, of about a dozen flavours developed and maintained by detectors
- Optical p2p link to ROBinNP cards

Readout System (ROS)

- Commodity computers hosting the ROBinNPs
- Transfers data to the High-Level Trigger farm over the network



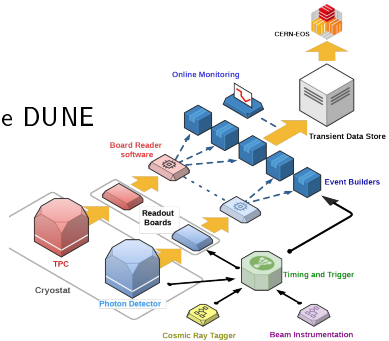
FELIX in the wild - protoDUNE-SP

ProtoDUNE single phase

- Demonstrator of design, construction, and operation of the DUNE TPC technologies
- Cryostat dimensions: 10x10x10 m, 750 ton of LAr
- Charged particle beam from SPS beam on target
- Test beam results published in arXiv:2007.06722

DAQ system[1,2]

- Continuous readout of TPC at 2 MHz via FELIX
- 15360 channels, 55 GB/s throughput
- Improved efficiency of FELIX firmware and software, merged into FELIX
- Data taking started in 2018

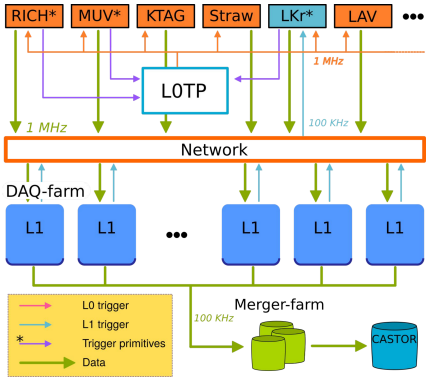


[1] https://doi.org/10.1051/epjconf/201921401013
[2] R. Sipos, IX Workshop on Streaming Readout

FELIX in the wild - NA62

- NA62 is a Kaon physics experiment located in the north area of the CERN SPS
- Kaon decay products detected using a wide range of detectors along a 270m long beamline.

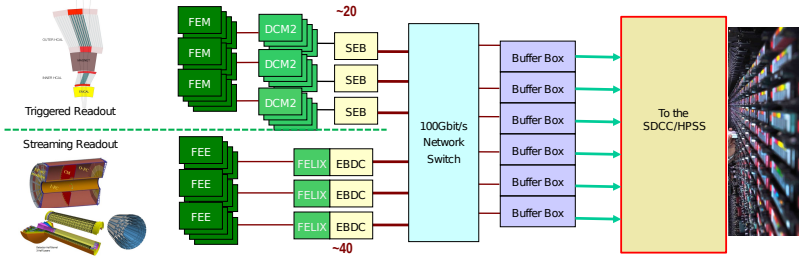
- L0 trigger in hardware, 1 MHz max event rate
- L1 software trigger reduces event rate to 100 kHz
- Moved to a FELIX based readout system [1]
- Felix is used to
  - buffer data: hits are indexed and trigger matching extracts relevant hits
  - distribute clock
  - manage synchronous communication for control and configuration



[1] M. Boretto, IX Workshop on Streaming Readout

FELIX in the wild - sPHENIX

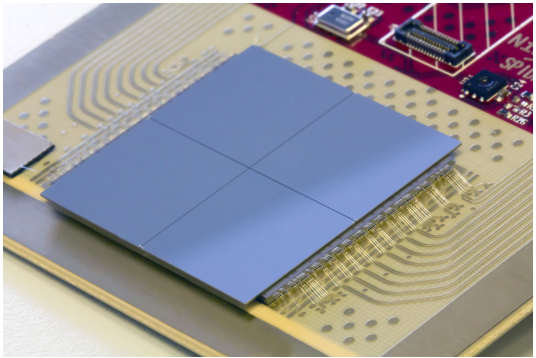
- Located at RHIC - BNL, sPHENIX studies QCD and QGP at different energy scales using p+p or Au+Au collisions
- Three sub-detectors read out with FELIX [1] in triggered and streaming mode
  - Pixel Vertex detector built with ALPIDE MAPS ($\sim$ 20 Gb/s)
  - Intermediate Silicon Strip Tracker ($\sim$ 7 Gb/s)
  - Compact Time Projection Chamber ($\sim$ 100 Gb/s)



[1] M. Purschke, IEEE-RT 2020

FELIX in the wild - SPIDR4 - Timepix4 readout

- Timepix4 is a large area pixel detector readout chip providing sub 200 ps timestamp binning [1]
- SPIDR4 is a readout solution for the Timepix4 pixel detector [2]
- Custom chip board, off-the-shelf PCIe card (BittWare)
- Modified version of FELIX firmware
  - Added custom decoder for 64b66b of TPX4 links
  - Unmodified FELIX CRToHost and Wupper PCIe DMA core
- DAQ software based on FELIX software, tailored for SPIDR4



[1] X. Llopart, Journal of Instrumentation 2022
[2] arXiv:2210.01442