

Coordinating Panel on Advanced Detectors – CPAD 2026

University of Washington □ Seattle □ WA

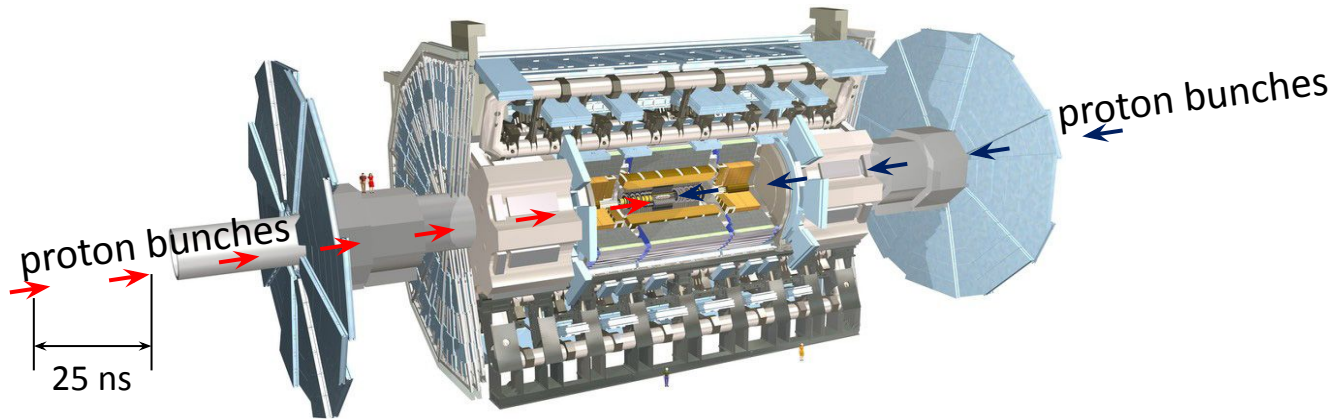
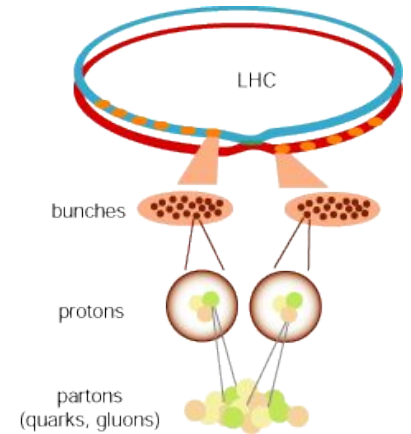
A Scalable Streaming Readout Architecture for Large-Scale FPGA-Based Trigger Systems

Tong Xu¹ on behalf of the ATLAS TDAQ Collaboration

¹Argonne National Laboratory, Lemont, IL, USA

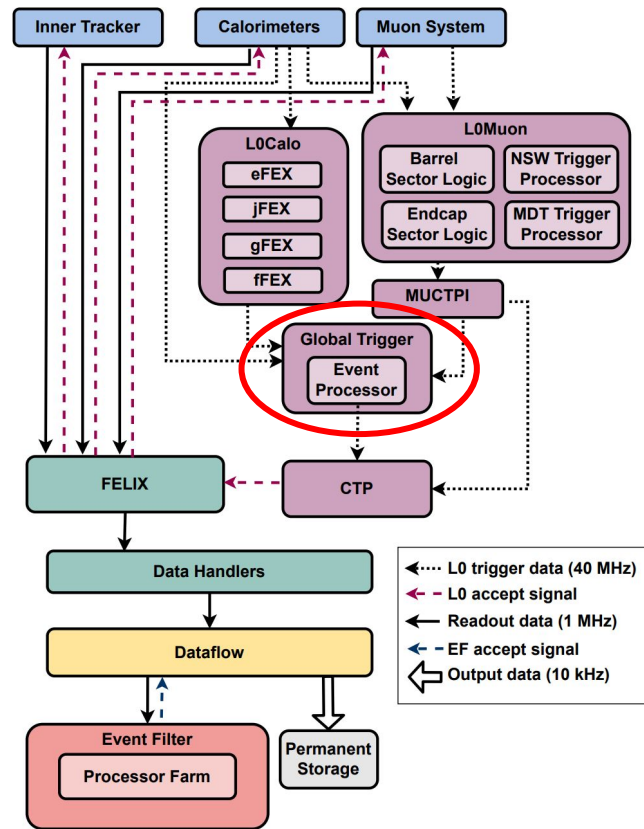
Collider Experiments and Trigger/DAQ Systems

- The Large Hadron Collider (LHC) accelerates protons to about 7 TeV and collides them every 25 ns. The ATLAS experiment at the LHC examines products from these collisions.
- Only a tiny fraction of collisions contain interesting physics. The ATLAS Trigger and Data Acquisition (TDAQ) system selects events of interest for further processing and storage.



Global Trigger

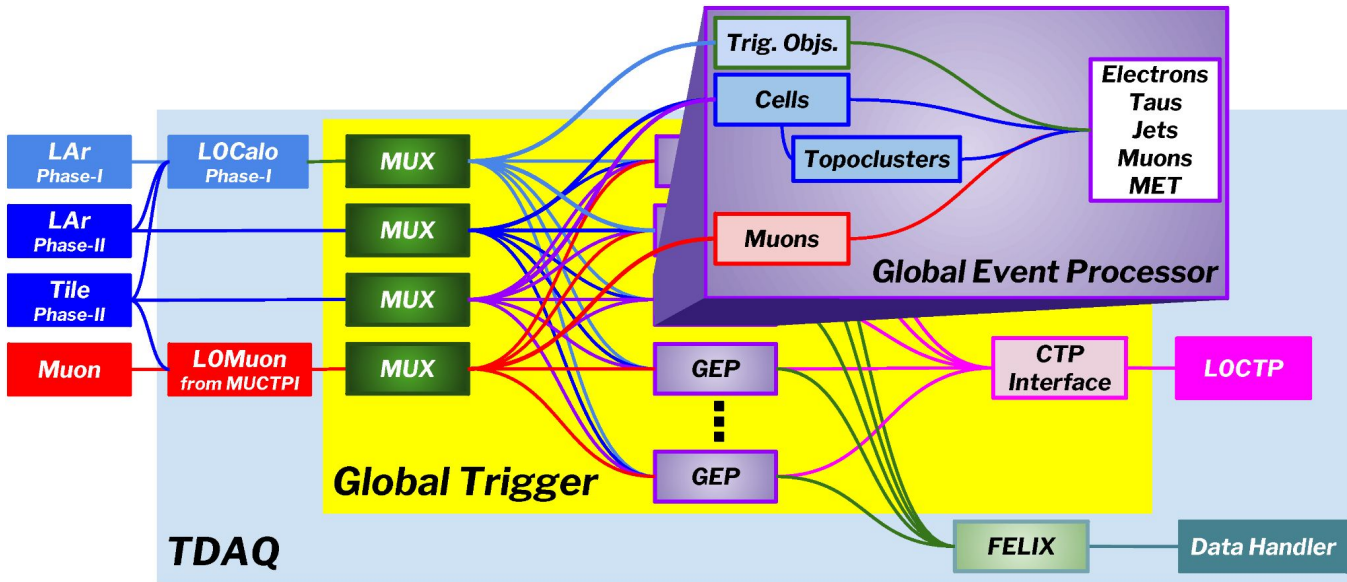
- Upgrade of the ATLAS Trigger System.
 - Up to **1 MHz output rate**, 10 μ s latency.
- Design focus:
 - Firmware-based implementation of trigger functionalities, enabling parallel execution of algorithms.
 - Unified hardware platform: **Global Common Module (GCM)**. (see the earlier talk from Shaochun)
- The Global Trigger analyzes the detector data to provide inputs to the Central Trigger Processor (CTP), which makes decisions on which events to read out.
- The Global Trigger will have multiple GCM nodes. Each node must be able to pass selected event data to off-line processing (FELIX). The challenge is to efficiently buffer event data for all collisions and transfer the selected data to the output links.



Upgraded ATLAS TDAQ architecture

Global Trigger system processing and trigger decision flow

- The Global Trigger system is organized in multiple layers, with firmware instances deployed on GCM nodes.
- 58 MUX nodes + 49 GEP nodes + 1 GCI node



MUX: Data Aggregation and Multiplexing

GEP: Global Event Processor

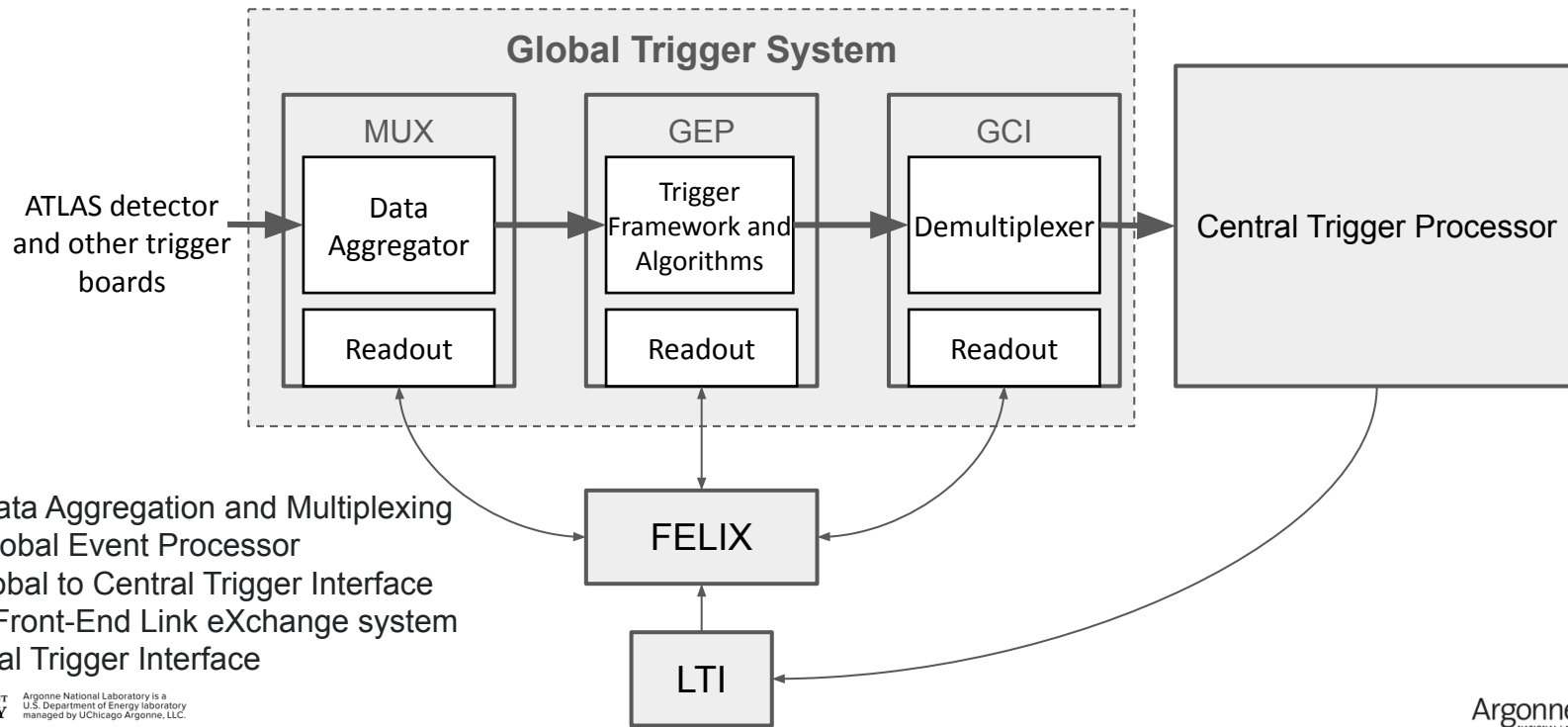
CTP: Central Trigger Processor

FELIX: Front-End Link eXchange system

Distributed Readout Across 108 GCM Nodes

The Global Trigger consists of 108 processing nodes deployed on the Global Common Module (GCM) FPGA boards.

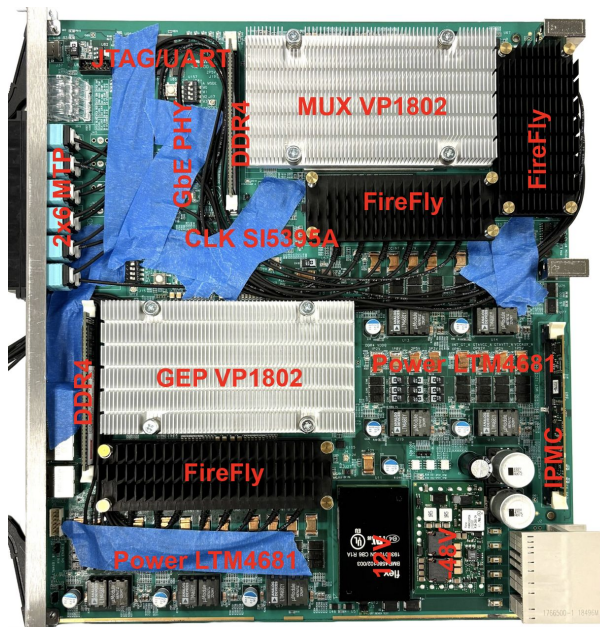
The readout firmware will be deployed on each GCM node. It will allow the transmission of the intermediate trigger data from the Trigger system to commodity computing.



FPGA boards: GCM and FELIX

The readout firmware will enable each GCM board to interface with the Front-End Link eXchange (FELIX) system via fiber-optic data links.

GCMv3



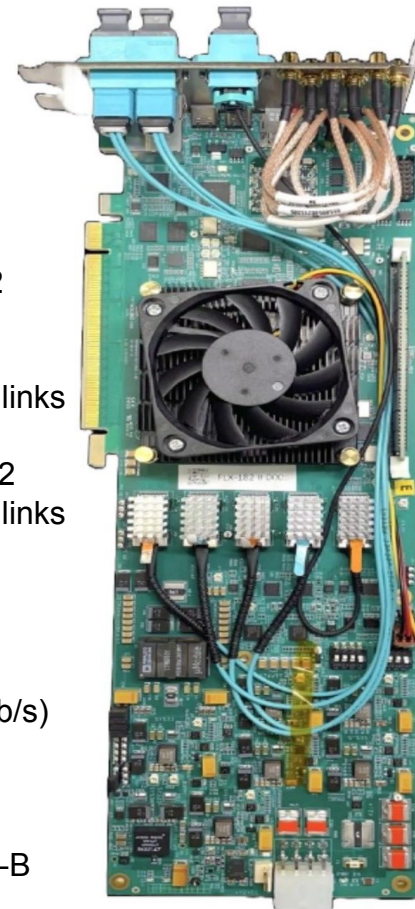
GCM Prototype V3

- FPGA: Two Versal Premium VP1802
- FireFly Modules
 - GEP/gCTPi: 1 T24 and 3 R24
 - 72 RX links and 24 TX links
 - Up to 25 Gb/s per link
 - MUX: 2 T24, 3 R24 and 1 Y12
 - 84 RX links and 60 TX links
 - Up to 25 Gb/s per link

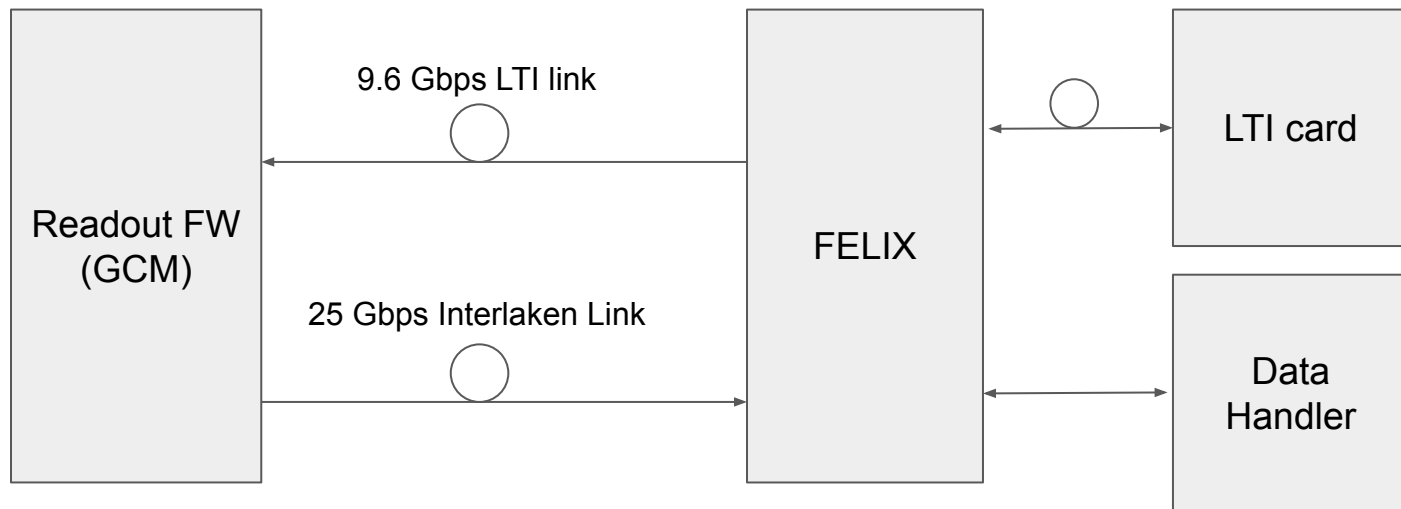
FELIX 182

- FPGA: Versal Prime VM1802
- 16-lane PCIe Gen4 interface (240 Gb/s)
- 4 FireFly transceivers (3 possible configurations)
 - 24 links up to 25 Gb/s
 - 24 links up to 10 Gb/s (CERN-B FireFly)

FLX-182



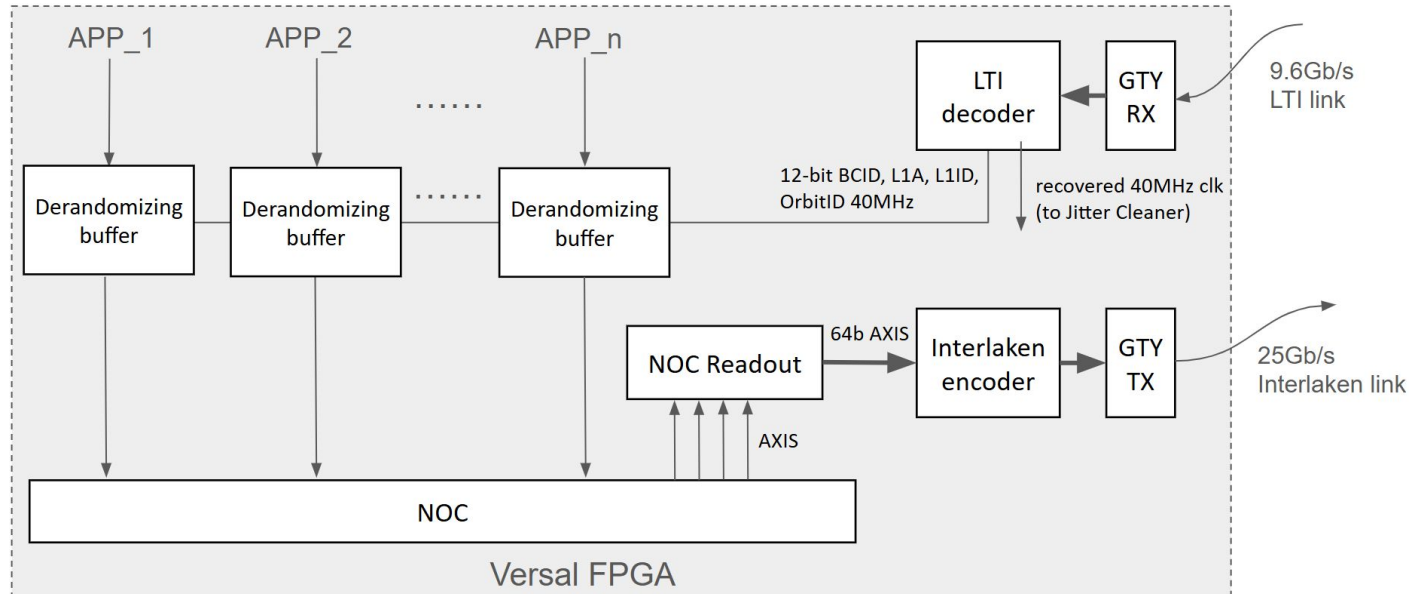
Connectivity to FELIX



- Each GCM node is connected to FELIX through two fiber-optic data links.
 - 9.6 Gb/s LTI link: synchronous to the 40.079 MHz LHC clock, L1A, BCID, OrbitID, L1ID, XOFF/XON signals, etc.
 - 25.78125 Gb/s Interlaken link: Selected event data, BUSY signal
- The data handler software will receive selected event data from the FELIX Card via the PCIe interface.
- The FELIX uses the Interlaken firmware with low-latency LTI fan-out

Streaming Readout Architecture for a GEP node

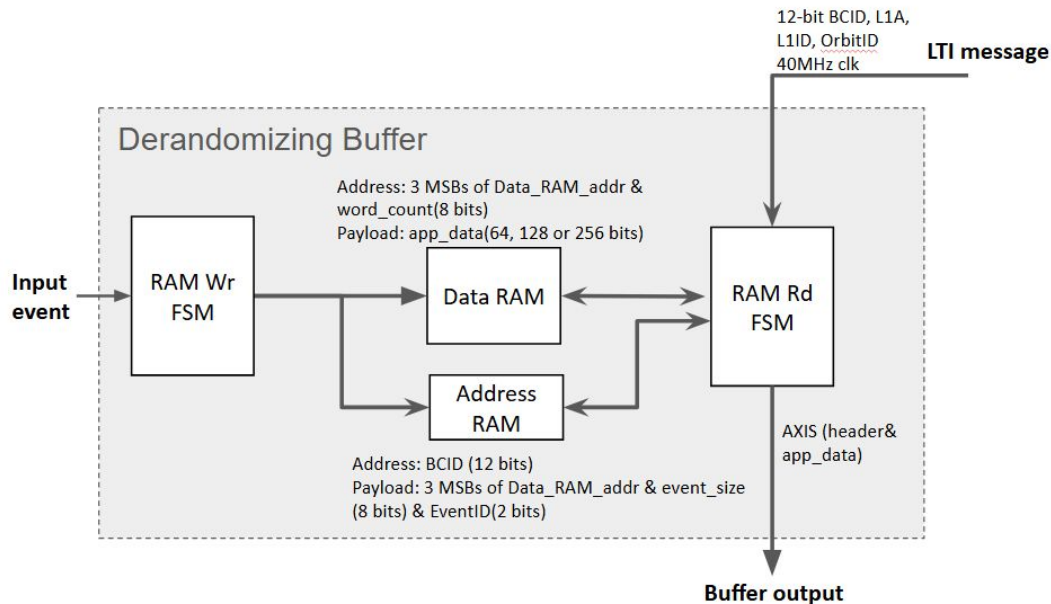
- The readout firmware buffers event data generated by trigger algorithms (APP). It then uses Bunch Crossing identifier (BCID) from the LTI link as the index for readout. The readout of selected data is initiated upon the arrival of Level-1 trigger accept (L1A) signal.
- The BCID identifies which 25 ns bunch slot (0–3563) within an LHC orbit a collision occurred in.



FELIX

Derandomizing buffer: decoupled event-data buffering and Readout

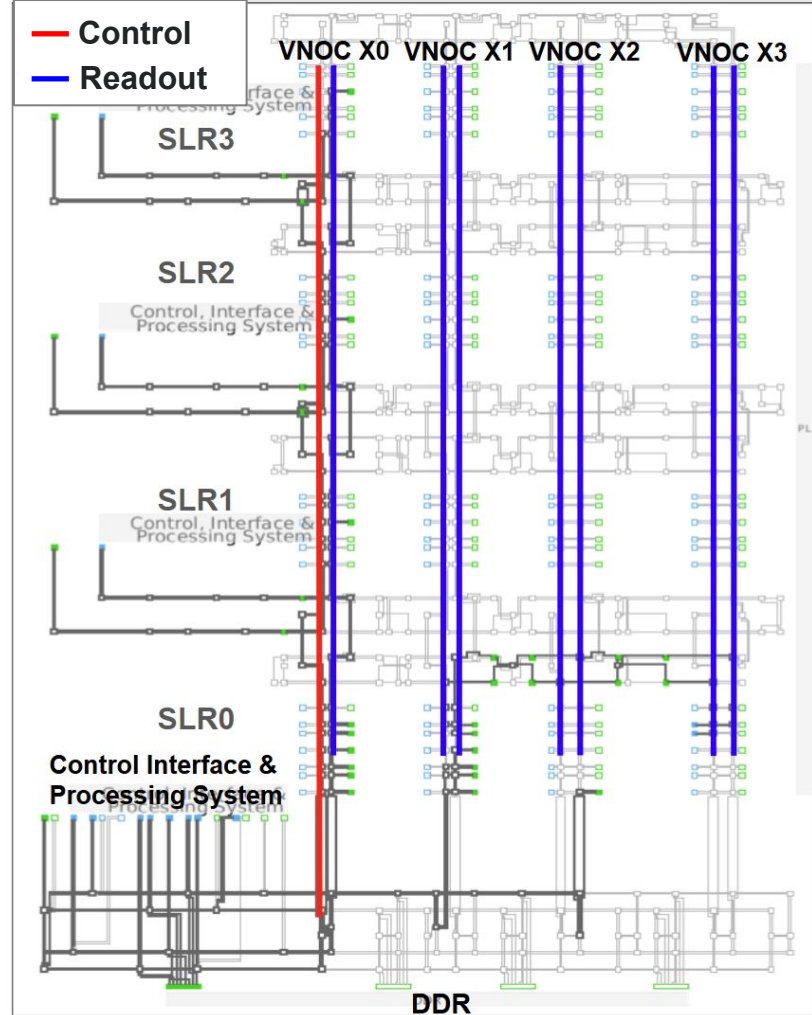
- A derandomizing buffer consists of dual-port RAMs and state machines that control the write and read operations.
- All input data is written continuously. Reading occurs upon arrival of an L1A.
- The buffer decouples continuous event data buffering from asynchronous trigger-driven readout, which enables burst-tolerant streaming operation.



- Generates a readout header for every L1A & BCID
- Readout data:
 - **BCID Match**: header + event data
 - **No match**: header only
- Configurable:
 - Input/output width
 - Event size
 - Buffer depth

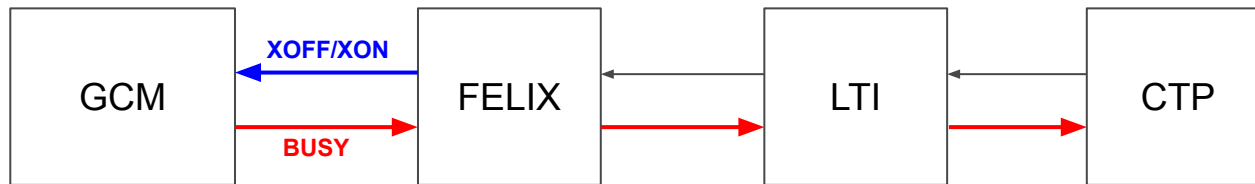
Scalable Inter-SLR data movement with NoC

- In GEP, the readout firmware uses Network-on-Chip (NoC) to transfer data from all Super Logic Regions (SLRs) to SLR0.
- Seven (out of eight) NoC lanes are used by the readout FW, and offer 860.48 Gb/s of maximum raw bandwidth ($7 \times 128b \times 960\text{MHz}$) towards SLR0. That exceeds the bandwidth (25 Gb/s) of the Interlaken link.
- NoC allows placement of the derandomizing buffers close to the trigger algorithm firmware. It supports scalable inter-SLR data transport as the number of buffers increases.
- Latest GCMv4 GEP readout implementation:
 - SLR3 (APPs + Derandomizing buffers) $\rightarrow$ NoC $\rightarrow$ SLR0 (Interlaken TX)



Flow control between Global, FELIX, and CTP

FLOW control is used to avoid uncontrolled data loss during temporary congestion.



FELIX → Global: XOFF/XON

- If buffers in FELIX are full, it sends **XOFF** to Global through the LTI link.
- Global pauses transmission through the Interlaken TX but continues processing the L1A.
- When the congestion is cleared, FELIX sends **XON**, and Global resumes transmission.
- **XOFF** is intended for short, exceptional periods to avoid data loss. It is unlikely to happen.

Global → FELIX → LTI → CTP: BUSY

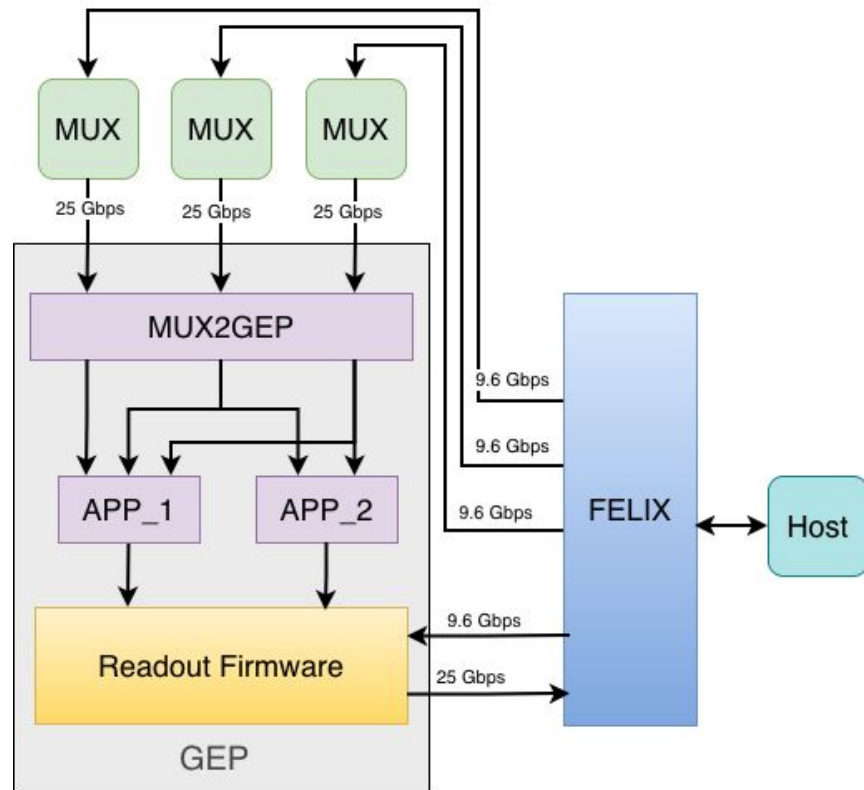
- If the Interlaken link FIFO in GCM reaches **prog_full**, Global asserts **BUSY** to FELIX through the Interlaken link.
- FELIX propagates **BUSY** to the CTP, which temporarily suppresses triggers.
- As the FIFO utilization drops below the threshold, Global de-asserts **BUSY**.
- FELIX propagates that to the CTP, and normal trigger operation resumes.

Testing & Verification

- Link tests between GCM and FLX182
 - LTI Link (9.6 Gb/s): Observed stable error-free link alignment and decoded LTI messages. CRC and 8b/10b encoding were used for error detection.
 - Interlaken Link (25 Gb/s): Also observed stable link alignment in FELIX and test data transfer from GCM to FLX182. The data was captured and inspected with software in the FELIX server.
- Error-free end-to-end verification of GEP readout
- 48-channel MUX readout test

End-to-End GEP Readout Verification

- Setup: 3 MUX → 1 GEP → FELIX
 - Each MUX uses a data emulator with LTI correlated BCID for data generation
 - In GEP, APP data was buffered and selected using L1A BCID
 - Both 25G and 9.6G MUX patterns tested
 - Validated > 3000 events. No mismatches observed in validated data
- Also verified:
 - Single-SLR GEP → FELIX
 - Multi-SLR GEP → FELIX



Example of GEP event data verification

```
[felix@dune rdo_data_Aug8]$ fdaq -t 10 -D -d 0 test
Opened FLX-device 0, FLX182-INTERLAKEN-2x12CH-250522-2337-GIT:rm=5.3/765, trailer=32bit ChunkHeaders, buffer=1024MB, D
Before startRec: Receiver@FLX0 DEBUG: head=1801216, tail=0, pwrite=01521b7c00
  start=0152000000, end=0192000000, psoftw=0152000000(0152000000), pfirwm=01521d8000, even-PC/DMA 0 0
  wraps=0, recvd=1801216, last-recvd=1801216, last-av=1801216, fullcnt=0

After startRec: Receiver@FLX0 DEBUG: head=11075584, tail=1801216, pwrite=0152a90000
  start=0152000000, end=0192000000, psoftw=01521b7c00(01521b7c00), pfirwm=0152c5d000, even-PC/DMA 0 0
  wraps=0, recvd=11075584, last-recvd=1156096, last-av=1801216, fullcnt=0

**START** using DMA #0 polling
-----
Secs | Recvd[MB/s] | File[MB/s] | Total[(M)B] | Rec[(M)B] | Buf[%] | Wraps
-----|-----|-----|-----|-----|-----|-----
1 | 1102.2 | 1099.2 | 1102.2 | 1099.2 | 0 | 1
2 | 1092.4 | 1070.4 | 2194.6 | 2169.7 | 2 | 2
3 | 1091.0 | 1110.7 | 3285.7 | 3280.4 | 0 | 3
4 | 1091.2 | 1072.7 | 4376.9 | 4353.1 | 2 | 4
5 | 1092.2 | 1104.9 | 5469.1 | 5458.0 | 1 | 5
6 | 1091.1 | 1076.1 | 6560.2 | 6534.1 | 2 | 6
7 | 1091.0 | 1114.1 | 7651.2 | 7648.2 | 0 | 7
8 | 1091.9 | 1088.4 | 8743.1 | 8736.6 | 1 | 8
9 | 1092.1 | 1091.0 | 9835.2 | 9827.6 | 1 | 9
10 | 1091.4 | 1072.9 | 10926.6 | 10900.5 | 2 | 10

**STOP**
After stopRec: Receiver@FLX0 DEBUG: head=109533568, tail=108383616, pwrite=015de4a400
  start=0152000000, end=0192000000, psoftw=015dd31800(015dd31800), pfirwm=015e054000, even-PC/DMA 0 0
  wraps=10, recvd=10936951808, last-recvd=1149952, last-av=3449856, fullcnt=0

--> Totals: Recvd 10936 MB, File 10917 MB (last 2312 MB), 3 files
--> File : ./test-250808-143451-3.dat
Exiting..
```

No matches
(FADE)

Matched event
(ABCD)

```
1 fa de 1d 10 b3 0f 54 fc 16 31 55 d4 ac 00 0d b0
2 fa de 2d 10 b3 0f 54 fc 16 31 55 d4 ac 00 0d b0
3 fa de 1d 38 b3 0f 54 fc 16 31 55 d4 b0 00 0d e0
4 fa de 2d 38 b3 0f 54 fc 16 31 55 d4 b0 00 0d e0
5 fa de 1d 60 b3 0f 54 fc 16 31 55 d4 b4 00 0d e0
6 fa de 2d 60 b3 0f 54 fc 16 31 55 d4 b4 00 0d 24
7 fa de 1d 88 b3 0f 54 fc 16 31 55 d4 b8 00 0d 24
8 fa de 2d 88 b3 0f 54 fc 16 31 55 d4 b8 00 0d 24
9 fa de 1d b0 b3 0f 54 fc 16 31 55 d4 bc 00 0d 54
10 fa de 2d b0 b3 0f 54 fc 16 31 55 d4 bc 00 0d 54
11 ab cd 1d d8 b3 0f 54 fc 16 31 55 d4 c0 00 0d 84
12 00 59 47 09 47 09 47 05 00 04 47 04 47 04 47 00
13 01 03 47 13 47 13 47 0f 00 ae 47 0e 47 0e 47 0a
14 01 ad 47 1d 47 1d 47 19 01 58 47 18 47 18 47 14
15 02 57 47 27 47 27 47 23 02 02 47 22 47 22 47 1e
16 05 59 47 09 47 09 47 05 05 04 47 04 47 04 47 00
17 06 03 47 13 47 13 47 0f 05 ae 47 0e 47 0e 47 0a
18 06 ad 47 1d 47 1d 47 19 06 58 47 18 47 18 47 14
19 07 57 47 27 47 27 47 23 07 02 47 22 47 22 47 1e
20 0a 59 47 09 47 09 47 05 0a 04 47 04 47 04 47 00
21 0b 03 47 13 47 13 47 0f 0a ae 47 0e 47 0e 47 0a
22 0b ad 47 1d 47 1d 47 19 0b 58 47 18 47 18 47 14
23 0c 57 47 27 47 27 47 23 0c 02 47 22 47 22 47 1e
24 0f 59 47 09 47 09 47 05 0f 04 47 04 47 04 47 00
25 10 03 47 13 47 13 47 0f 0f ae 47 0e 47 0e 47 0a
26 10 ad 47 1d 47 1d 47 19 10 58 47 18 47 18 47 14
27 11 57 47 27 47 27 47 23 11 02 47 22 47 22 47 1e
28 14 59 47 09 47 09 47 05 14 04 47 04 47 04 47 00
29 15 03 47 13 47 13 47 0f 14 ae 47 0e 47 0e 47 0a
30 15 ad 47 1d 47 1d 47 19 15 58 47 18 47 18 47 14
31 16 57 47 27 47 27 47 23 16 02 47 22 47 22 47 1e
32 19 59 47 09 47 09 47 05 19 04 47 04 47 04 47 00
33 1a 03 47 13 47 13 47 0f 19 ae 47 0e 47 0e 47 0a
34 1a ad 47 1d 47 1d 47 19 1a 58 47 18 47 18 47 14
35 1b 57 47 27 47 27 47 23 1b 02 47 22 47 22 47 1e
36 1e 59 47 09 47 09 47 05 1e 04 47 04 47 04 47 00
37 1f 03 47 13 47 13 47 0f 1e ae 47 0e 47 0e 47 0a
```

- GEP event data were captured by FELIX and verified offline. No issues were observed.

48-Channel MUX Readout Verification

- The MUX (Multiplexer) aggregates data from multiple detector sources and forwards them to GEP nodes for further processing.
- Setup: Data Source Sink (DSS) → MUX → FELIX
 - DSS generates emulated data using the 25G MUX data pattern, with a 200-BC offset relative to the to LTI BCID
 - Full 48-channel MUX readout; 48 derandomizing buffers were instantiated
 - Data captured by FELIX were validated word-by-word
 - Validated > 4,000 events
- Also tested MUX readout:
 - Setup: MUX with data emulator → FELIX
 - Verified for 1-ch, 6-ch and 48-ch

Example of MUX event data verification

```
[felix@dune mux_rdo_data]$ fdaq -t 5 -D -d 0 test
Opened FLX-device 0, FLX182-INTERLAKEN-2x12CH-260328-1252-GIT:rm-5.4/653, trailer=32bit ChunkHeaders, buffer=1024MB, DMA=0
Before startRec: Receiver@FLX0 DEBUG: head=0, tail=0, pwrite=01a2000000
  start=01a2000000, end=01e2000000, psoftw=01a2000000(01a2000000), pfirwm=01a2000000, even-PC/DMA 0 0
  wraps=0, recvd=0, last-recvd=0, last-av=0, fullcnt=0

After startRec: Receiver@FLX0 DEBUG: head=115712, tail=76800, pwrite=01a201c400
  start=01a2000000, end=01e2000000, psoftw=01a2012c00(01a2012c00), pfirwm=01a201e400, even-PC/DMA 0 0
  wraps=0, recvd=115712, last-recvd=5120, last-av=76800, fullcnt=0

**START** using DMA #0 polling
  Secs | Recvd[MB/s] | File[MB/s] | Total[(M)B] | Rec[(M)B] | Buf[%] | Wraps
-----|-----|-----|-----|-----|-----|-----
  1    | 4.3         | 4.2         | 4.3          | 4.2        | 0       | 0
  2    | 4.2         | 4.2         | 8.5          | 8.5        | 0       | 0
  3    | 4.2         | 4.1         | 12.7         | 12.6       | 0       | 0
  4    | 4.2         | 4.2         | 16.8         | 16.8       | 0       | 0
  5    | 4.2         | 4.2         | 21.0         | 21.0       | 0       | 0

**STOP**
After stopRec: Receiver@FLX0 DEBUG: head=21080064, tail=21070848, pwrite=01a341a800
  start=01a2000000, end=01e2000000, psoftw=01a3418400(01a3418400), pfirwm=01a341b800, even-PC/DMA 0 0
  wraps=0, recvd=21080064, last-recvd=4096, last-av=84992, fullcnt=0

-> Totals: Recvd 21 MB, File 20 MB (last 20 MB), 1 files
-> File : ./test-260514-110917-1.dat
Exiting..
```

```
1 ab cd 10 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
2 10 11 19 91 19 91 19 91 10 00 19 90 19 90 19 90
3 10 33 19 93 19 93 19 93 10 22 19 92 19 92 19 92
4 10 55 19 95 19 95 19 95 10 44 19 94 19 94 19 94
5 10 77 19 97 19 97 19 97 10 66 19 96 19 96 19 96
6 ab cd 25 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
7 25 11 19 91 19 91 19 91 25 00 19 90 19 90 19 90
8 25 33 19 93 19 93 19 93 25 22 19 92 19 92 19 92
9 25 55 19 95 19 95 19 95 25 44 19 94 19 94 19 94
10 25 77 19 97 19 97 19 97 25 66 19 96 19 96 19 96
11 ab cd 11 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
12 11 11 19 91 19 91 19 91 11 00 19 90 19 90 19 90
13 11 33 19 93 19 93 19 93 11 22 19 92 19 92 19 92
14 11 55 19 95 19 95 19 95 11 44 19 94 19 94 19 94
15 11 77 19 97 19 97 19 97 11 66 19 96 19 96 19 96
16 ab cd 26 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
17 26 11 19 91 19 91 19 91 26 00 19 90 19 90 19 90
18 26 33 19 93 19 93 19 93 26 22 19 92 19 92 19 92
19 26 55 19 95 19 95 19 95 26 44 19 94 19 94 19 94
20 26 77 19 97 19 97 19 97 26 66 19 96 19 96 19 96
21 ab cd 12 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
22 12 11 19 91 19 91 19 91 12 00 19 90 19 90 19 90
23 12 33 19 93 19 93 19 93 12 22 19 92 19 92 19 92
24 12 55 19 95 19 95 19 95 12 44 19 94 19 94 19 94
25 12 77 19 97 19 97 19 97 12 66 19 96 19 96 19 96
26 ab cd 27 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
27 27 11 19 91 19 91 19 91 27 00 19 90 19 90 19 90
28 27 33 19 93 19 93 19 93 27 22 19 92 19 92 19 92
29 27 55 19 95 19 95 19 95 27 44 19 94 19 94 19 94
30 27 77 19 97 19 97 19 97 27 66 19 96 19 96 19 96
31 ab cd 13 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
32 13 11 19 91 19 91 19 91 13 00 19 90 19 90 19 90
33 13 33 19 93 19 93 19 93 13 22 19 92 19 92 19 92
34 13 55 19 95 19 95 19 95 13 44 19 94 19 94 19 94
35 13 77 19 97 19 97 19 97 13 66 19 96 19 96 19 96
36 ab cd 28 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
37 28 11 19 91 19 91 19 91 28 00 19 90 19 90 19 90
38 28 33 19 93 19 93 19 93 28 22 19 92 19 92 19 92
39 28 55 19 95 19 95 19 95 28 44 19 94 19 94 19 94
40 28 77 19 97 19 97 19 97 28 66 19 96 19 96 19 96
41 ab cd 14 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
42 14 11 19 91 19 91 19 91 14 00 19 90 19 90 19 90
43 14 33 19 93 19 93 19 93 14 22 19 92 19 92 19 92
44 14 55 19 95 19 95 19 95 14 44 19 94 19 94 19 94
45 14 77 19 97 19 97 19 97 14 66 19 96 19 96 19 96
46 ab cd 29 19 93 79 22 8a 84 d5 d7 b0 00 40 00 00
47 29 11 19 91 19 91 19 91 29 00 19 90 19 90 19 90
48 29 33 19 93 19 93 19 93 29 22 19 92 19 92 19 92
49 29 55 19 95 19 95 19 95 29 44 19 94 19 94 19 94
50 29 77 19 97 19 97 19 97 29 66 19 96 19 96 19 96
```

- MUX readout is performed for less than 0.1% of L1As, so the data rate is much smaller than for GEP.
- Each MUX node sees all BCIDs, so a match was identified for every LTI BCID specified by L1A. No issues were observed.

Summary

- A scalable streaming readout architecture has been developed to collect intermediate processed data from the Global Trigger System and transmit the selected data to commodity computing through FELIX.
- The architecture is implemented in firmware that continuously buffers all the intermediate event data and performs trigger-driven readout. Upon receiving the trigger decision, the firmware retrieves selected events from the buffers and streams them to output data link. It uses the NoC for inter-SLR data transfer in the FPGA.
- The architecture is designed to support over 100 FPGA processing nodes of the Global Trigger. The firmware has been implemented and verified for GEP and MUX nodes.
- End-to-end data flow across GCM, FELIX, and host software has been verified.

Thank you!



U.S. DEPARTMENT
of ENERGY

Argonne National Laboratory is a
U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC.

