FELIX: readout upgrade for the ATLAS Trigger DAQ system in HL-LHC*

Mengqing Wu[a,*], on behalf of the ATLAS TDAQ Collaboration

[a]*Institute for Mathematics, Astrophysics and Particle Physics, Radboud University/Nikhef, 6500 GL Nijmegen, The Netherlands*

Abstract

The High Luminosity LHC (HL-LHC) will start to operate in 2029 (Run 4) with an instantaneous luminosity five times higher than the LHC nominal value, meaning the ATLAS experiment will be operated in an increasingly harsh collision environment. This has motivated a series of upgrades, to be installed during the long experimental shutdown period from 2026 to 2029. One key change among these is the upgrade of the Front-End Link eXchange (FELIX) system, which was developed for the third major LHC data taking period (Run 3) from 2022 to 2025 to improve the capacity and the flexibility of the detector readout system for selected ATLAS systems. After the HL-LHC upgrade, all ATLAS systems will be read out via the FELIX system, with sub-detector specific processing taking place in a common software processing framework. One major upgrade challenge for the system in Run 4 is to support the trigger rate increase from 100 kHz in Run 3 to 1 MHz in Run 4, along with an overall increase in event size. This presentation will cover the FELIX design for Run 4, as well as the results of a series of performance tests carried out with Run 3 hardware at Run 4 rates, which will inform the next stages of development.

Keywords: CERN LHC, ATLAS, DAQ, FELIX

1. Introduction

The Large Hadron Collider (LHC) collides proton bunches at a 40 MHz rate and ATLAS detects collision products with the trigger systems selecting physics events of interest. The new FELIX (Front-End Link eXchange) system [1] was installed for selected ATLAS systems at the current Run 3. It functions as a router between custom serial links from front-end electronics to data collection and processing components via a commercial off-the-shelf (COTS) switched network, and also forwards TTC (Timing, Trigger and Control) signals to front-end electronics. The trigger input rate to the FELIX system is capped at 100 kHz at Run 3. The Run 3 FELIX systems consist of COTS servers, each hosting up to two FPGA-based PCIe I/O cards (FELIX cards). The FELIX card interfaces through a switched internet with the SoftWare ReadOut Driver (SW ROD) that runs on commodity servers for event building and aggregation, and detector-specific data processing.

At the High Luminosity LHC (HL-LHC), the instantaneous luminosity will be increased to five times the LHC nominal value. This imposes data acquisition challenges for the ATLAS experiment, with a three times higher average number of interactions per bunch crossing, a 10 times higher hardware trigger rate, and a five times higher data recording rate.

An upgrade of the FELIX system for HL-LHC is currently under development and it will be installed for all ATLAS systems, as shown in Figure 1. Similarly to Run 3, its baseline design is based on commodity servers hosting up to two FPGA-based PCIe I/O FELIX Cards (Gen 4 or later). Each FELIX host transfers data through a high-bandwidth switched network to server farms that run a newly developed multi-threaded software platform called 'Data Handler', as successor to the Run 3 SW ROD.

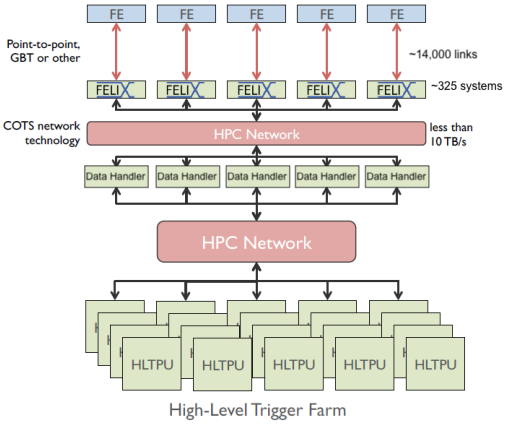

Figure 1: HL-LHC FELIX system diagram. [2]

2. Performance results in HL-LHC environment

One major upgrade challenge for the FELIX system in Run 4 is to support the trigger rate increase from 100 kHz (Run 3) to 1 MHz (Run 4), along with an overall increase in event size. This motivated a series of performance tests carried out with the Run 3 FELIX hardware at the Run 4 trigger rates. The test setup is shown in Figure 2, where a Run 3 FELIX server equivalent

*speaker
Email address: mengqing.wu@cern.ch (Mengqing Wu)

has been used to host a Run 3 FELIX card (FLX-712) with an external TTC system. An external data generator is used to generate data in configuarable sizes and send them to FELIX via different link protocols. Data were processed either in the FELIX server or transferred across a network to a SW ROD.

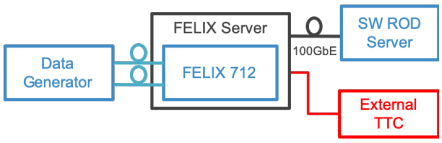

Figure 2: Performance test setup with Run 3 hardware.

FLX-712 card. The FLX-712 card [3] is equipped with a Xilinx Kintex UltraScale FPGA, 4/8 MiniPOD optical transceivers to support 24/48 bidirectional optical links, and 16-lane PCIe Gen3 divided into two 8-lane endpoints with a switch. It is equipped with Flash and Micro-controller to support firmware update and capable to accomodate multiple TTC architectures with BUSY signal. The firmware supports two data taking modes: GigaBit Transceiver (GBT) mode and FULL mode. The GBT mode is designed to interface with the radiation-hard GBTx [4] ASICs developed at CERN. One FLX-712 card can support up to 24 4.8 Gb/s GBT links each divided into maximum 40 E-links with configurable bandwidth. The FULL mode supports communication with other FPGA-based systems over an in-house protocol implemented as a single wide data stream with no handshaking or logical substructure (i.e. no E-links). Each FLX-712 card supports up to 24 FULL mode links with each link 8b10b encoded with a line speed of 9.6 Gb/s.

FELIX readout software for the Run 3 upgrade. The FELIX readout software uses RDMA (Remote Direct Memory Access) technology for low overhead transfers based on the custom 'netio-next' library. Low level software is developed for basic configuration and monitoring, while high level software (such as the SW ROD) is used for high rate data taking and channel monitoring.

Test results. Both data taking modes were tested with various data packet size, number of links to find the maximum trigger rate with stable operation. The data rate was initially measured with data processing locally in the FELIX server with increasing trigger rate over around half an hour, see example measurements in Figure 3. The results show stable operation with 24 links up to 1.1 MHz for GBT mode with each link configured with eight 8 bit E-links receiving 26 byte data packet; 1.2 MHz for FULL Mode with each link receiving 364 byte data; and thus satisfying the required 1 MHz in the HL-LHC environment.

Tests were then carried out also with data transferred via the network to the SW ROD in both data taking modes. The measured maximum data rate in this case is 0.65 MHz for GBT mode and close to 1 MHz for FULL Mode (but with 12 links and a smaller packet size of 192 byte). The limitation in this case proved to be the network bandwidth, meaning upgraded hardware will be needed to move to higher rates.

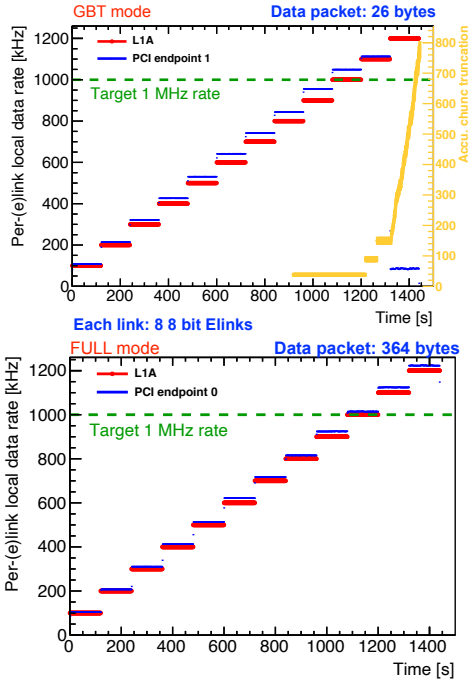

Figure 3: Data rate measured with data processed in the FELIX server in GBT mode (top) and FULL mode (bottom). Red lines are trigger rate input to the FELIX card while blue lines show the data rate measured in the FELIX server. Orange lines for 'chunk truncation' indicate that FELIX has backpressure from e.g. network saturation and had to truncate data packets to maintain operation.

3. Summary

The FELIX system will be upgraded to tackle the readout challenges in the HL-LHC environment. A series of performance tests have been carried out with the Run 3 FELIX hardware at the 1 MHz Run 4 trigger rate. The results show stable operation up to 1.2 MHz with 364 byte data packets in FULL mode and 1.1 MHz with 26 byte data packets in GBT mode with data processing locally in the FELIX server. Backpressure from network bandwidth was found to prevent the current test setup from reaching the required 1 MHz trigger rate with data transferred over the network to the SW ROD. This gives confidence that the design will scale to the required 1 MHz trigger rate, given the expected technology evolution over the coming years. A series of follow-up tests is planned to maximize throughput with data transfers to the SW ROD across the network.

References

[1] J. Anderson, et al., FELIX: a PCIe based high-throughput approach for interfacing front-end and trigger electronics in the ATLAS upgrade framework, Journal of Instrumentation 11 (12) (2016) C12023–C12023. doi:10.1088/1748-0221/11/12/c12023.

[2] Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System, Tech. rep., CERN, Geneva (Sep 2017). doi:10.17181/CERN.2LBB.4IAL.

[3] K. Chen, et al., A generic high bandwidth data acquisition card for physics experiments, IEEE Transactions on Instrumentation and Measurement 69 (7) (2020) 4569–4577. doi:10.1109/TIM.2019.2947972.

[4] P. Moreira, et al., The GBT Project (2009). doi:10.5170/CERN-2009-006.342.