The FELIX detector interface for the ATLAS TDAQ upgrades and its deployment in the ITk demonstrator setup

C.A. Solans Sanchez, on behalf of the ATLAS Collaboration

CERN, Geneva, Switzerland

E-mail: `carlos.solans@cern.ch`

Abstract. The ATLAS experiment at the LHC at CERN will move to use the Front-End Link eXchange (FELIX) system in a staged approach for LHC Run 3 (2021) and LHC Run 4 (2026). FELIX will act as the interface between the data acquisition; detector control and TTC (Timing, Trigger and Control) systems; and new or updated trigger and detector front-end electronics. FELIX functions as a router between custom serial links from front end ASICs and FPGAs to data collection and processing components via a commodity switched network. Links may aggregate many slower links or be a single high bandwidth link. FELIX also forwards the LHC bunch-crossing clock, fixed latency trigger accepts and resets received from the TTC system to front-end electronics. The FELIX system uses commodity server technology in combination with FPGA-based PCIe I/O cards. The FELIX servers run a software routing platform serving data to network clients. Commodity servers connected to FELIX systems via the same network run innovative multi-threaded software for event fragment building, processing, buffering and forwarding. This proceeding will describe the design and status of the FELIX based readout for the Run 3 upgrade, during which a subset of the detector will be migrated. It will also show how the same concept has been successfully introduced into the demonstrator test bench of the ATLAS Pixel Inner Tracker, acting as a proof of concept towards the longer term Run 4 upgrade in which all detectors will adopt a FELIX based readout.

1. Introduction

The ATLAS experiment [1] will undergo a series of upgrades known as Phase-I (2019) and Phase-II (2024) to cope with the increase in luminosity expected for the LHC Run 3 (2021) and LHC Run 4 (2026). These upgrades will result in an increase of two orders of magnitude in the number of electronic channels that the Trigger and Data AQcuisition (TDAQ) system will have to read out. During these upgrades the Front End LInk eXchange (FELIX) system will be introduced for all ATLAS readout [2]. One of the systems that will make significant use of FELIX will be the ITk [3], that is an all silicon tracking detector made out of 5 pixel layers and 4 strip layers. The work described in this proceeding is part of a programme to demonstrate the calibration of ITk front-end prototypes, such that the needs and requirements for such procedures can be clearly established.

2. Upgrades of the ATLAS DAQ System

The ATLAS DAQ system during Run 2 was based on custom point-to-point links from the different sub-detector front-end electronics to the Read-out Driver (ROD). The ROD was an

ATLAS wide element with sub-detector specific implementation, that provided the interface with the Read-out System (ROS) through point-to-point high speed S-LINKS [4], and included among other functionalities the trigger handling, the busy generation to halt new triggers, and local monitoring capabilities. Similarly, FELIX is the read-out system component which implements the interface to all detector specific electronics via custom point-to-point serial links, functioning as a router between them and a commodity multi-gigabit network, allowing routing of event data to the DAQ system, distribution of the TTC (Timing, trigger and control), and the calibration and monitoring of the detector. It is detector agnostic and encapsulates common functionality without the requirement to decode or process the data.

In terms of data throughput, by the end of Run 2 the throughput of the ROS was 50 GB/s, the average recording rate 1.5 GB/s, and the event size 2 MB, which is not expected to change after the Phase-I upgrade. However, after the Phase-II upgrade the network throughput is expected to increase up to 290 GB/s, with an increase of the recording rate of 50 GB/s, and an event size of 5 MB. A new data handler or Software ROD (SWROD) element will be added to the DAQ chain that will perform detector specific data processing without data buffering including configuration, calibration, control, and monitoring. A diagram of the TDAQ architecture for the Phase-II upgrade is shown in figure 1.
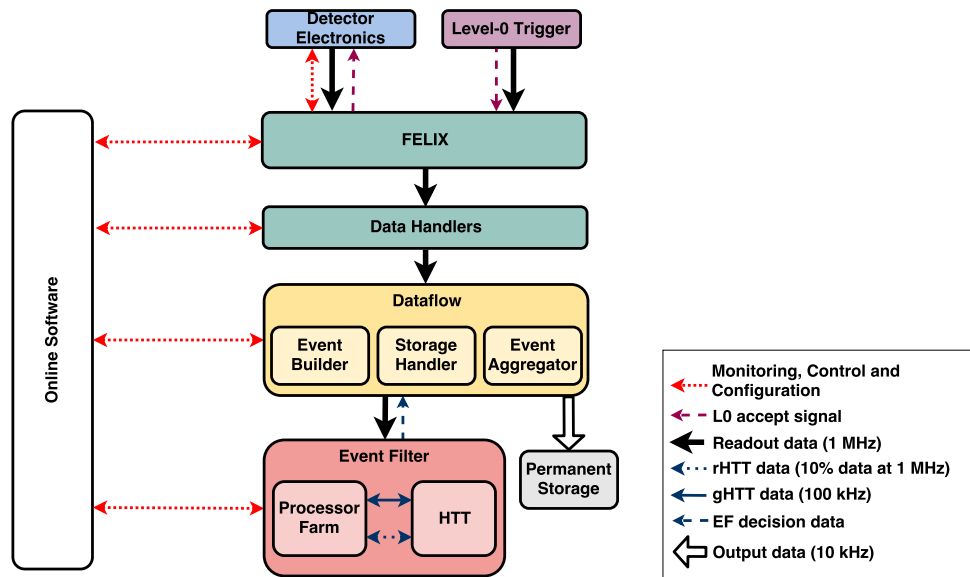

Figure 1. TDAQ architecture for Phase-II upgrades [2].

FELIX-based readout will be introduced by the Phase-I upgrade on several ATLAS sub-detectors including the New Small Wheel, the Liquid Argon Calorimeter, the Level 1 Calorimeter Trigger, and the Muon Trigger System, replacing the ROD and ROS by FELIX, which will reduce the number of sub-detector specific components. It is foreseen that the rest of the ATLAS sub-detectors replace their ROD and ROS elements by FELIX-based readout and SWROD during the Phase-II upgrade.

FELIX supports different uplink (detector to FELIX) and downlink (FELIX to detector) protocols. The FULL mode is designed to transfer packets that are multiples of 32-bit words, with no maximum packet size over a 9.6 Gb/s link with a payload of 7.68 Gb/s due to 8b/10b encoding. The GBT mode [5] transmits 120-bit packets over a 4.8 Gb/s speed link, which includes a 4-bit header, a 4-bit slow control word, 80 bits of user data (3.2 Gb/s), and a 32-bit Forward Error Correction (FEC) block. The 80 bits of the user data can be further split into E-

links, which is a variable width logical link on top of the GBT protocol that makes it possible to separate different streams on the same physical link. This mode is designed to be radiation hard, and it is implemented in the GBTX 130 nm CMOS ASIC [6] designed for front-end electronics.

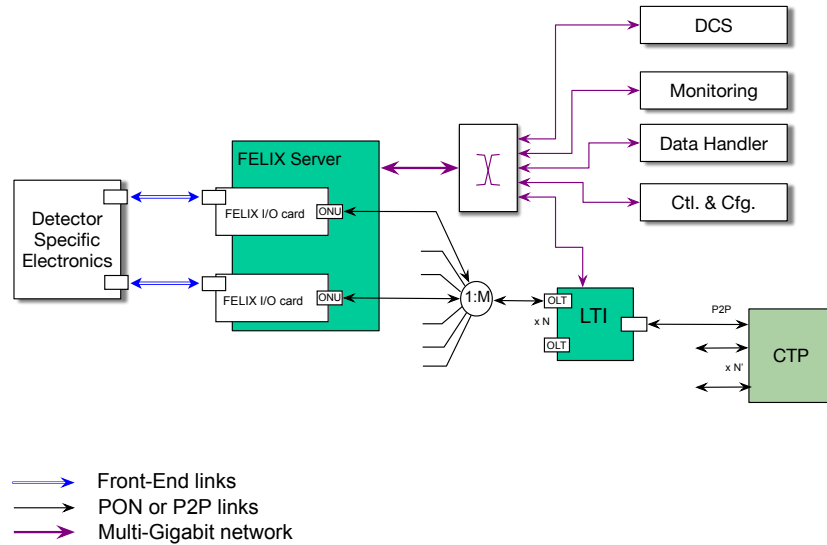## 3. The FELIX detector interface

Figure 2. FELIX structure for Phase-II upgrades [2].

The FELIX structure, shown in figure 2, is based on custom I/O cards hosted in a commodity server with a high bandwidth network connection. There are up to 48 links on each FELIX card, that can handle 48 links in GBT mode or up to 24 links in case of FULL mode in the Phase-I system. For Phase-II, it is foreseen to handle up to 24 lpGBT (Low Power GBT) links. The FELIX development platform is based on a SuperMicro X9SRL-F or X10DRG-Q mainboard with 2 PCI express generation 3 slots equipped with a Broadwell CPU E5-1650V4 running at 3.6 GHz, along with a Xilinx Virtex 7 evaluation board, VC709, so-called Mini-FELIX, that is used as the FELIX I/O card with 4 SFP+ optical links. The system is completed with a FGPA Mezzanine Card (FMC) known as TTCfx equipped with an ADN2814 clock recovery chip and a SI5345 jitter cleaner, and a Mellanox ConnectX-3 PCI express card with 56 Gb/s Infiniband connectivity. The FELIX hardware selected for Phase-I is based on a similar SuperMicro motherboard (X10SRW-F), but in this case the FELIX I/O card is a custom design (FLX-712) based on a Xilinx Kintex Ultrascale XCKU115 with 48 MiniPOD optical links with the clock recovery and jitter cleaner chips incorporated. This board also accepts a Timing mezzanine card to support TTC in its actual form, or the TTC-PON or white-rabbit standards. Finally the interconnection with the wider DAQ system is done through a Mellanox ConnectX-5 card that provides 25 Gb/s or 100 Gb/s Ethernet.

The FELIX group provides the full software stack for Scientific Linux 6 and CERN CentOS 7 operating systems. Including the kernel drivers, the low level API, the command line interface tools, and the FELIX core application, which itself provides a network interface communication layer called NETIO for the high level connectivity to the SWROD, or the ITk calibration software. In the most recent version of the software, the FELIX core application

has been streamlined to an event driven software architecture, where the event loop has no synchronisation, no queues, and no data copy. The application has two threads running per FELIX I/O card for the uplink, since there is one process per DMA and 2 DMA per PCI endpoint. The downlink has independent threads for slow control, monitoring, and busy.

The performance of the system in GBT mode is evaluated for the Phase-I requirements using a FELIX system equipped with 2 I/O cards. Each card is configured to have 24 GBT links, and 8 E-links per link (10-bit wide E-links). Upon external trigger input, each card generates an event with 40 bytes of data that is pushed through the system out to a receiving PC. Under these conditions, the FELIX system is expected to handle 100 kHz trigger rate, which corresponds to 38 Mb/s per E-link or 12.2 Gb/s per FELIX. The measured performance is 3.5 times higher than expected, reaching 112 Mb/s per E-link or 42 Gb/s per FELIX. Correspondingly, FULL mode for Phase-I is evaluated using a FELIX with only one I/O card, with 12 FULL mode links, and 5 kB data blocks generated upon every trigger request. Similarly, the FELIX system is expected to handle 100 kHz trigger rate, which corresponds to 4 Gb/s per FULL mode link, and the system is capable of handling 6.4 Gb/s, or the equivalent of 160 kHz trigger rate. A useful test on the way to meeting Phase-II requirements are evaluated using FULL mode with 464 bytes of data per trigger. Considering the expected trigger rate for Phase-II to be 1 MHz, the current FELIX implementation outperforms this requirement by a factor 2, reaching 74 Gb/s rate per FELIX.

4. The ATLAS ITk

The ATLAS ITk will be an all silicon tracker detector to replace the current ATLAS inner detector. The main characteristics of the ITk are the extreme radiation tolerance up to 100 MRad, the high granularity, and the low material budget with a tracking coverage up to $|\eta|=4$. The ITk layout will consist of a strip detector made of four barrel layers and six end-caps per side at a distance larger than 362 mm from the beam line, and a pixel detector with five barrel layers from a radius of about 39 mm to about 271 mm. The latter will be completed by a forward structure of concentric rings arranged at different radii. The full surface foreseen for the pixel detector will be of about 14 m^2. The pixel layers of the ITk will be based on a hybrid pixel detector technology, in which a sensor is glued to a read-out chip. Two types of sensors are considered, 3D-Sensors in the innermost layer, and planar sensors in the other layers. The pixel sizes vary from 100×25 μm^2 or 50×50 μm^2. The trigger requirement in the outer layers varies from 1 to 4 MHz depending on the scenario, and the data transmission is up to 4×1.28 Gb/s per front-end.

Within the context of the ATLAS ITk, the ITk Pixel outer barrel demonstrator program has been set up to drive the integration of an outer barrel stave. The program includes mechanical, thermal and electrical tests of the stave itself, while also demonstrating readout and off-detector control capabilities. The demonstrator stave is 1.8 m long with 6 serial powering chains and 2 cooling lines, equipped with up to 120 of the existing FEI4 ASIC [7] used in the innermost layer of the current ATLAS Pixel detector, arranged as 32 inclined dual and 14 flat quad sized modules.

The integration of the FELIX detector interface with the ITk demonstrators has been carried out in small scale systems that represent a vertical slice of the detector, from the FEI4 front-end modules, up to the ITk calibration software. As depicted in Figure 3, the FEI4 front-end modules receive commands at 40 Mb/s, and send data at 160 Mb/s. The data from up to 20 FEI4 front-ends is aggregated into one GBT link at 3.2 Gb/s through means of a GBTX ASIC, and plugged into a FELIX I/O card that is controlled by the FELIX core application, and interfaced to the ITk software through the NETIO protocol. Because the FEI4 front-end modules are AC coupled, a sub-detector specific module is required in the the FELIX firmware for the transmission of commands, that is based on a 'Manchester' encoding where each bit is
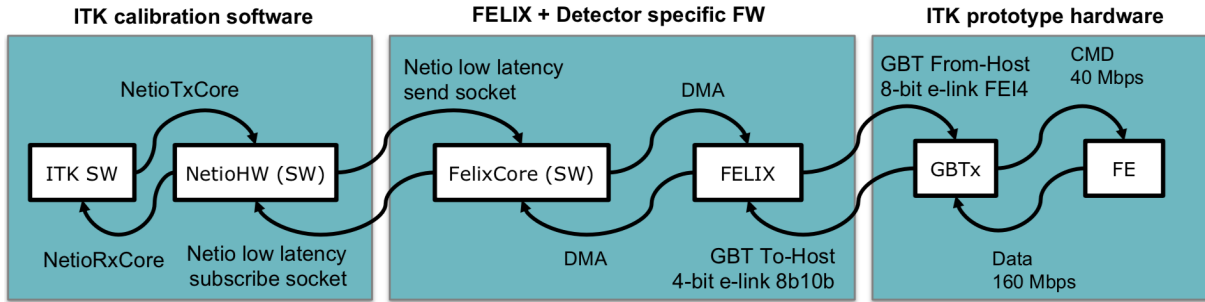
Figure 3. Schematic representation of the command and data paths during a scan of the ITk front-end prototypes

either low then high, or high then low, for equal amounts of time in the communication.

The demonstrator software is based on the implementation of what is known as the FEI4 'tuning sequence' using FELIX. This tuning is specific to the FEI4 and is an iterative process that involves many sub-processes known as 'scans'. Each scan is a computing intensive process that handles a large data volume. For example, one of the goals of the tuning is to define the threshold at which the pixel will produce a hit, which should be the same for all pixels in the chip. This is done with two different scans, one that sets the threshold globally for the whole front-end, and one that adjusts the threshold per pixel. The requirement for the first one is that the mean value of the threshold distribution matches the expectation, the second scan requires that all pixels have a similar threshold distribution. It is possible that the mean value of the threshold distribution changes after the per pixel threshold scan, thus requiring the execution of a second global threshold scan.

The ITk software required for calibration runs is a framework that supports the implementation of multiple read-out systems like FELIX or the RCE [8]. It will also allow the implementation of several front-ends like the FEI4, or the RD53 [9], which is the next pixel front-end prototype for ITk. The components of the software required for a calibration run are the scan, that steers the calibration in a series of nested loops over the parameter of the chip, and gathers the data, a decoder that processes the raw data into event data, a histogrammer that fills histograms from event data, and the analysis that accumulates the results from the histograms and provides feedback to the scan in several cases. One of the characteristics of the current implementation is the fact that the communication with each front-end at each step of the scan is done in series.

A FELIX-based tuning sequence has been successfully implemented, with the performance measured as a function of the number of front-ends, with up to two GBT links at a time. The scan time measured as user, real or system time, shows a clear linear dependence with the number of front-end chips for both the digital and analog scans, as shown in Figures 4 and 5 respectively. The CPU load as a function of number of front-ends shows a steep increase for lower number of chips, and a high value for relatively few front-end chips. Naively extrapolating from these results, if it takes 30 seconds to scan 10 front-ends, it would take approximately 4 hours to make a single scan on all of the 5000 modules of the ITk Pixel. Complementary measurements on the full scale demonstrator will provide valuable insight to the current implementation. However, areas of improvement have been already identified. In particular in the implementation of the scan. As mentioned before, at each step of the scan the settings of each front-end are addressed individually. However, FELIX allows for addressing each front-end independently and in an asynchronous way. Since the tuning of each front-end is a completely independent operation, it means that the tuning of each front-end could be implemented in a fully parallel manner, were
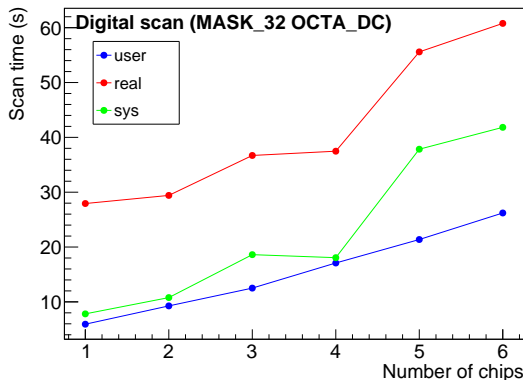
Figure 4. Scan time as a function of the number of front-ends for a digital scan.
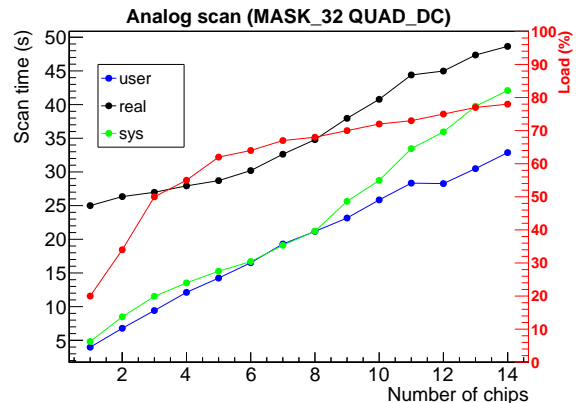

Figure 5. Scan time as a function of the number of front-ends for an analog scan.

each front-end would run in an independent process. In conclusion, the linear dependency with the number of chips is expected, and the optimal scaling would be the one where there is no dependency on the number of front-ends in the scan.

5. Outlook and next steps

The FELIX detector interface for ATLAS will be installed in staged fashion, first into few sub-detectors in the Phase-I upgrade of ATLAS by 2021, and then into all sub-detectors by the Phase-II upgrade in 2026. The current implementation achieves the required Phase-I performance with emulators, and the integration with ITk pixel is progressing well.

The integration with the next Pixel detector prototype chip for the ATLAS ITk Pixel, the RD53A read-out chip, has already started. It requires an intermediate PILUP [10] board for data formatting, and allows the read-out through FELIX in FULL-mode. The plan is to develop a full sized demonstrator by the end of 2019 based on the RD53A chip. This will require a 20 fold increase in the data bandwidth to FELIX, thus requiring the installation of a 100 Gb/s data network that is planned for the surface integration site at CERN. This will allow further integration steps with Phase-I FELIX and SWROD components.

References
[1] The ATLAS Collaboration 2008 *JINST* **3** S08003
[2] The ATLAS Collaboration 2017 TDR for the Phase-II Upgrade of the ATLAS TDAQ System Tech. Rep. CERN-LHCC-2017-020 CERN
[3] The ATLAS Collaboration 2017 TDR for the ATLAS Inner Tracker Pixel Detector Tech. Rep. CERN-LHCC-2017-021 CERN
[4] O Boyle et al The S-Link interface specification URL `https://cern.ch/s-link`
[5] P Moreira et al 2009 *CERN Yellow Reports* URL `https://cds.cern.ch/record/1235836`
[6] K Wyllie, et al 2012 *Physics Procedia* **37** 1561–1568
[7] M Garcia-Sciveres et al 2011 *NIM A* **636** S155 – S159
[8] R Herbst et al 2014 *2014 IEEE NSS/MIC* pp 1–4
[9] RD53 Collaboration 2017 The RD53A Integrated Circuit Tech. Rep. CERN-RD53-PUB-17-001 CERN
[10] A Gabrielli et al 2018 *Proceedings of Science* **321** 076