UC Irvine
UC Irvine Previously Published Works

Title

A new ATLAS muon CSC readout system with system on chip technology on ATCA platform

Permalink

Journal

ISSN

Authors

Bartoldus, R
Claus, R
Garelli, N
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

A new ATLAS muon CSC readout system with system on chip technology on ATCA platform

R. Bartoldus,[a] R. Claus,[a] N. Garelli,[a] R.T. Herbst,[a] M. Huffer,[a] G. Iakovidis,[b]
K. Iordanidou,[b,c,1] K. Kwan,[d] M. Kocian,[a] A.J. Lankford,[e] P. Moschovakos,[b,f] A. Nelson,[e]
K. Ntekas,[b,f] L. Ruckman,[a] J. Russell,[a] M. Schernau,[e] S. Schlenker,[g] D. Su,[a] C. Valderanis,[h]
M. Wittgen[a] and S.C. Yildiz[e,2]

[a]*SLAC National Accelerator Laboratory, Stanford, CA 94309 U.S.A.*

[b]*Brookhaven National Laboratory, Upton, NY 11973, U.S.A.*

[c]*University of Athens, 106 79 Athens, Greece*

[d]*The Chinese University of Hong Kong, Hong Kong, China*

[e]*University of California, Irvine, CA 92697, U.S.A.*

[f]*National Technical University of Athens, 159 73 Athens, Greece*

[g]*European Organization for Nuclear Research (CERN), Genève, Switzerland*

[h]*Johannes-Gutenberg-Universität Mainz, 55099 Mainz, Germany*

E-mail: cenk.yildiz@cern.ch

---

[1]Current Address: Nevis Laboratory, Columbia University, Irvington NY, U.S.A.
[2]Corresponding author.

ABSTRACT: The ATLAS muon Cathode Strip Chamber (CSC) backend readout system has been upgraded during the LHC 2013-2015 shutdown to be able to handle the higher Level-1 trigger rate of 100 kHz and the higher occupancy at Run-2 luminosity. The readout design is based on the Reconfigurable Cluster Element (RCE) concept for high bandwidth generic DAQ implemented on the Advanced Telecommunication Computing Architecture (ATCA) platform. The RCE design is based on the new System on Chip XILINX ZYNQ series with a processor-centric architecture with ARM processor embedded in FPGA fabric and high speed I/O resources. Together with auxiliary memories, all these components form a versatile DAQ building block that can host applications tapping into both software and firmware resources. The Cluster on Board (COB) ATCA carrier hosts RCE mezzanines and an embedded Fulcrum network switch to form an online DAQ processing cluster. More compact firmware solutions on the ZYNQ for high speed input and output fiberoptic links and TTC allowed the full system of 320 input links from the 32 chambers to be processed by 6 COBs in one ATCA shelf. The full system was installed in September 2014. We will present the RCE/COB design concept, the firmware and software processing architecture, and the experience from the intense commissioning for LHC Run 2.

KEYWORDS: Electronic detector readout concepts (gas, liquid); Modular electronics; Data acquisition concepts

Contents

1 The ATLAS Cathode Strip Chambers

1.1 Chambers

ATLAS [1] Cathode Strip Chambers (CSC) are multiwire proportional chambers designed to detect muons in the high pseudorapidity region: $2.0 < |\eta| < 2.7$.[1] The system consists of two endcaps placed in the ATLAS Small Wheel (figure 1a), with 16 chambers in each endcap (8 big, 8 small chambers for seamless coverage). Each chamber has 4 precision (η) layers with 192 strips in each layer, with pitch 5.31 and 5.36 mm for big and small chambers, and 4 transverse (ϕ) layers with 48 strips in each [2]. As the CSC is very close to the interaction point, there is a high rate of particles going through the chambers, making the readout a challenging task. The position of the CSC in ATLAS can be seen in the display of the cosmic event in figure 1b.

1.2 On-detector electronics

The on-detector (or front-end) electronics of the CSC consist of 5 Amplifier Storage Module (ASM-II) board for each chamber. For each channel a bipolar pulse with 70 ns peaking time is created and sampled every 50 ns. Each board handles 192 channels and stores the samples on a 144 cell analog

---

[1]ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upwards. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.
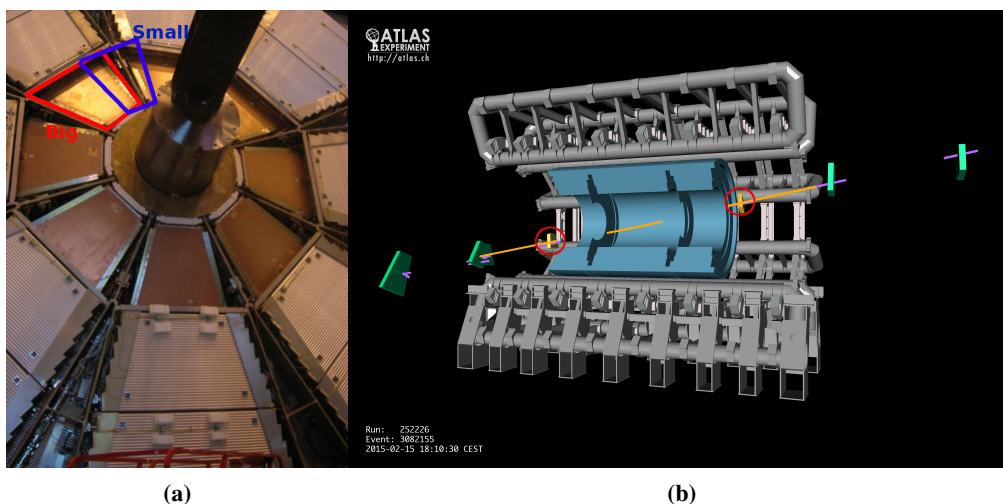
Figure 1. (a) ATLAS Small Wheel and the CSC chambers. Small chambers are not seen in the photo as they overlap the big chambers from behind (b) Cosmic event that passes through the CSC detectors on both endcaps. The CSC chambers that were hit are shown in red circles.

memory Switched Capacitor Array (SCA). Each sample is digitized to 12 bits of data, and the front-end can ship two or more of those samples to the off-detector readout system, or so called Readout Driver (ROD) via high speed fiberoptic G-Links. There are a total of 10 G-Links from a chamber to ROD, that carry the data of 96 channels each, and 5 G-Links from ROD to chamber for controlling ASM-II boards [3]. As there is no zero-suppression on the front-end, 5.76kB of data is created per event, creating an input of 4.4Gbits/sec per chamber at 100kHz of Level 1 (L1) trigger rate. The ROD then must perform threshold cut and cluster finding, resulting in a data reduction up to a factor of 60.

The Run-1 ROD was limited to run at maximum rate of 75kHz, inducing high deadtime as L1 rates went beyond 75kHz [4].

For Run 2, LHC bunch spacing was to be decreased from 50ns to 25ns, energy to be increased to 13 TeV and luminosity increased by having higher number of collisions per bunch crossing. Thus development of a new ROD was needed to handle L1 rate of 100kHz and higher luminosity, that correspond to an average of 50 interactions per beam crossing (compared to average of 20 interactions per beam crossing in the Run 1) [5].

2 ATCA based generic Data Acquisition system

The development of the CSC Run-2 off-detector readout system was based on a generic Data Acquisition (DAQ) concept that has been developed at SLAC since 2007. This section will explain the concept, while the next section will focus on the implementation on the CSC.

2.1 Platform

Advanced Telecommunication Computing Architecture (ATCA), which was originally developed for the telecommunication industry, was selected as the host platform for the development of the DAQ system. It is becoming increasingly popular in high energy physics due to the high speed backplane,
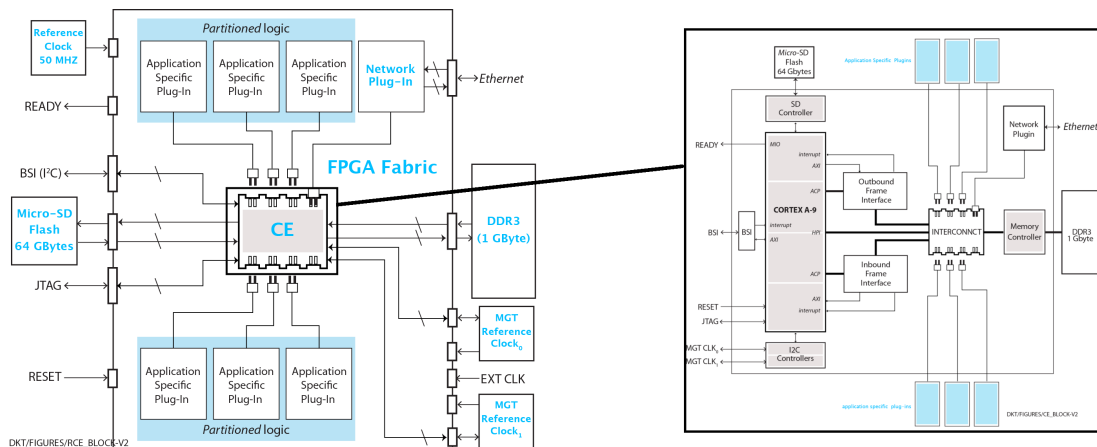
Figure 2. Block diagram of the Reconfigurable Cluster Element and Cluster Element.

its high availability, hot swapability of boards and the Intelligent Platform Management Interface (IPMI) based shelf management infrastructure. The backplane is protocol agnostic, providing pin-to-pin connections between cards, allowing the user to select the communication protocol, and the Rear Transition Modules (RTM) allow separating the data processing from the input/output.

2.2 Reconfigurable Cluster Element

Reconfigurable Cluster Element (RCE) is the computational element, which is a bundled set of hardware, firmware and software components. It is based on System-On-Chip (SoC) technology (Xilinx ZYNQ) and aims to provide a high performance and efficient use of resources. It has a large FPGA fabric, high speed I/O channels, large memory banks, and a strong interconnect that allows the CPU to access all these resources at high speed. The heart of the RCE is the Cluster Element (CE) and it consists of a dual-core A-9 ARM processor clocked at 800 MHz, 1 GB of DDR3 RAM, and up to 64 GB of flash memory. One has several options to run as operating system such as: Real-Time Executive for Multiprocessor Systems (RTEMS), Linux or bare metal.

RCEs are implemented in two forms. The Data Processing Module (DPM) RCE is used by the application and consists of a Zynq XC7Z045-2FFG900E with 16 Multi-Gigabit Transceivers (MGT) to perform data processing and high speed I/O. The Data Transfer Module (DTM) RCE consists of Zynq XC7Z030-2FBG484E with 4 MGTs. It is shared among several DPMs, and is responsible for distributing application specific signals to DPMs and managing the networking among the RCEs. Block diagram of the RCE and CE can be seen in figure 2.

2.2.1 Protocol Plug-In model

Protocol Plug-in (PPI) is defined as an arbitrary set of application specific logic, that resides in the RCE's FPGA fabric and requires the exchange of information with its CE. The plug-ins work with a plug and socket model. There are 8 predefined sockets on the RCE, and specific applications can be implemented in the firmware/software and plugged into these sockets using a wrapper. The PPIs can either act as an I/O device, defining the data transfer protocol, or take advantage of Digital Signal Processing (DSP) tiles and combinatoric logic of the FPGA to process data. The whole model allows one to implement hardware solutions in firmware/software with much lower footprint and power consumption.
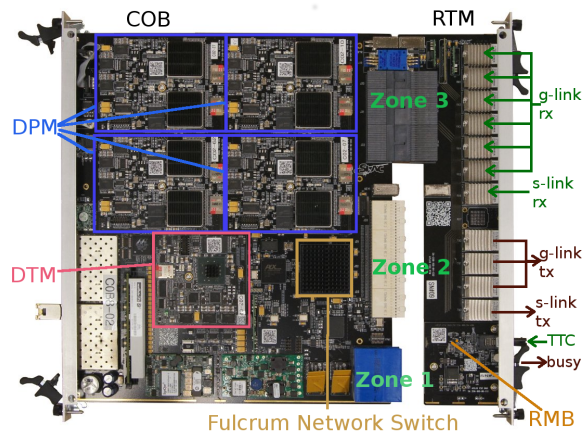
Figure 3. The CSC COB and RTM.

2.3 Cluster-On-Board card

The carrier board developed to host RCEs is called the Cluster-On-Board (COB) card. It hosts up to 8 processing RCEs on 4 DPMs and one control RCE on DTM.

The COB has a Cluster Interconnect (CI) section that consists of an embedded 24-port 10GbE low latency Fulcrum switch as its main feature. A network connection at the front panel Ethernet port of a single COB in the ATCA shelf can be distributed to all RCEs on all COBs on the shelf. The COB has 3 rear connection zones. The Zone 1 connector is used for power and management, the zone 2 connector is the data transport interface for communicating between boards, and the zone 3 provides connectivity between COB and RTM in a 96-channel high density connector. This arrangement allows the active processing elements to remain on the COB as generic resources while the RTM only hosts the system specific configuration of I/O connection ports. The COB can be seen in figure 3.

3 CSC Run-2 off detector readout system

3.1 Overview

The CSC Run-2 ROD is an implementation of the RCE concept. The COBs that host the RCEs reside in a standard 6 slot ATCA shelf, each COB hosting three active DPM bays (6 RCE, each related to a single chamber), and one DTM bay (One RCE). The CSC specific RTM is the physical interface between the chambers and the ATLAS Trigger and DAQ (TDAQ) system. For each DPM RCE, there are 15 G-Link fiber connections to a CSC chamber, 10 to receive data from chambers and 5 to send control signals to the on-detector electronics. Each DPM RCE also has two S-Links to the Read Out Subsystem (ROS); one to receive commands, and one to transmit formatted event fragments. Finally, for each COB, a single Busy signal is sent to the ATLAS Busy Module to halt trigger in case any of the RCEs are busy; trigger, Timing and Control (TTC) signal is received from the TTCex module to be distributed to all DPM RCEs.

Each DPM RCE handles a single chamber and runs RTEMS as OS. There is a dedicated firmware logic called the SCA controller for requesting 4 or more time slices from the front-end whenever there is a L1 trigger. The G-link PPI is used to decode the data, which is immediately sent

to the feature extraction (FEX). The FEX PPI performs the threshold cut, while the FEX software does the clustering and builds the fragments according to the CSC data format. The formatted data is sent out to the ROS, using the S-Link PPI. The interaction with the TTC system is done via TTC Receive and Busy PPIs, with the use of the DTM, which is shared by all DPMs in the COB.

The DTM runs Arch Linux as OS, and its main responsibilities are: managing the network traffic; combining all the Busy signals from the DPM RCEs and directing it to the Busy output of the RTM; and receiving the TTC signal from the TTC input of the RTM and distributing it to the DPMs. Furthermore, a TTC emulator functionality is implemented in the DTM, allowing it to reproduce the functionality of the Local Trigger Processor (LTP) as an internal TTC source so that the COB can be a more versatile standalone system.

The G-Link PPI, S-Link PPI, and the DTM's ability to handle the TTC signal are perfect examples of how the large FPGA fabric is used to replace classical hardware solutions. For G-Links and S-Links, G-link ASICs and High-speed Optical Link (HOLA) cards are replaced by DPM firmware, while for the TTC handling, the former 9U VME TTC Interface Module is replaced by DTM firmware. All this allows not only to reduce the footprint, so that one can squeeze more components in unit space, but also to have much lower power consumption, so that these components don't generate more heat than the system can handle.

The control processor (CP) is a server PC with dual network interface to act as a gateway between the ATCA shelf and the external world, interacting with both sides. Only a single COB is connected to the CP via a dedicated switch, while the network assignment of the remaining COBs are done internally in the ATCA fabric.

3.2 Integration with ATLAS

The system is integrated with the TDAQ infrastructure and Detector Control System (DCS). The ATLAS TDAQ is a collection of systems that combine the triggering, event selection, data collection from detectors, event building and data recording. With all the filtering, the collisions at 40MHz are filtered down to 1kHz to be written to storage. On the other hand the DCS allows one to monitor and control the detectors and experimental infrastructure.

The TTC and Busy transmission between the CSC and the ATLAS Central Trigger Processor (CTP) is done through the intermediate unit LTP via the connections on the RTM.

The formatted data is sent to the ROS via S-Links, where it is buffered, and sent out for event building upon request of High Level Trigger (HLT).

The interaction with the ATLAS data taking finite state machine is done on the CP, which communicates with the RCE in a server/client model using Remote Procedure Calls (RPC) as the inter-process communication method. During an ATLAS data taking, each RCE is represented by a separate process in the CP. The CP can send configuration parameters to RCEs, and retrieve the status of each RCE. In addition, the status of a complete set of internal state registers are gathered and sent to the Information Service (IS is a ATLAS service that is used to share variables between DAQ applications, and log them) every 5 seconds for debugging and diagnosis. The CSC system makes use of ATLAS TDAQ automatic recoveries, such as resynchronizing an RCE during data taking, or reconfiguring the whole system.

For the pedestal runs, a dedicated histogramming service runs on the FEX software to perform a fast calibration for threshold calculation.
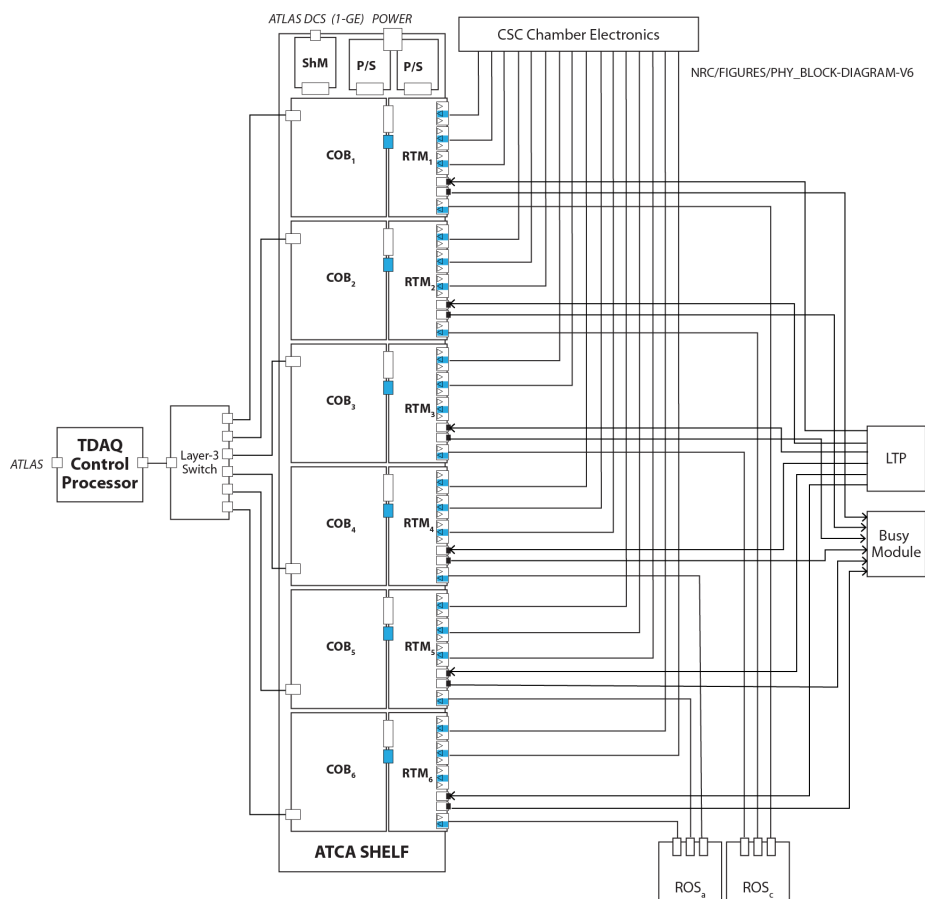
Figure 4. Schematics of the CSC off-detector readout system connections.

Finally, the health of the ATCA shelf, COBs and RCEs is monitored in DCS via the shelf manager on the ATCA making use of the IPMI Controller of the COB based on the Pigeon Point System solution. Informations such as power supply status, COB health or temperature at several points on the shelf is monitored, with specific alarms implemented to warn experts if any of the parameters go out of working limits.

All the connections between CSC off-detector readout system, CSC detector, LTP, ROS and the network are shown in figure 4.

3.3 Performance

Performance depends on three fundamental components of the system: front end, FEX and ROS. The front-end performance has not changed in the Run 2, because it depends on the buffering capacity and G-link bandwidth, which limits the system to 111 kHz for 4 samples readout.

The data transfer speed to ROS depends on the Read Out Link (ROL) output bandwidth, which was doubled by expanding to twice as many actual S-links. The S-link PPI driving at 100MHz has also extended the individual S-link capacity to a maximum of 200MB/s.

The FEX performance is the main improvement in the Run-2 readout system. FEX on the new system is implemented both in the firmware and software, threshold cut being done on the FEX
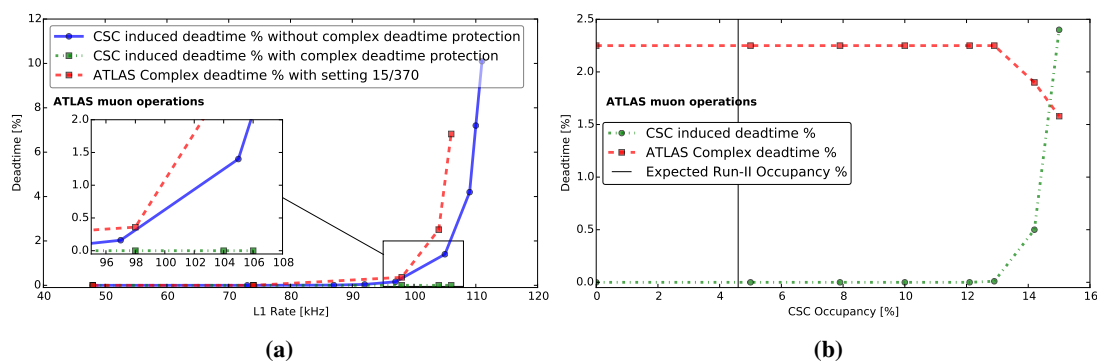
Figure 5. (a) The CSC induced deadtime and complex deadtime with Run-2 occupancy of 4.6%. With the complex deadtime protection of the CSC setting 15/370, no additional deadtime is induced by the CSC even at the maximum possible rate allowed by the complex deadtime setting. Without any protection, the CSC still runs without causing deadtime until 96kHz. Note that dashed and solid lines are obtained from two separate tests (b) The CSC induced deadtime vs. the occupancy of the CSC for input rate of 100 kHz. The CSC starts to induce deadtime in much higher occupancies compared to the expected occupancy in Run 2. As CSC start to induce deadtime, the trigger rate, thus the complex deadtime starts to drop.

PPI, while clustering and formatting is done on the software. FEX uses On-Chip Memory (OCM) for faster access, and makes use of assembler for the most performance critical parts.

As the intervals between triggers are randomly distributed during data taking, detector readout systems should be protected against a burst of triggers that would cause it to run out of front end buffers. In ATLAS, a complex deadtime mechanism that works in a leaky bucket model is used to provide this protection [6]. ATLAS complex deadtime system in Run 2 can have 4 different complex deadtime settings, where each setting is represented by two parameters: "bucket size"(X) and "leak rate"(R). The CSC has adapted the setting 15/370 that was already used by other systems, thus bringing no extra complex deadtime to data taking. The mechanism works in the following way: each L1 trigger is added to bucket, and one is removed every R bunch crossings. Whenever the bucket is full (contains X triggers), CTP asserts deadtime until the next leak. Due to this, the deadtime induced by the mechanism only depends on the time distribution of the triggers.

With the L1 rate of 100kHz, and expectation of 50 interactions per beam crossing (which corresponds to 4.6% occupancy in the CSC) the readout system is required to perform without inducing any additional deadtime. ATLAS in Run 2 so far hasn't reached such L1 rates and luminosities. Therefore, for the performance tests random triggers are generated on CTP to test the system, and the CSC occupancy is set using specific threshold patterns that simulates fake clusters.

The system was tested with L1 rates up to 106kHz at Run-2 occupancy using the CSC complex deadtime setting, and no additional deadtime was observed due to the CSC. Furthermore, fixing the input trigger rate at 100kHz, the CSC occupancy was varied from 0 to 15%, and no dead time was observed due to the CSC until 14% occupancy, which is much higher than the Run-2 expectation for the CSC (figure 5).

4 Conclusion

With the increased luminosity and L1 rate in ATLAS in Run 2, a new readout system was developed for the CSC. It is the first deployment of ATCA and RCE based modern DAQ systems in ATLAS. The main reason for development is performance improvements. The system features a powerful hardware, and matching firmware/software framework, to handle high rates and luminosities. The CSC Run-2 readout system has been running stably to date and has demonstrated the main goal of running at 100kHz and Run 2 occupancies with no deadtime.

References

[1] ATLAS collaboration, *The ATLAS experiment at the CERN Large Hadron Collider*, 2008 *JINST* **3** S08003.

[2] T. Argyropoulos et al., *Cathode strip chambers in ATLAS: installation, commissioning and in situ performance*, *IEEE Trans. Nucl. Sci* **56** (2009) 1568.

[3] I.G. Eschrich et al., *Readout electronics of the ATLAS muon cathode strip chambers*, in proceedings of *Topical Workshop on Electronics for Particle physics*, Naxos Greece September 15–19 2008 [ATL-MUON-PROC-2009-003].

[4] R. Murillo, M. Huffer, R. Claus, R. Herbst, A. Lankford, M. Schernau et al., *Software design of the ATLAS muon cathode strip chamber ROD*, *J. Phys. Conf. Ser.* **396** (2012) 012031.

[5] ATLAS collaboration, N. Garelli, *The evolution of the trigger and data acquisition system in the ATLAS experiment*, *J. Phys. Conf. Ser.* **513** (2014) 012007.

[6] ATLAS collaboration, M. Stockton, *The ATLAS level-1 central trigger*, *J. Phys. Conf. Ser.* **331** (2011) 022041.

2016 JINST 11 C01059