A Streaming Data Acquisition System for Mu2e

Mark Bowden, Kurt Biery, Ryan Rivera Fermi National Accelerator Laboratory*

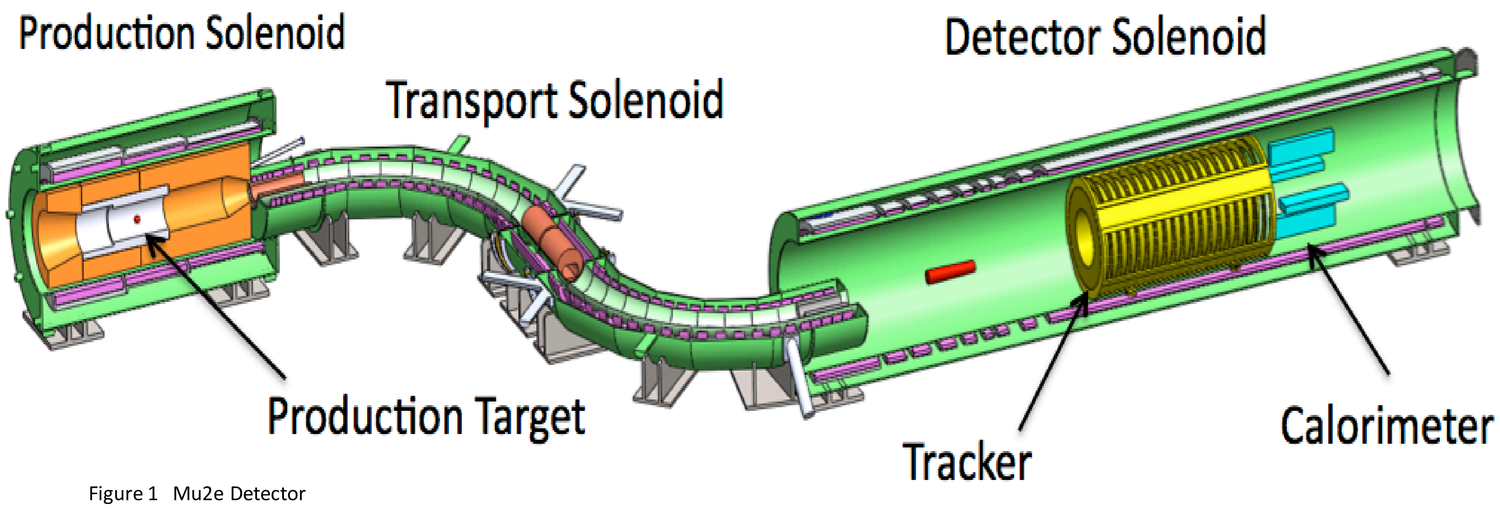

Figure 1 Mu2e Detector

Introduction

Mu2e is a proposed high-sensitivity muon to electron conversion experiment (figure 1). The detector includes a Tracker with 21,600 straw tubes and a Calorimeter with 2112 crystals. Additional data comes from the Cosmic Ray Veto system and beamline monitors.

The DAQ system for Mu2e is based on a streaming readout (no on-detector triggering). Motivations for a streaming DAQ include greater use of commodity technology (networking, processing), more flexibility in the development of trigger and reconstruction algorithms, and higher trigger efficiency.

Streaming DAQ systems with bandwidths of a few GBytes/sec are now in production (e.g., NOvA). Mu2e will increase this rate by another 50X.

Data Volume

Mu2e operates at a 1695ns pulse repetition rate (figure 2). Data is zero-suppressed on the detector, and is output in fixed timeslices containing multiple microbunches. The Tracker and Calorimeter will produce an estimated 100GBytes/sec after front-end zero-suppression. With continuous readout, the off-detector data rate will be similar to that of the CMS and Atlas detectors.

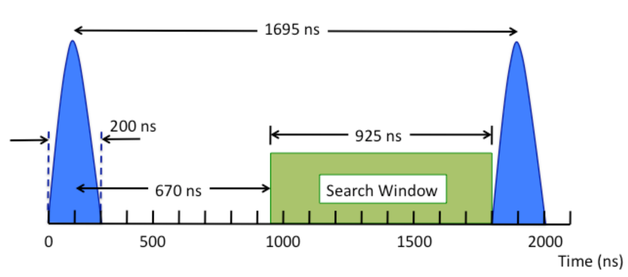

Figure 2 Mu2e Microbunch Beam Structure

Dataflow

The proposed DAQ architecture is shown in figure 3. Data Sources (Readout Controllers) are located inside the detector vacuum. Each source drives four 2.5 Gbps optical links. Each of the links connects to an independent DAQ Block, and data timeslices are transmitted "round-robin" to each Block. Links are bundled in groups of twelve fibers per MTP cable and sent to an optical patch panel in the counting room where they are routed to the DAQ Blocks. Three DAQ blocks are implemented in the initial system configuration. The 4th link from each data source is used for redundancy, but can also drive another DAQ block if rates increase. Segmenting the DAQ into blocks allows the system to continue operation at a reduced rate if a block is offline. It also reduces the cost of the event building network.

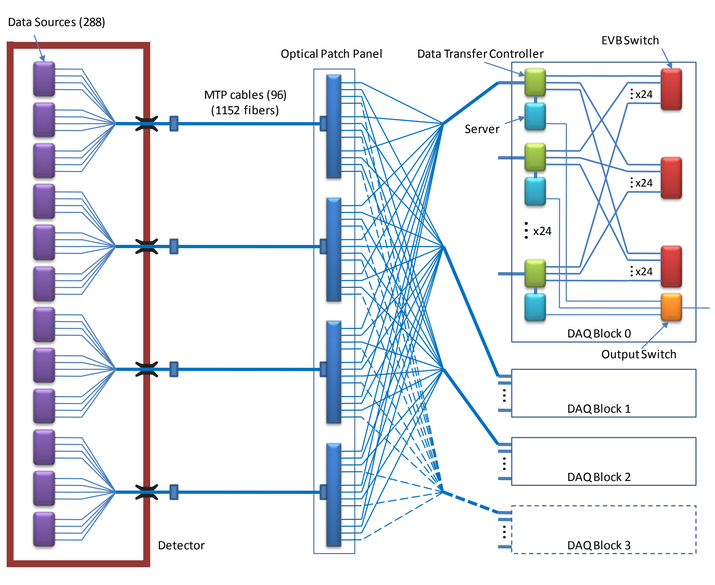

Figure 3 Proposed Mu2e DAQ Architecture

Data Transfer Controller

The Data Transfer Controller (DTC, figure 4) has three ports;

1) Input Port - a 12 channel X 2.5 Gbps optical (MTP) connection to receive data from the detector readout controllers.
2) Output Port - a 3 channel X 10 Gbps copper (SFP+) Ethernet connection to send data to the online processors.
3) EVB Port – a 3 channel X 10 Gbps copper (SFP+) Ethernet connection to exchange data with other DTCs via the Event Building (EVB) Network.

The DTC can operate in either 2-port or 3-port modes. In the 3-port mode, event building is done entirely in DTC hardware and fully assembled events are delivered to the online processors. In this mode the DTC can also pre-process the event data in FPGA firmware to offload the processors. In the 2-port mode, the EVB Port is bypassed and the Output Port is connected to the EVB Network. The output of the EVB Network connects to the online processors and all event assembly and processing is done in the processors.

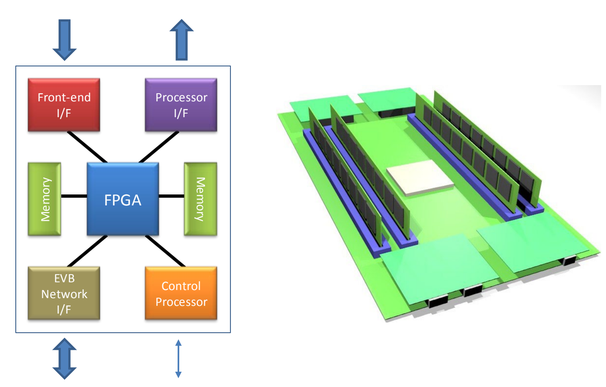

Figure 4 Data Transfer Controller

Event Building Network

The Event Building Network in each DAQ Block consists of three 24-port 10G Ethernet switches, for a bandwidth of ~70 GBytes/sec per block (~200 GBytes/sec total). Connections are direct-attach SFP+ copper cables. The event building function is distributed, with the first stage (3:1) taking place in the on-detector readout controllers, the second stage (also 3:1) in the DTCs, and the final stage (24:1) in the EVB network. This reduces the number of switches required to implement the final event building stage (figure 5).

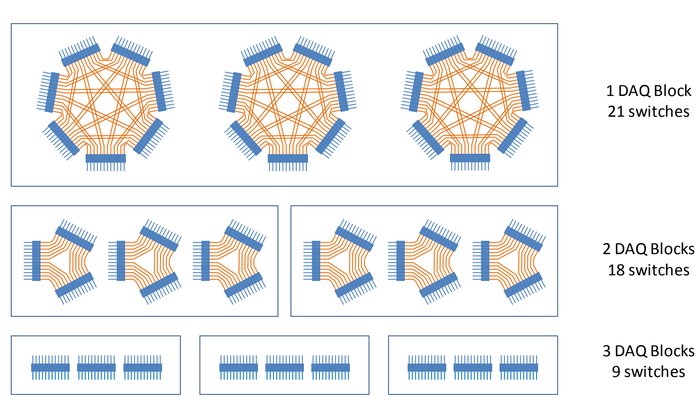

Figure 5 EVB Network vs. Number of DAQ Blocks

Processing

"Many-Core" servers designed for large data centers may also be useful in HEP online processing farms. Pictured (figure 6) are examples of 512 core servers from Quanta (Tilera) and SeaMicro (Intel). Servers based on ARM and Intel MIC processors are under development. These systems typically use ½ to ¼ the power and space of traditional servers, and include integrated "top-of-rack" networking infrastructure.

Figure 6 "Many-Core" Servers

DAQ

The Mu2e DAQ subproject is responsible for most of the electronics external to the detector. This includes the optical links, DTCs, Event Building Network, Online Processing Farm, Fast Control and Timing, Slow Control Network, and associated software (figure 7).

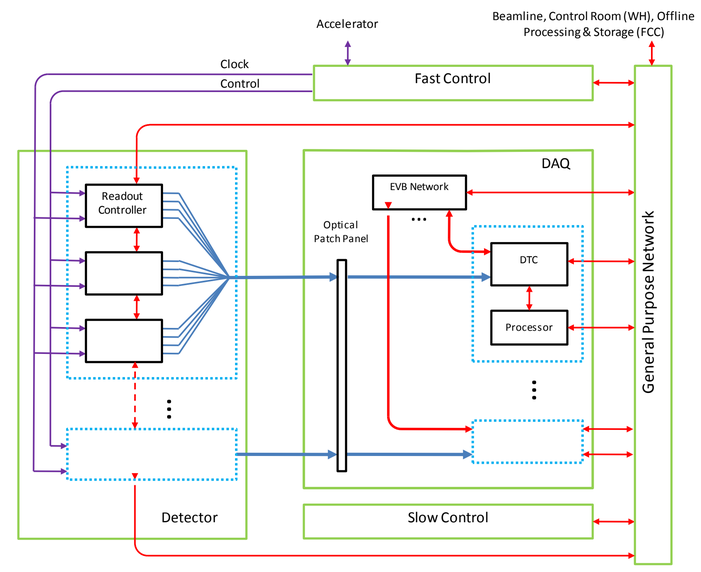

Figure 7 DAQ System Interfaces

Control and Timing

The Timing system (figure 8) is relatively simple, consisting of a low jitter 1695ns clock synchronized to the beam and a 2.5Gbps serial control link. Clock phase alignment/calibration is handled by the detector electronics. The Control Link supports multicast addressing, and control messages can be synchronized to the system clock or timestamp counter at the receivers. The Slow Control network is based on Ethernet (or EtherCAT).

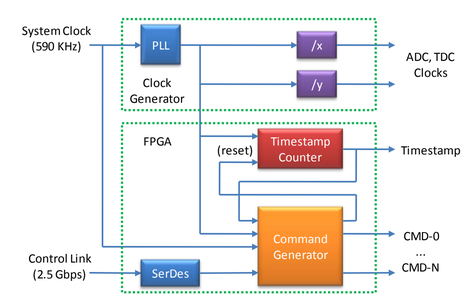

Figure 8 Timing System (front-end interface)

Software

The online software trigger will use the same event analysis framework as the offline data analysis, with suitable modifications to handle data input and output in an online environment. A simplified model of the data flow inside the processor farm is shown in Figure 9. Many events will be reconstructed and analyzed in parallel on each processor node to make full use of the multiple cores that will be present, and to achieve the necessary event rejection factor of several thousand.

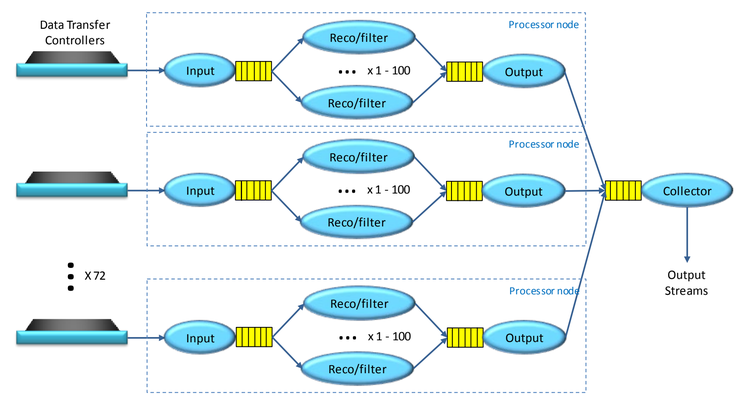

Figure 9 Software

Control and monitoring applications in the DAQ will make use of components and design patterns from other recent experiments at Fermilab (e.g., NOvA), and the Slow Controls system will be based on an existing framework such as EPICS.

Summary

The MU2e DAQ system will implement a streaming ("triggerless") readout of detector data at a rate of 100 GBytes/sec. Continuous improvements in processing and networking technology make this feasible.

* Operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy.