Date: December 11, 2019

EIC Detector R&D Progress Report

Project ID: eRD23

Project Name: Streaming readout for EIC detectors

Period Reported: from 7/1/2019 to 12/30/2019

Project Leader: M. Battaglieri and J. C. Bernauer

Contact Person: M. Battaglieri and J. C. Bernauer

Project members

J. Huang, M.L. Purschke
Brookhaven National Laboratory, Uptown, NY

S. Ali, V. Berdnikov, T. Horn, M. Muhoza, I. Pegg, R. Trotta
Catholic University of America, Washington DC

M. Battaglieri, M. Bondí, A. Celentano, L.Marsicano, P.Musico, S. Vallarino
INFN, Genova, Italy

F. Ameli
INFN, Roma La Sapienza, Italy

D. K. Hasell, C.Fanelli, R. Milner
Massachusetts Institute of Technology, Cambridge, MA

J. C. Bernauer
Stony Brook University, Stony Brook, NY and Riken BNL Research Center, Uptown, NY

E. Cline
Stony Brook University, Stony Brook, NY

C. Cuevas, M. Diefenthaler, R. Ent, Y. Furletova, G. Heyes, B. Raydo
Thomas Jefferson National Accelerator Facility, Newport News, VA

Abstract

The detectors foreseen for the future Electron-Ion Collider will be some of the few major collider detectors to be built from scratch in the 21^{st} century. A truly modern EIC detector design must be complemented with an integrated, 21^{st} century readout scheme that supports the scientific opportunities of the machine, improves time-to-analysis, and maximizes the scientific output. A fully streaming readout (SRO) design delivers on these promises, however, it can also impose limitations on the sub-detectors and electronics. The streaming readout consortium will research the design space by evaluating and quantifying the parameters for a variety of streaming readout implementations and their implications for sub-detectors by using on-going work on streaming readout, as well as by constructing a few targeted prototypes particularly suited for the EIC environment.

1 Past

The fifth streaming readout workshop was held at the Riken BNL Research Center (RBRC) from November 13 to November 15, 2019 to discuss current and future work in the streaming readout community. Detector and electronics groups presented progress in their field and discussed future development in the context of experiments under construction and looking forward as EIC development and design progresses.

Over the course of this very successful workshop a wide-ranging set of streaming readout topics were discussed by the community.

Data Rate Requirement & Management: The larger experiments represented in the workshop have a definitive advantage in handling their data rate. For experiments with the appropriate funding, it is possible to interface with companies directly and have chips built to custom specifications. ATLAS, for example, is prototyping a new Versal FPGA chip that is not currently commercially available. The scale of the EIC will allow for close contact with industry professionals to develop electronics suitable for our needs. However, the turn-around time for such developments are long, and input for developments required for EIC need to be defined and communicated without much delay to be compatible with the envisioned timeline of an EIC.

Data Format: The format chosen must be robust and fault tolerant. Small instances of data corruption should not render an entire file unreadable. The data format should not have size limitations that make handling too difficult. Data from sub-detectors will be organized in different streams and written to different files in parallel. We discussed optimal file sizes in the context of tape storage and the risk of corruption. Instead of creating and closing files synchronized over all streams, *i.e.* a given time segment of one file for one stream would have a corresponding file with the same time segment for all other streams; files are created and closed asynchronously to accommodate the vastly different sub-detector rates and realize files of optimal size for tape archival. We estimate the risk from lost tapes minimal if some care is taken in arranging files of the same time period to be stored on the same tape.

Online/Offline Convergence: Traditionally, online and offline environments have been very distinct with calibrations and Level-3 (L3) triggering somewhere in-between. Driven by a combination of the physics being studied and detector technologies, the offline-online distinction is being blurred, and in some cases removed: 1) Detectors such as TPCs or pixel vertex detectors that are difficult to integrate into the traditional online data flow defer integration processing until later. 2) Detector topologies that are hard/expensive to deal with in electronics, defer to software. 3) L3 triggers require partial or complete reconstruction to deduce the rate to mass storage, and trigger rates and detector topologies that are hard/expensive to deal with in electronics, defer to software. For the EIC, we propose to remove the offline-online distinction completely and extend the machine-detector interface (MDI) to a machine-detector-analysis interface where the DAQ and analysis are fully integrated with the MDI to optimize physics reach. In this session, we have discussed a prototypes for a seamless data processing from DAQ to data analysis based on streaming readout.

Detector Support: Specific detector designs depend on the experimental needs. Certain experiments need readout electronics to be close to the sensitive area of a detector due to material length requirements. We discussed available platforms and the opportunity for cooperation with the private sector to close gaps in the current capabilities.

Accelerator interface: Some data needs to be made available to accelerator controls almost immediately for proper beam diagnostic and tuning. We discussed several ways how this can be realized in a streaming readout setup, and are confident that this can be handled appropriately. At the same time, it is clear that the detector readout design has to be coordinated with the accelerator design to make sure that the required data can be provided within the required maximum delay.

Timing We discussed the timing accuracy requirements and possible solutions. While some detectors might have special needs, the timing synchronization between detector doesn't have to be much better than the beam clock. We compared possible distribution schemes like White Rabbit, but agreed that a simple, dedicated clock distribution network, phase locked to the accelerator clock, is the simplest and most useful solution. Care has to be taken that PLL locks inside the system have enough locking range to allow for frequency shifts during energy ramps, if existent at the final EIC design.

Industry: Representatives from CAEN attended the workshop and expressed willingness to interface with the community and develop a board to the mutually agreed upon specifications. The community stressed the importance of an open, modular and well documented design and emphasized that a fully working software suite provided by the manufacturer is far less important than the ability to write own driver and firmware code.

High Rate Experiences: The high performance computing sector has experience handling huge amounts of data and performing complicated calculations. However, the benefits of modern compute accelerators can only be exploited with suitable data organization which favors arrays of vectors instead of vectors of (complex) arrays. The latter is the prevalent data model in nuclear physics software, but streaming readout naturally favors the first.

To improve the collaboration with the HPC community, we identified having a workfest-type meeting between streaming readout enthusiasts and computing groups to work on some concrete and similar problems as an ideal kick-off point.

Organization: It would be beneficial to start on documentation to define the streams from different EIC sub-detectors. Individual groups have been working independently but it would be beneficial to the community if there were a central repository of documentation for all groups to see. It would make implementation more difficult if every detector creates a completely different format for readout. A single protocol for every detector would be the logical way forward for the EIC. A prototype design has been defined and is being tested in different groups. Feedback from the EIC Detector R&D community is highly welcome to develop the design further.

Electronics: It would be very helpful to organize all available ASICs or FE boards into a table that contains not only their operating specifications but also their time to market. Waiting for a chip that is many years away from production is not productive, but encouraging the industry to fast track something the community finds useful should be targeted.

Differential Cost/Complexity/Capabilities: An accurate cost comparison between a triggered system and a streaming readout system is non-trivial as it strongly depends on the final detector layout, channel count and detection principle. However, some general statements can be made. The simple, traditional triggered systems without provisions to reduce dead time are not suitable for a modern detector design. This comparison will only discuss the more complex modern triggered designs, hybrid designs, and full streaming readout.

In most modern triggered designs which avoid dead-time, the signal is converted continuously and without the presence of an external trigger. The often zero-suppressed data is then stored in a memory buffer. The fixed-latency trigger then selects the memory region to read out, while the other content is dropped. In a streaming system, either all of the data is stored, or the memory buffer is moved from the front end to the main memory of a computer.

There are few readout ASICs which use memory on the analog side, to reduce the required sampling speed or the required number of the ADCs. However, only very few can hide the dead-time of this approach, and do this only up to a certain event rate.

In the context of EIC, we assume that all front end electronics will be specialized developments. While some of this development may be based on existing designs, substantial design work is needed independent on the general readout philosophy. It is the opinion of the members of this eRD project, which includes the leading electronics engineers of BNL and JLAB, that the development of a streaming system is easier than the development of a dead-time reduced triggered front end, as the latter is essentially the former, with the addition of the memory buffer, trigger interface and data selection logic.

The current projections for the physics data rate, including moderate detector noise, puts the total at 100 GBits/s. This is substantially lower than the rates at the running LHCb experiments or what is planned for sPHENIX, which will run before EIC. Such a rate is straight forward to handle, and all digitized detector signals could be recorded by a simple streaming readout system, without any dead-time. As an upper limit, it has been estimated that a pixel vertex detector with 20 to 50 million channels would produce about 240 GByte/s of raw data. While it is not realistic to save the full data stream of such a detector to disk—area/time of interest readout or on-the-fly analysis has to reduce it—the actual data transport is easily feasible even with today systems.

We want to note that the majority of the data rate of this estimate stems from the TPC or pixel detector. For the former and at the rates of EIC, in combination with the drift times for typical TPC sizes, only a streaming readout solution is sensible—the events overlap often enough that the TPC has to be readout essentially all the time anyway.

The crucial question is the detector noise rate—hits above reasonable thresholds which stem from background or dark current. If this noise is large, it is not feasible nor helpful to save all of the noise. In a triggered system, the amount of noise related data is reduced, as the detector information is only stored if a trigger condition was fulfilled.

In the equivalent implementation in a streaming readout system, the CPUs/FPGAs reading out the detector participating in the trigger decision to identify data of interest not via an electric signal, but by distributing a list of the time slices of interest to the other readout CPUs via the network. While conceptually similar, such a system is more flexible – for example, the readout can keep data from not-interesting time slices if the data rate maximum is not reached, to facilitate calibration etc. The trigger decision can be more complex and take longer since the maximum latency is not limited by on-chip or on-front-end memory, but by the memory available in the readout servers.

Further, the decision doesn't have to be fixed-latency, since the time slices are identified by their number, not by the occurrence of a signal with a fixed offset. Such a system is simpler, as no dedicated trigger lines have to be distributed to every front-end board and subsequently timed in. Instead, the communication happens over the COTS network, and the decisions are made with software, reducing the debugging and development complexity.

The downside is an increased minimum latency for the decision, and with that, the minimum amount of memory required, as the network incurs delays on the millisecond scale. However, this disadvantage only comes into play for extreme data rates. For example, a typical desktop PC with 16GB can store the whole physics rate suggested above for more than a second, and even a hundred-fold increase requires only about $5k of RAM at *current* prices. Such memory could buffer the upper-limit estimate above for multiple seconds.

This means that RAM-based temporary storage is feasible even for high-level event reconstruction on the multiple-second scale if the noise rate is not too high, allowing for very selective data selection criteria and reduced long-term storage requirements without loss of physics.

Next Steps: We discussed the next steps (see below), and agreed that it is not necessary for the streaming readout community to wait for site selection to begin making progress on designs and development. Test stands can be created from existing hardware and taken to test beams to assess performance and measure noise and background levels. We believe that a CD1 detector design will need to discuss the read out electronics, making timely progress on a detector-wide readout design a pressing requirement.

Accomplishments by the individual groups:
The following activities were not supported by the eRD23 funds directly, but are relevant for the development of the field.

- SBU/RBRC: RBRC organized the workshop described above. The protocol specification was tested on a test-beam experiment at DESY, at which a calorimeter was read out simultaneously using a traditional triggered DAQ and a streaming DAQ. Analysis is ongoing.

- JLAB: FADC250 based system (half crate: 128 FADC 250Msps 12bit channels) was used in INDRA-ASTRA Lab to demonstrate that digitized pulses could have charge/time extracted and streamed the JLAB VTP. The VTP framed hits into 50 us time bins of variable sized TCP packets sent over a single 10 Gbps Ethernet link to a server using the JLAB CODA SRO (Streaming Readout Protocol). Plans are to scale the system up to use a full crate (and multiple crates) as well as multiple 10 Gbps Ethernet links from each crate in a Hall B beam test in February 2020. Figure 1 show the VXS crate setup with FADC250 and VTP modules and the Ethernet data rate vs. FADC250 hit rate. At 4Gbps an efficiency loss begins to develop - this point will be much closer to 10 Gbps when the hardware accelerated TCP stack is recompiled with a larger window size (even as-is, a $\sim$1 MHz hit rate per channel is a very conservative limit for most applications).

  A project sponsored by hall-A is evaluating SAMPA based readout of a GEM detector using hardware based on designs originally developed for ALICE/ATLAS. A cosmic ray test stand uses two scintillators and two 384 channel GEM detectors. The test is operational and being read in streaming mode. JLAB is also working with INFN to integrate their TRIDAS framework, discussed later, with hardware and software being developed at JLAB.

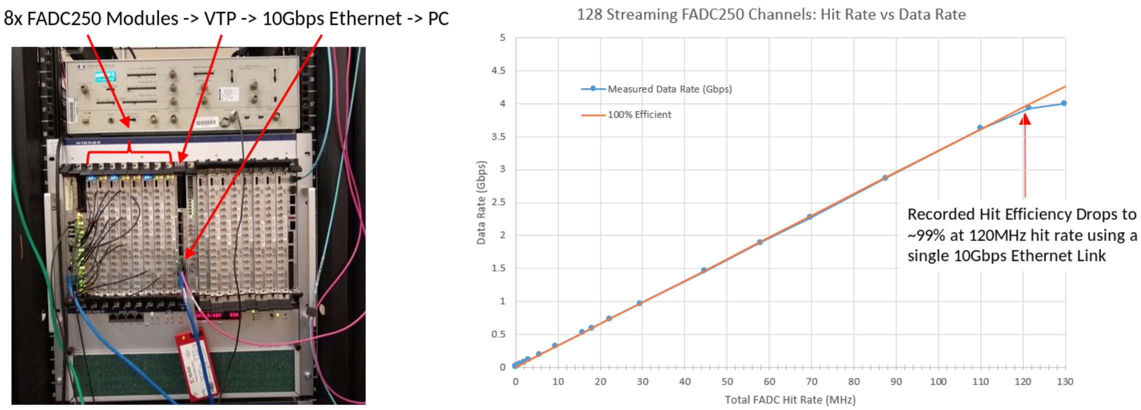

Figure 1: Left: Test-setup with eight FADC250 modules. The data stream is sent via a VTP card and 10G Ethernet to a PC. Right: Hit-rate vs. recorded data rate. Up to more than 100MHz hit rate, all data can be recorded. At 120MHz, about 1% of hits are lost. This can be mitigated in the future via an optimized FPGA firmware.

- BNL: At the Relativistic Heavy Ion Collider (RHIC), the forward silicon tracking detector (FVTX) at PHENIX and the e-going Time-of-Flight (eToF) at STAR are already read out in a streaming mode. Our current research focuses on adopting the state-of-the-art readout technology and large-scale streaming detector system for an EIC. Specific areas include following:
  - the ATLAS FELIX DAQ interface is designed to bridge custom front-end with commodity computing with 48x 10 Gbps bi-directional optical link and 100 Gbps PCIe Gen3 link to a hosting server, which would be suitable for the use in an EIC streaming DAQ. Since the last report, the manuscript on FELIX including its application in the Nuclear Physics experiments is published in *IEEE Transactions on Instrumentation and Measurement*, [doi: 10.1109/TIM.2019.2947972]. Production of FELIX v2.1 for both sPHENIX and CBM experiments has stared and we are following the next generation FELIX interface prototype development which will support of PCIe Gen4 and 25 Gbps optical transceivers.
  - the RCDAQ Data Acquisition system[1], in use by sPHENIX and several eRDx groups, supports a number of streaming-readout hardware components, such as the FELIX-based readout of the SAMPA ASICs, or the DREAM readout. The DREAM electronics, while not strictly a trigger-less system, is read out as a stream because there is no concept of a deterministic end of the data of a given trigger. At different test beam setups at Fermilab earlier this year, GEM-based detectors have been read out using the DREAM system, and a TPC prototype has been read out in streaming mode using the FELIX system.
  - Real-time computing: Streaming loss-less data compression and throughput for the `ZeroMQ` protocol over a 100 GbE network have been studied varying a comprehensive set of parameters as shown in Figure 2.

[1]`https://github.com/sPHENIX-Collaboration/rcdaq`

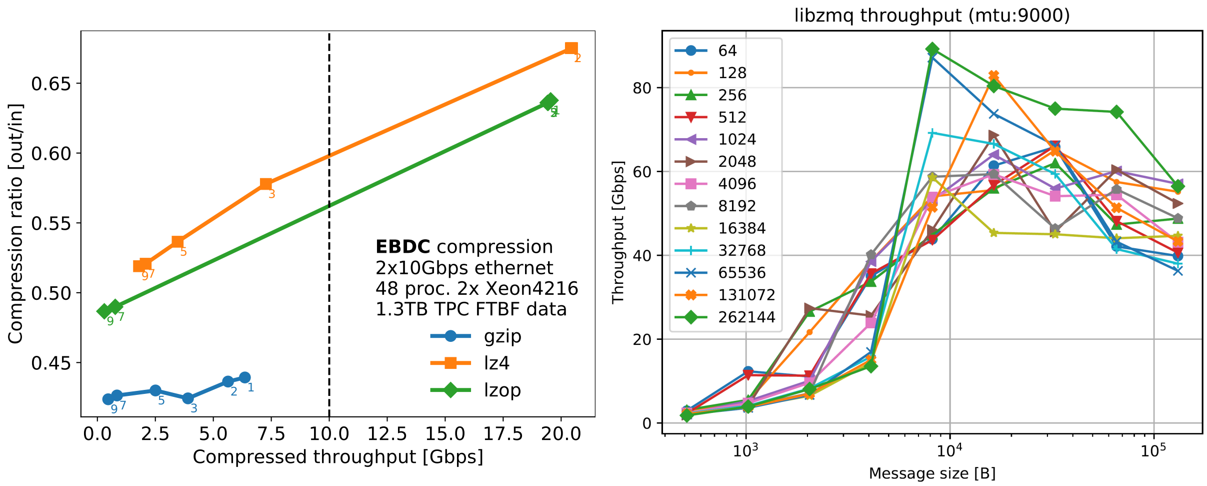

Figure 2: Throughput tests. Left: lossless data compression of the streaming SAMPA data from in the June 2016 sPHENIX TPC beam test. Both `lzo` and `lz4` algorithms at low compression levels surpass 10 Gbps throughput with a single server. Right: throughput using the `ZeroMQ` protocol via 100 GbE NIC and switch versus various message size (horizontal axis) and batch buffer size (curves), courtesy Dr. Brett Viren (BNL) and BNL LDRD 19-028. The EIC collision signal data at 100 Gbps can be processed by a fleet of such servers (e.g. sPHENIX TPC will use 24 such servers to handle a ~ 200 Gbps peak data rate).

  - Streaming digitizer development: A successful multi-project wafer prototype has been completed for the SAMPA v5 ASIC which is aimed to provide streaming readout for the sPHENIX Time Projection Chamber (TPC). Comparing to the v4 chip used by the ALICE experiment, the version-5 SAMPA will support shorter (80ns) shaping time and is optimized for 20 MHz waveform sampling rate, which would be more suitable for both sPHENIX and many EIC tracking applications. In addition, we have been developing a generic streaming digitizer platform based on the Xilinx Zynq Ultrascale RFSoC ZCU111 board[2] as shown in Figure 3. It integrates eight 4.1Gbps 12-bit ADCs, UltraScale+ FPGA fabric and ARM CPUs in a single chip, which provides a powerful testing ground for streaming signal processing at the front-end such as background suppression and feature extraction.

- MIT: Funding was not received until very late in the fiscal year and thus we were not able to collaborate with the ASIC design firm Alphacore to test their new ADC and pre-amplifier designs as originally planned.

  However, we did build a small, 3×3 array of lead tungstate crystals (loaned to us by Tanja Horn from CUA) with PMT readout. This calorimeter was tested at the DESY test beam in September. Both a triggered and a streaming readout system were used simultaneously by splitting the PMT signals. The signals from each crystal were gain matched using 5.2 GeV electron beam and then measurements were made scanning

[2] `https://www.xilinx.com/products/boards-and-kits/zcu111.html`

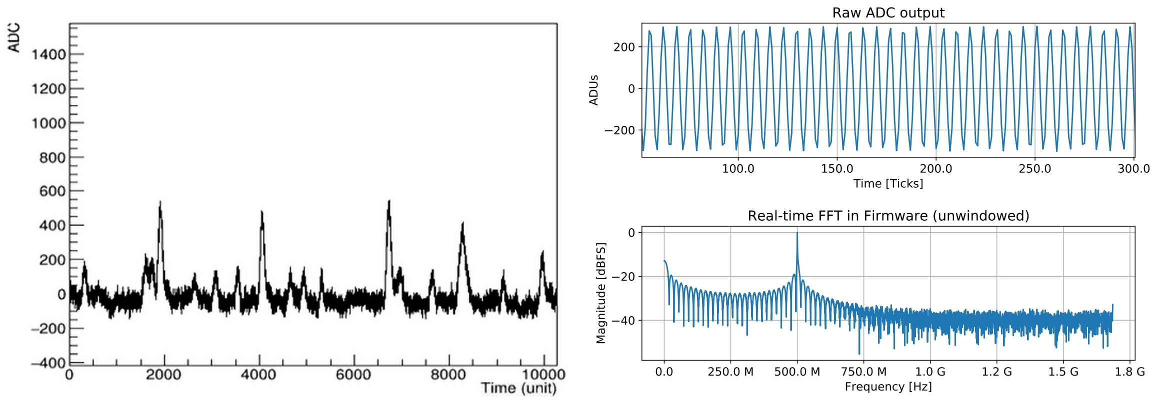

Figure 3: Work in progress with ZCU111 RFSoC board, a powerful platform prototyping streaming ADC and digital signal processing. Left: a segment of ADC waveform stream at 3.2 GSps from SiPMs amplified and shaped by the sPHENIX EMCal pre-amplifier in the high gain mode. Both the few-pixel LED pulser and single-pixel noise peaks are visible in the raw ADC waveform. Right: 500 MHz sinusoidal signal digitized at 3.3 GSps and real-time FFT on FPGA, courtesy Shreeharshini Murthy (BNL) and BNL LDRDs.

the electron beam across the central region of the calorimeter with incident energies of 2, 3, 4, and 5 GeV.

The triggered readout configuration used thin scintillators with SiPM readout upstream of the calorimeter to trigger the readout of a CAEN QDC module. To realise this configuration the PMT signals had to be delayed by approximately 75 ns to be in time with the trigger logic signal.

The streaming readout system used CAEN digitizers that recorded the signal waveform as well as determining the time and charge integral for each channel. Since the digitizers were not triggered valid signals as well as signals from cosmic rays and noise were also recorded.

- INFN: During the last 6 months we focused on three items: FE electronics, readout software and bench tests set up in Genova.
  We are currently testing the v2.0 of the WaveBoard, a custom 12 ch modular 12/14 bit 60/125/250 MHz digitizer streaming compatible designed by INFNRM and INFNGE. The WaveBoard v1.0 has been successfully used in BDX-MINI spring tests at Jefferson Lab. The v2.0 corrected some minor issues of the previous version and it is currently under tests in Genova and Rome.
  On the DAQ software front, we are currently working on the implementation of JANA libraries (currently used at JLAB by the GLUEX Collaboration as reconstruction software) within the TRIDAS framework. This would allow to implement sophisticated L2 algorithms in order to reduce the data rate to record as well as implement on-line calibration and advanced data processing AI-supported.
  The third item concerns the implementation of a streaming RO framework to test scintillating glass in collaboration with EIC-eRD1. The plan is to use an already assembled

set up that includes three ALICE MRPCs to cover a wide area (about a squared meter) to track cosmic muons with good spacial resolution. The specimen detector will be placed within the MWPC acceptance and synchronized at few ns level with the MRPC output by a time stamp provided by a GPS system. The absolute time recorded by the two subdetectors will allow to combine the different information. This set up represents a good example of the streaming RO capability.

2 Future

The available funding will be used for travel support for the workshops, outreach to other groups and other meetings. We further plan a strong engagement in the Yellow Report initiative, especially in the subgroup tasked with DAQ/electronics.
The next workshop will be held in late spring, organized by JLAB at Cristopher Newport University (Newport News). Results of the current activity at BNL, JLAB, MIT and INFN will be reported. A significant amount of time will be devoted to the discussion with the other EIC R&D consortia. In particular the eRD20 (Software Development) will be involved in the workshop organization.

The following group activities are not supported by eRD23 funds.

- SBU/RBRC: Dr. Bernauer and his group will further their role in the sPHENIX readout development. sPHENIX DAQ is similar in scope to an EIC, and will use a hybrid approach, with large streaming components. Further, the group will continue the work in collaboration with MIT and INFN on the DESY test beam, and coordinate software development with the software eRD and working group. Undergraduate student Marisa Petrusky will work with Dr. Bernauer during the spring semester on an VHDL project related to streaming readout.

- JLAB/INFN: (Note: With the position change of M. Battaglieri, the JLAB and INFN efforts will merge in the near future. We therefore list them here together).

  FADC250/VTP system will update the front-end pulse processing firmware to provide new data format modes: raw waveform and fine time resolution. These modes are intended to provide improved timing and charge resolution for test setups or experiments that require it. Further refinements of the firmware will be made to improve data transfer efficiency so that the full 40Gbps bandwidth is usable.

  In collaboration with JLAB team, the INFN group will instrument an electromagnetic calorimeter in streaming mode and test it with CEBAF beam. The TRIDAS sofware has already been installed on INDRA Lab and experimental counting house's machines. The WB v2.0 will be used as bench test for the whole system. When the DAQ will be fully working, we will instrument a crate containing JLAB fADC250. Via the VTP board, data will be streamed to the TRIDAS sofware, filtered by a CPU-based L2 trigger and recorded on tapes. Unbiased data will be also collected for further checks. Results we'll be reported to the summer meeting. If available, CUA scintillating glasses instrumented with sipms will be used. In alternative, PbWO crystals will be used.

Similar tests are also planned in collaboration with the MIT team. The same DAQ system will be used at DESY to read out PbWO crystals or scintillating glasses.

- BNL: We plan to carry out a test for the timing distribution system by installing our timing module prototype at the RHIC RF system clock source, which would provide a low glitter bunch crossing clock source under the realistic collider operation conditions such as beam dump, ramp and tuning. We will continue the work on the ZCU111 platform, in joint R&D with Packed Ultra-wideband Mapping Array (PUMA)[3]. As the sPHENIX construction has started, we will continue to gain experience on constructing a large detector system with three streaming readout detectors that constitutes the full tracking system.

- MIT: Again we plan to collaborate with the ASIC design firm Alphacore to test their latest ADC and pre-amplifier boards with a view to providing a range of useful specifications for test and development front-end electronics.

  We will also test a larger, 5×5, array of lead tungstate crystals with PMT readout at the DESY test beam in April. During this test we will use only a streaming readout scheme with CAEN digitizers. The new calorimeter will have several design improvements and will permit the crystals to be cooled. It is hoped that we will also be able to replace some of the lead tungstate crystals with new ceramic glass crystals produced at CUA by Tanja Horn.

3 Manpower

All personnel is currently funded by external sources. We report here time spend on SRO related activities, whether directly EIC-related or not.

- SBU/RBRC: Ethan Cline joined Bernauer's group and will work with him on streaming readout. Further, the undergraduate student Marisa Petrusky will work in the group during spring 2019.

- INFN/CUA: the personnel involved in the aforementioned activities at INFN is: Marco Battaglieri (senior staff scientist), Andrea Celentano (staff scientist), Luca Marsicano (PhD student), Simone Vallarino (master thesis student) and Paolo Musico (senior staff engineer) in Genova and F.Ameli (senior staff scientist) in Rome. Each of us has spent approximately 30% of the time on this activity, partially shared on synergistic activities, in particular the BDX experiment at Jefferson Laboratory.

- BNL: J. Huang, M. Purschke will commit 10-20% time developing the SRO system for EIC, which will be supported under BNL LDRD 19-028 and in synergy with on-going work on sPHENIX SRO tracking system. This work will be supported by an experienced engineering team at BNL including J. Kuczewski, J. Mead, and A. Dellapenna and in collaboration with the ATLAS DAQ team at BNL.

[3] arXiv:1907.12559 [astro-ph.IM]

- JLAB: Streaming readout work is performed by staff scientists and engineers from the JLAB DAQ and Electronics groups. G. Hayes and C. Cuevas are the respective group leaders with B. Raydo, E. Jastrzembski and J. Gu working 10% of their time on various streaming readout activities. Eric Pooser, hall-A postdoc has been working on streaming GEM readout. In the last six months the Hall-B and Hall-D DAQ experts S. Boiarinov and A. Somov joined the team. The streaming readout development activities are an ongoing extension of the hardware developed and implemented for the 12GeV experimental programs at JLAB. These activities are collaborative with the SRO consortium groups and complementary work to implement real-time calibration and analysis is also supported by JLAB LDRD-2014.

- MIT: Main effort will be through D. Hasell, R. Milner, I. Friscic, S. Lee, P. Moran, and B. Johnston working on the DESY test beam. I. Friscic will assume a shared role at JLAB and MIT working on streaming readout in the spring. C. Fanelli will work together INFN and JLAB teams to implement high level JANA-based algorithms within the TRIDAS framework.

4 External Funding

The Riken BNL Research Center supported the Streaming Readout Workshop V. BNL LDRD 19-028 supported many streaming DAQ research at BNL as discussed in Section 1.

5 Publications

Proceedings in the form of DVDs with the talks and additional material from the Streaming Readout Workshop V are in production.