

Progress Report: Streaming readout for EIC detectors

Jan C. Bernauer
for the eRD23 EIC Streaming Readout Consortium

EIC R&D meeting, BNL, January 2020



RBRC
RIKEN BNL Research Center



Stony Brook
University

Who are we: SRC members

- ▶ **Brookhaven National Laboratory:** J. Huang, M. Purschke,
- ▶ **Catholic University of America:** S. Ali, V. Berdnikov, T. Horn, M. Muhoza, I. Pegg, R. Trotta
- ▶ **INFN Genova:** M. Battaglieri, M. Bondi, A. Celentano, L. Marsicano, P. Musico, S. Vallarino
- ▶ **INFN Roma:** F. Ameli
- ▶ **Massachusetts Institute of Technology:** D. Hasell, C. Fanelli, I. Frišćić, R. Milner
- ▶ **Stony Brook University:** J. C. Bernauer (also RBRC), E. Cline
- ▶ **Thomas Jefferson National Accelerator Facility:** C. Cuevas, M. Diefenthaler, R. Ent, Y. Furletova, G. Heyes, B. Raydo

Additionally many regulars

→ We welcome new members! ←

Workshop



Workshop

- ▶ Supported by the Riken BNL Research center
- ▶ November 13-15, 2019
- ▶ **Proceedings:** Talks published on DVD
- ▶ Many topics – short summary

WS: Data Rate Requirements & Management

- ▶ Heard from current/near term experiments with large data amounts
- ▶ Scale of experiment helps: **Can talk to chip vendors**
Example: ATLAS tests Versal FPGAs not commercially available
- ▶ This is relevant e.g. for FELIX NG-type of development.
- ▶ Have to mind turn-around time. **Talk to vendors very soon!**

WS: Data format I

- ▶ Discussed file sizes and organization
 - ▶ Stream sub-detectors data to different files in parallel
 - ▶ **Do not** synchronize file creation over all files
 - ▶ **Instead**, optimize file size for tape backend
- ▶ Make sure that files are written to tape which correspond to similar times
 - ▶ Minimizes loss if tape is **unrecoverable**

WS: Data format II

- ▶ Discussed optimal data format to leverage modern compute architectures
 - ▶ **Structs of arrays (of simple structs)** instead of **arrays of (complex) structs**
- ▶ The latter is the current model, the former is natural for streaming
- ▶ A lot of experience in the HPC community, but limited cross-pollination
- ▶ **Plan to have a work-fest type meeting to kick-off collaboration**

WS: Online/offline convergence

- ▶ Working together with eRD20 / EICUG Software WG
- ▶ Goal is to develop machine-detector interface (MDI) to machine-detector-analysis interface
- ▶ Integrate DAQ and analysis to have high level analysis available at data taking
 - ▶ For data verification
 - ▶ Data reduction / experiment optimization
- ▶ Discussed AI approaches

WS: Accelerator interface

- ▶ Accelerator operation requires fast and stable information from detector
- ▶ Typically rather low level.
- ▶ Discussed how this would be realized, **no principle problems found.**
- ▶ **But will need continued coordination between detector and accelerator design**
 - ▶ Specify required information and maximum latency

WS: Detector support

- ▶ Started to compile list of available solutions for different detectors
- ▶ Generally: Some detectors will need specialized electronics to meet performance and cost requirements
 - ▶ Example: GEM-TRD: Works beautifully with streaming-capable fADC125. Does not work with APV, DREAM, VMM. Might work with modified SAMPA.
- ▶ Still enough time to work on ASICs, but time windows is closing. Need to come up with requirements now.

WS: Timing

- ▶ Discussed timing requirements and possible solutions
- ▶ Discussed **White Rabbit** (time via Ethernet) as a distribution model
- ▶ But decided that an **accelerator-synchronous approach is easier and better**
 - ▶ I.e. dedicated clock distribution
 - ▶ Need to check PLL lock and acquire freq. ranges if accelerator ramps

WS: Industry

- ▶ Representative from CAEN presented their upcoming designs
- ▶ All streaming based.
- ▶ Agreed to work on open up platform.
 - ▶ We need less out-of-the-box solutions, but ability to modify to our needs.

WS: IP available to others

- ▶ We collected a list of what IP or even hardware is available for others for test setup etc.
- ▶ BNL: Gerbers, drivers, firmware, especially for FELIX and SAMPA interfacing
- ▶ JLAB: Several hardware designs, e.g. 192 FPGA ROC board with MAROC3 ADC readout and 1ns TDC

WS: Next Steps

- ▶ We believe that a CD1 detector design needs to include a readout solution
- ▶ Continue our current efforts, and participate in the yellow report process
- ▶ Continue developing test stands and small scale streaming experiments

WS: Differential Cost/Complexity/Capabilities

- ▶ A full-scale cost comparison of a triggered vs streaming readout is not feasible without fixing final detector layout etc.
- ▶ Can make general statements:
 - ▶ Simple triggered systems won't work at 0.5MHz collision rate. **Too much deadtime or lost physics.**
 - ▶ Compare modern triggered design, hybrid designs and full streaming readout

WS: Differential Cost/Complexity/Capabilities

- ▶ A full-scale cost comparison of a triggered vs streaming readout is not feasible without fixing final detector layout etc.
- ▶ Can make general statements:
 - ▶ Simple triggered systems won't work at 0.5MHz collision rate. **Too much deadtime or lost physics.**
 - ▶ Compare modern triggered design, hybrid designs and full streaming readout
- ▶ To avoid deadtime, modern triggered designs have streaming frontends.
 - ▶ Fixed latency trigger selects window out of memory buffer
- ▶ Some ASICs do this on the analog side – but hard to hide dead time
- ▶ Assume that EIC development is "from scratch". In this case, streaming readout incurs **less cost on front end electronics development.**

WS: Differential Cost/Complexity/Capabilities

- ▶ Current projections for physics rate ≈ 12 GByte/s, below sPHENIX.
 - ▶ Can stream everything to disk
- ▶ Upper end: pixel vertex with 20-50 mio channels would produce 240 GByte/s of raw data
 - ▶ Still reasonable to handle in realtime, but cannot write it out.
 - ▶ Need noise suppression, Region/Time of Interest read-out
- ▶ In this case, rate is fully driven by TPC or pixel vertex. Other detectors have small rates.

Cross over point between Streaming vs. Hybrid

- ▶ In a streaming readout, the data reduction from a trigger would be achieved by tagging time segments for storage – others are dropped.
 - ▶ But better (max latency, data selection,...)
- ▶ Typical PC today can buffer > 1 second of the physics rate.
- ▶ 100 times higher rates (or 5 times worst case) is about \$5k/second today.

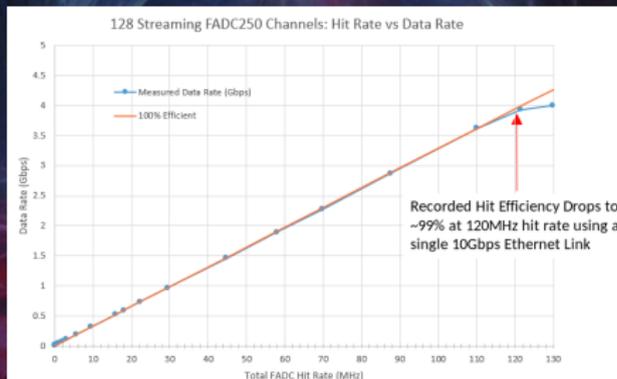
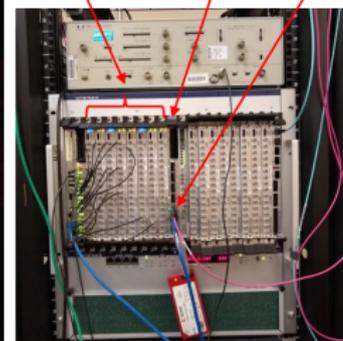
Progress at the institutes

The following activities are not funded by eRD23.

Progress: JLab

- ▶ fADC250 based system used in INDRA-ASTRA lab for streaming tests. So far 128 channels (half crate).
- ▶ Streamed via JLAB CODA SRO protocol
- ▶ Efficiency drop above 4 Gbps. Optimized firmware will hit 10 Gbps.

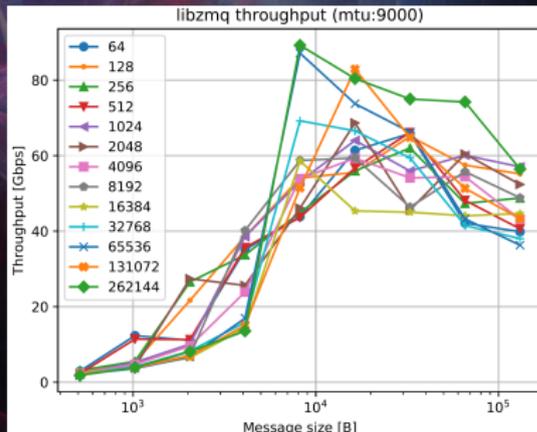
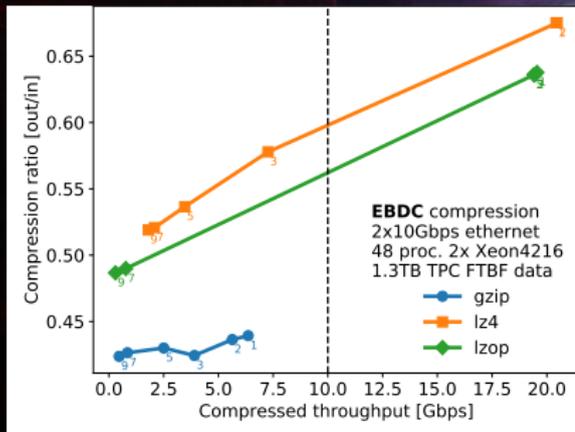
8x FADC250 Modules -> VTP -> 10Gbps Ethernet -> PC



- ▶ On-beam test this/next week: FT calorimeter/hodoscope with ≈ 500 ch, measure $eA \rightarrow (X)e'\pi_0 \rightarrow (X)e'\gamma\gamma$
 - ▶ Different data selection modes, from "all" to online cut on π_0 mass
- ▶ Test in April: PbWO prototype with offline coincidence between a triggered pair spectrometer and calorimeter(triggerless)

Progress: BNL

- ▶ FVTX@PHENIX and eTOF@STAR already streaming in hybrid configuration
- ▶ ATLAS FELIX: 48x10GBps bi-directional interface card
- ▶ RCDAQ DAQ software
- ▶ SAMPA based streaming digitizer
- ▶ ZeroMQ-based streaming over 100 GbE + lossless data compression



Progress: SBU/RBRC and MIT

- ▶ SBU/RBRC: Organized the workshop :)
- ▶ MIT: Funding from DOE was received late, so could not work with Alphacore on ASIC design
- ▶ Test beam time at DESY, parallel read out of calorimeter via triggered and streaming electronics
- ▶ Calorimeter was constructed from tungstate crystals on loan from Tanja Horn (CUA)
- ▶ **Analysis ongoing**

Progress: INFN

- ▶ Testing v2.0 of WaveBoard digitizer board (12 ch, 12/14 bit, 60-250 MHz)
- ▶ JANA integration of TRIDAS
- ▶ SRO framework for testing eRD1's scintillating glass
- ▶ Preparing tests at JLab and DESY with TRIDAS and WB 2.0

How much time do you envision to complete your ongoing project(s)

- ▶ We will continue developing streaming readout for the EIC continuously. This is open ended.
- ▶ Conceptual design of a streaming readout system naturally has the same timeline as the conceptual design of a full detector design.
- ▶ Unfunded streaming readout projects have different time lines, many of which will end before 2023.

What achievements are required for TDR readiness 2023

- ▶ All labs are planning test setups + experiments in the near future which will make use of streaming readout. All of these activities should have been carried to fruition before 2023.
- ▶ CLAS12
 - ▶ Aims to convert to fully streaming in the time scale of few years.
 - ▶ Then will make full use of SRO capabilities
- ▶ sPHENIX
 - ▶ Hybrid system
 - ▶ Similar conditions: Integration time \times collisions >1 , and similar data rates
 - ▶ Uses RHIC RF clock source, which will likely basis of EIC RF source.
- ▶ These experiments will demonstrate the benefits of SRO and the readiness at EIC scale.

Future: Next workshop and other activities

- ▶ Next workshop will be at **Christopher Newport University**, organized by JLAB
 - ▶ **May 13-15**
 - ▶ Existing and future developments in SRO hardware
 - ▶ Software framework / connection with EIC software
 - ▶ Validation: Experiments performed/planned to test SRO in real life
 - ▶ Requirements and constraints for EIC detectors w.r.t. SRO
- ▶ **YR: SRO represented by A. Celentano (subconvener), process to define detector requirements and integration has started.**
- ▶ **Will use funding to support travel to workshop, to yellow report meetings and to test beam times.**

Thank you!



Backup slides

Hic sunt dracones

Last review

An interesting beginning of a cost-benefit analysis of possible readout architectures focusing on minimizing custom trigger hardware and relying, as much as possible, on commercial computing for data selection has been presented. Without any particular set of requirements from a particular set of detector subsystems it is not clear that one can yet make any firm conclusions on where an optimum design might lie. It should be noted that all the variants mentioned assumed on-detector zero suppression and so some level of local “triggering”. This topic will be of increasing interest as detector ideas converge and as real detector design communities emerge.

Recommendation: Modest funding is suggested to continue to encourage discussions of options and development of the means of completing a detailed cost benefit analysis once more is known about the EIC detector(s).

Current State

- ▶ We are on the verge of this.
 - ▶ Require detector and accelerator specifics to make stronger cost statements.
- ▶ Full cost analysis and, more importantly, design of a complete readout solution will come with complete detector design.
- ▶ So far, we can make general statements:
 - ▶ SRO will have many advantages over the full stack from FEE to analysis.
 - ▶ SRO is viable for almost all detectors.
 - ▶ SRO required for some detectors.
 - ▶ In fact, a modern triggered system is very often streaming with added-on trigger.
 - ▶ SRO requires less hardware, more cost effective except in rare/extreme cases
- ▶ We are prepared to transition from broad development of basic concepts to focused development in coordination with the detector design

Cross over point between Streaming vs. Hybrid, ext vers.

- ▶ In a streaming readout, the data reduction from a trigger would be achieved by tagging time segments for storage – others are dropped.
 - ▶ No requirement for fixed latency, just bound latency.
Large buffers → Ability to use higher-level selection codes.
 - ▶ More flexible: Can retain non-tagged data if bandwidth is available. Or write only high-level data for non-tagged.
- ▶ But if the rate is too high, need suppression on frontend.

Cross over point between Streaming vs. Hybrid, ext vers.

- ▶ In a streaming readout, the data reduction from a trigger would be achieved by tagging time segments for storage – others are dropped.
 - ▶ No requirement for fixed latency, just bound latency.
Large buffers → Ability to use higher-level selection codes.
 - ▶ More flexible: Can retain non-tagged data if bandwidth is available. Or write only high-level data for non-tagged.
- ▶ But if the rate is too high, need suppression on frontend.
- ▶ Typical PC today can buffer > 1 second of the physics rate.
- ▶ 100 times higher rates (or 5 times worst case) is about \$5k/second today.