EIC Yellow Report Summary Report

A. Celentano

VI Streaming readout workshop

The EIC Yellow Report process

- The **CERN Yellow Reports** series provides a **medium for communicating CERN-related work** where publication in a journal is not appropriate. Reports include material having a large impact on the future of CERN, as well as reports on new activities which do not yet have a natural platform. The series includes **reports on detectors and technical papers, criteria being that the audience should be large and the duration of interest long**. The term Yellow Reports is now used frequently for documents with similar purpose in various physics communities unrelated to CERN.

The EIC GOAL:

- **Advance the state of documented (i) physics studies** (White Paper, INT program proceedings) **and (ii) detector concepts** (Detector and R&D Handbook) in preparation for the EIC.

This will provide both the basis for further development of concepts for experimental equipment best suited to the EIC science needs, including complementarity of the two detectors/interaction regions, and input towards future Technical Design Reports (TDRs) of the experimental equipment.

EIC Yellow Report strategy

Quantify physics measurements for existing or new physics topics and implications for detector design ("Physics WG")

- Go beyond physics motivation to implication for detector requirements.
- Physics considerations for two independent complementary detectors.

Study detector concepts based on the requirements defined above, and quantify implications for the physics measurements ("Detector WG")

- Balance detector concepts versus impact on physics measurements.
- Document complementarity (+ reduction of systematics) of detectors.
- Fold in ancillary detectors, measurements (polarimetry, luminosity, ...).
- Engage EIC-detector R&D consortia.

EIC Yellow Report Timeline

Timeline as discussed in Early 2020 – there is an ongoing effort to speed-up the activities and have a first draft ready by the end of this year.

January 2020	Software tutorials are given, all activities are underway
March 19-21	First workshop at Temple University – Philadelphia *Goal: present progress for various groups and sub-groups, with much discussion and work time, initiate detector complementarity study based on detector technologies*
May 22-24	Second workshop at U of Pavia – Pavia, Italy *Goal: present initial physics measurements and detector requirements following five chosen processes/tools (inclusive measurements, semi-inclusive measurements, jets and heavy quarks, exclusive measurements, diffractive measurements & tagging), present detector concepts and implications for physics measurements. Complete detector requirements table including segmentation needs.*
August 3-7	Status reports at EICUGM @ FIU – Miami, FL *Goal: Conveners/sub-conveners inform community about status and progress. Conveners identify possible issues (if any) in meeting with EICUG Steering Committee.*
September 17-19	Third workshop at CUA – Washington, DC *Goal: present mature studies of detector requirements from physics processes, balance detector concepts versus impact on physics measurements. Discuss possible systematics reduction among complementary detector choices. Complete final "to-do" list for YR(s).*
November 19-21	Fourth workshop at UCB/LBL – Berkeley, CA or Final Meeting (assembly of Yellow Report(s)) *Goal: distribute draft YR sections before meeting*
January 2021	(optional) Final Meeting

EIC Yellow Report Timeline

Timeline as discussed in Early 2020 – there is an ongoing effort to speed-up the activities and have a first draft ready by the end of this year.

January 2020	Software tutorials are given, all activities are underway
March 19-21	First workshop at Temple University – Philadelphia *Goal: present progress for various groups and sub-groups, with much discussion and work time, initiate detector complementarity study based on detector technologies*
May 22-24	Second workshop at U of Pavia – Pavia, Italy *Goal: present initial physics measurements and detector requirements following five chosen processes/tools (inclusive measurements, semi-inclusive measurements, jets and heavy quarks, exclusive measurements, diffractive measurements & tagging), present detector concepts and implications for physics measurements. Complete detector requirements table including segmentation needs.* WE ARE HERE!
August 3-7	Status reports at EICUGM @ FIU – Miami, FL *Goal: Conveners/sub-conveners inform community about status and progress. Conveners identify possible issues (if any) in meeting with EICUG Steering Committee.*
September 17-19	Third workshop at CUA – Washington, DC *Goal: present mature studies of detector requirements from physics processes, balance detector concepts versus impact on physics measurements. Discuss possible systematics reduction among complementary detector choices. Complete final "to-do" list for YR(s).*
November 19-21	Fourth workshop at UCB/LBL – Berkeley, CA or Final Meeting (assembly of Yellow Report(s)) *Goal: distribute draft YR sections before meeting*
January 2021	(optional) Final Meeting

EIC Yellow report – detector working group

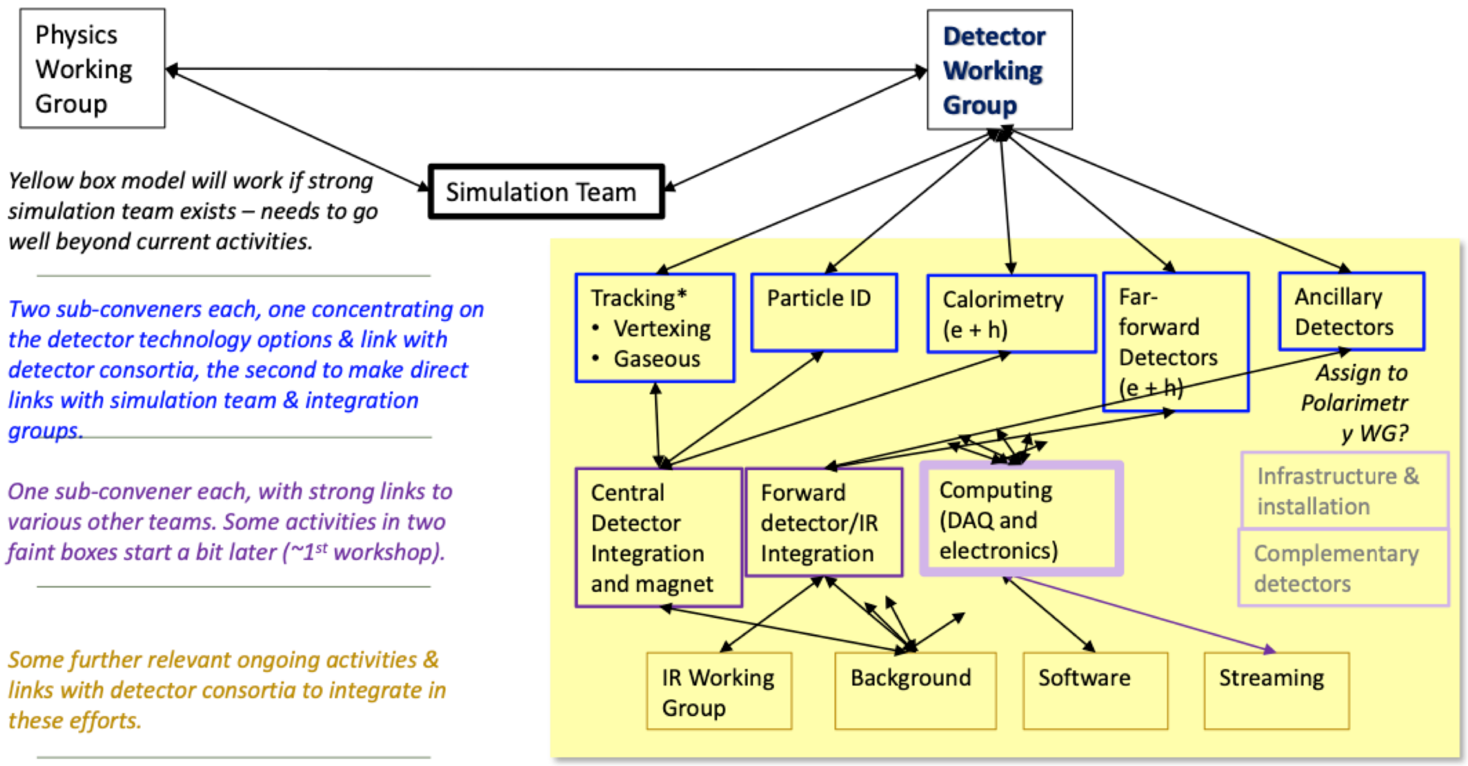
The detector working group is organized in different sub-groups, with a focus on a specific detector element.

Each sub-group has one or more sub-conveners who organize the activity and act as a contact point with the WG conveners.

A strong interaction among sub-working groups is foreseen and encouraged.

The readout and DAQ sub-working group has a connection with (almost) all the others packages!

The working group organization chart



*One additional sub-convener (to cover each of these distinct and evolving detector technologies)

EIC Yellow report – detector working group organization

Who we are

Working group conveners	K. Barish, T. Horn, P. Jones, S. Dalla Torre, M. Diefenthaler

System	Sub-conveners
Tracking	K. Gnanvo, L. Greiner, A. Mastroserio
PID	P. Rossi, T. Hemmick
Calorimetry	E. Chudakov , V. Berdnikov
Far-forward detectors / IR integration / ancillary detectors	Y. Furletova, M. Murray, A. Jentsch, D. Gaskell, E. Aschenauer
Central detector integration and magnet	A. Kiselev, W. Brooks
DAQ and Electronics	D. Neyret, A. Celentano
Detector complementarity	E. Aschenauer

Readout and DAQ working group

Motivations and methods

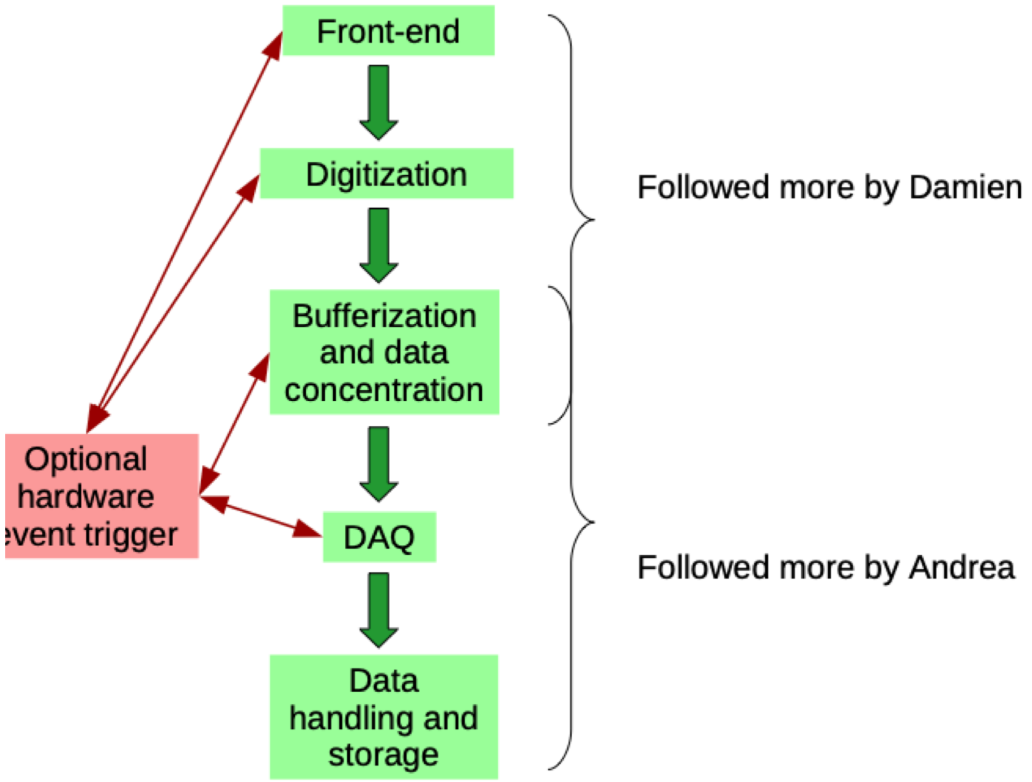
- *Bring peoples from different laboratories together to imagine realistic scenarios for the readout electronics and DAQ system of the future EIC experiments*

Strong links to build with:

- Detector WG → What detector we will have to read? Expected signal flux ?
- Physics and simulation WG → What signal to read ? What background to reject ? Which rate for each?

Keep a strong link with EIC R&D groups, in particular eR&D21, *streaming readout consortium*

TOPICS:



Readout and DAQ working group

Motivations and methods

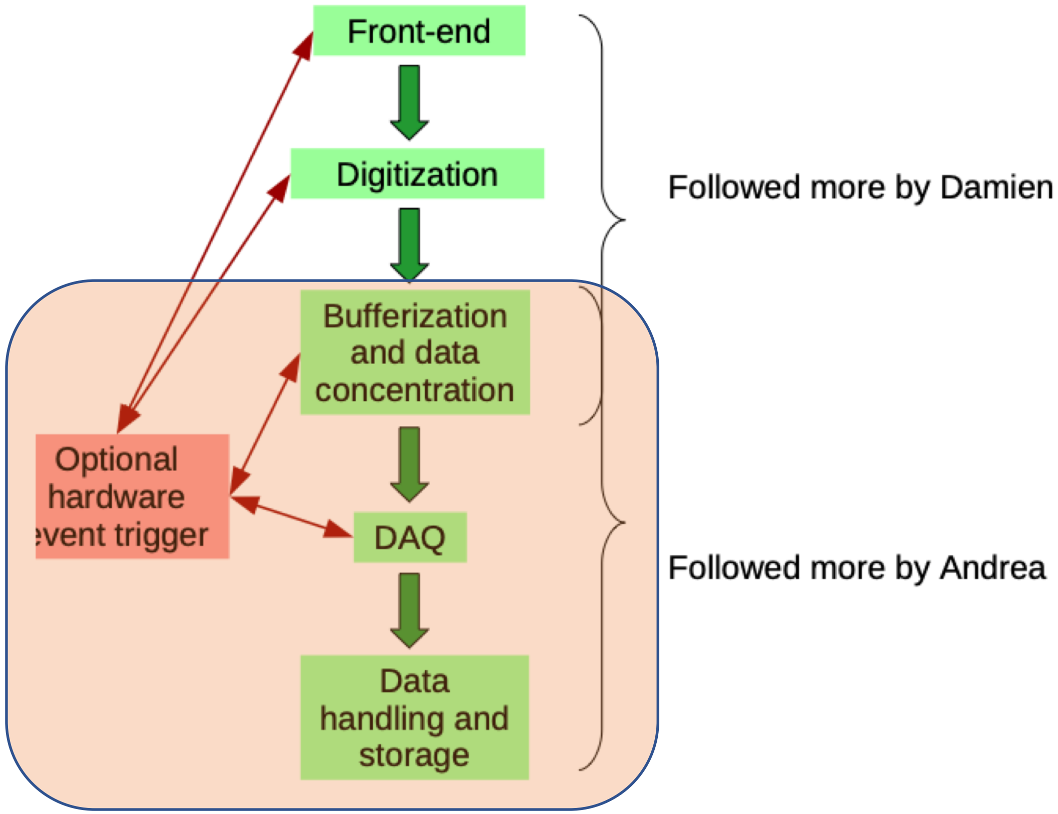
- *Bring peoples from different laboratories together to imagine realistic scenarios for the readout electronics and DAQ system of the future EIC experiments*

Strong links to build with:

- Detector WG → What detector we will have to read? Expected signal flux ?
- Physics and simulation WG → What signal to read ? What background to reject ? Which rate for each?

Keep a strong link with EIC R&D groups, in particular eR&D21, *streaming readout consortium*

TOPICS:



Main topics of this talk

Front-end

Main discussion topics

What do we need to readout? The answer strongly depends of kind of detector to read!

- Amplitude and shape of the signals
- Detector capacitance
- Number of channels
- Measurement to be done: amplitude, timing, position (barycenter of channels)
- What resolution for each kind of measurement ? What peaking time ?
- What context: particle fluxes, electronics occupancy, electronics noise level

First steps to reach:

- Overview on kind of detectors to read: contacts with detector WG
- Present state of the art on read-out electronics: existing chips for each kind of detectors, foreseen evolutions
- Projects on future read-out chips: do we need to?

DAQ

Main discussion topics

Triggered or triggerless DAQ?

- Identify pro/cons for each strategy - focus on general arguments rather than technical/economical details. Possibly provide experimental data (or at least simulations) to support these arguments.
  - Technical/economical reasons are important, but all numbers will be very different when the EIC will be built.
- Event definition and construction?
  - Are we ok with the paradigm “1 trigger == 1 event” in the triggered case?
  - Do we save “events” in the triggerless case, or just time-stamped hits / reconstructed quantities?
  - Effect on the physics to be measured / strategies to validate the trigger?

DAQ system

- Do we need a new DAQ infrastructure or can we adapt/reuse existing systems? *Depends on answer to questions before.*
- Complexity of trigger/filter decision? → How many trigger/filter levels?
- How to implement online software filtering (relevant for triggerless system and for a L3 software trigger level)?
  - Interconnection with the reconstruction software is critical
  - How to simulate the trigger / filtering?

EIC DAQ
Ongoing discussion

The current focus of the working group, concerning the DAQ system, is to conclude if (i) a software event triggering and building is necessary and (ii) if we should save raw data or reconstructed data.

Conveners' mail after Temple meeting:

Dear colleagues,
at our last meeting (see: https://indico.bnl.gov/event/8222/) we had contributions from Markus Diefenthaler, who introduced the EIC software requirements and strategies, and from Rosen Matev, who presented the ongoing efforts at CERN for streaming-readout based systems. In particular, Markus presented some interesting thoughts regarding the points that we, as Readout and DAQ working group, should think about concerning the software side of the Trigger and DAQ system. You can find these in the last slides of Markus presentation.

- Is software event triggering and building necessary?
- Could we drop raw data and keep only the reconstructed part?
- Same question concerning the event selection?
- What sort of calibrations would be necessary?

In the following, we'll briefly review the status of the ongoing discussion.

Our goal is to come to a conclusion for this at the Pavia meeting, next week.

DAQ: where we are?

We started the discussion regarding the strategy for EIC detector readout, to (try to!) solve the questions reported in the previous slide.

Starting point: the detailed rate estimate provided by J. Huang for the sPHENIX-EIC detector model

	EIC	RHIC	LHC → HL-LHC
Collision species	$\vec{e}+\vec{p}, \vec{e}+A$	$\vec{p}+\vec{p}/A, A+A$	$p+p/A, A+A$
Top x-N C.M. energy	140 GeV	510 GeV	13 TeV
Bunch spacing	10 ns	107 ns	25 ns
Peak x-N luminosity	10^{34} cm^{-2} s^{-1}	10^{32} cm^{-2} s^{-1}	$10^{34} \to 10^{35}$ cm^{-2} s^{-1}
x-N cross section	50 μb	40 mb	80 mb
Top collision rate	500 kHz	10 MHz	1-6 GHz
$dN_{ch}/d\eta$ in p+p/e+p	0.1-Few	~3	~6
Charged particle rate	4M N_{ch}/s	60M N_{ch}/s	30G+ N_{ch}/s

- EIC luminosity is high, but collision cross section is small ($\propto \alpha_{EM}^2$) → low collision rate
- Lower collision rate and small event size → signal data rate is low
- But events are precious and have diverse topology. Background and systematic control is crucial

From Jin's talk at Temple meeting

DAQ: where we are?

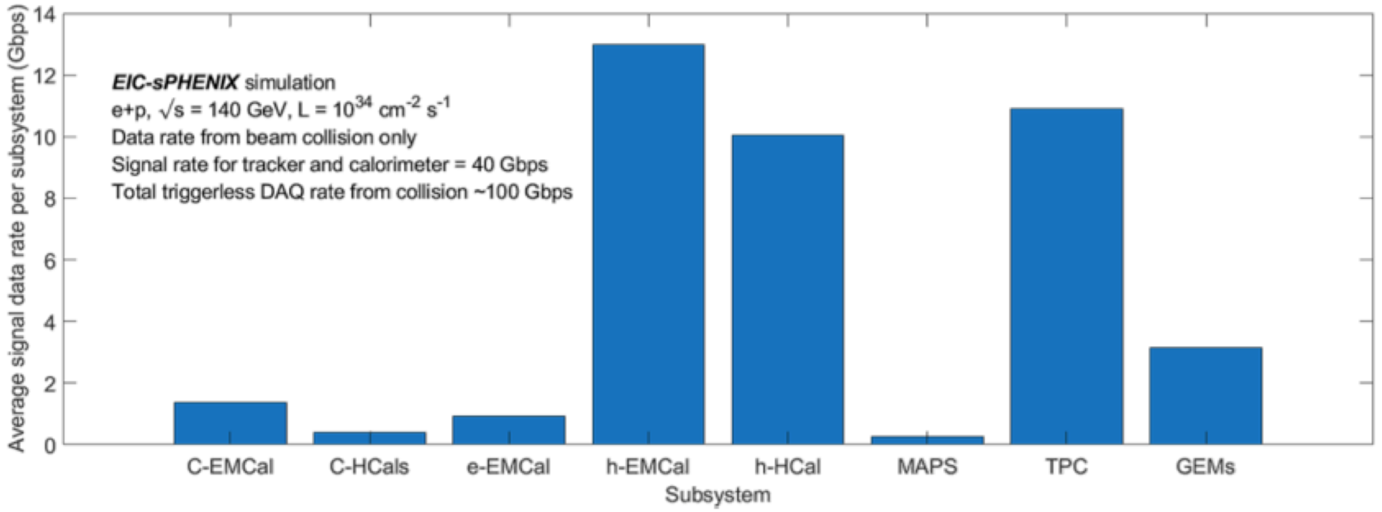
We started the discussion regarding the strategy for EIC detector readout, to (try to!) solve the questions reported in the previous slide.

Starting point: the detailed rate estimate provided by J. Huang for the sPHENIX-EIC detector model

From Jin's talk at Temple meeting

Signal data-rate estimate:

- Tracker + calorimeter ~ 40 Gbps
  - + PID detector + 2x for noise ~ 100 Gbps
  - Similar rate to an earlier estimation T. Ljubicic for BeAST (2015)
- Unlike LHC or RHIC, signal-collision data rate of 100 Gbps seems quite manageable:
  - Smaller than sPHENIX TPC peak disk rate of 200 Gbps
- Far-forward spectrometers/instruments not yet included



A more robust evaluation of the detector readout noise is currently in progress

DAQ: where we are?

We started the discussion regarding the strategy for EIC detector readout, to (try to!) solve the questions reported in the previous slide.

Starting point: the detailed rate estimate provided by J. Huang for the sPHENIX-EIC detector model

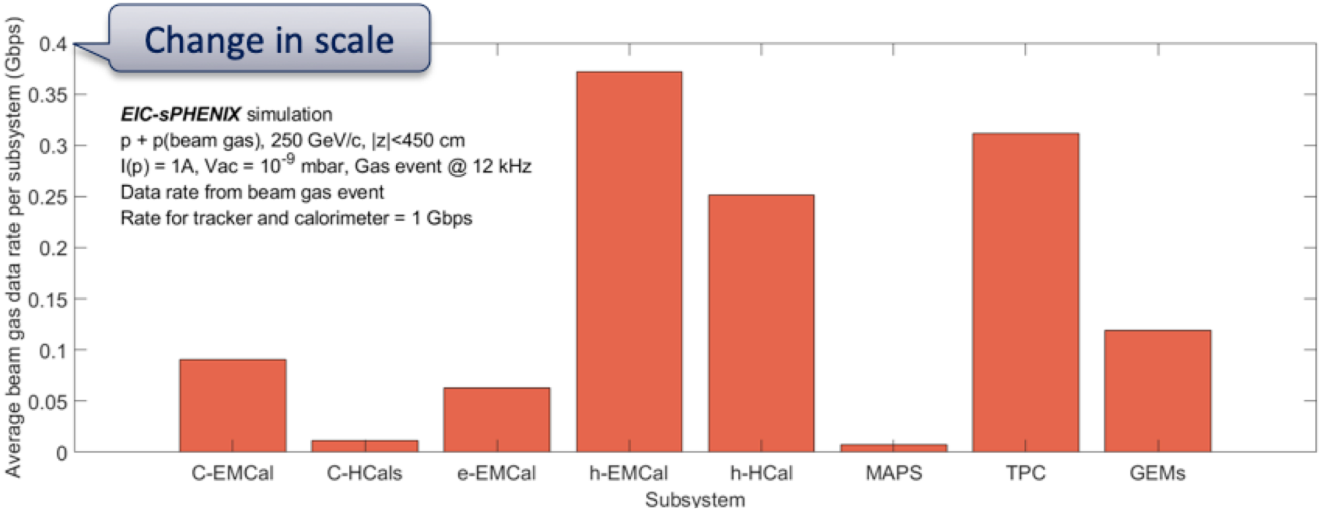
Bottom-line message: in principle, we can stream to disk directly all raw data from the full detector!

From Jin's talk at Temple meeting

Beam-gas interaction data rate estimate:

- Very similar rate distribution among subsystems when compared with EIC collisions
- With an assumed vacuum profile
  - Based on HERA experience, assuming 10^{-9} mbar flat within experiment region (|z|<450 cm)
  - Overall ~ 1 Gbps @ 12kHz beam gas << EIC collision signal data rate

  Thanks to the discussions with E. Aschenauer, A. Kiselev, and C. Hyde
- Further investigation needed:
  - In the experimental region : Dynamic vac profile
  - Beyond experiment region: beam gas profile, possible passive shielding and active veto



- What we want to record: total collision signal ~ 100 Gbps @ 10^{34} cm^{-2} s^{-1} less than sPHENIX peak disk rate
- For YR studies over the coming ½ year
  - Need to include realistic PID simulation and far forward instrucmentation
  - Need to include as much as other source of background and noises (e.g. synchrotron)

DAQ

Ongoing discussion

Markus's slides

Discussion: Is software event triggering and building necessary?

Comments from ongoing discussion with Sergey Furletov (JLAB), David Lawrence (JLAB), Chris Pinkenburg (BNL), Maurizio Ungaro (JLAB), and Torre Wenaus (BNL):

Is software event triggering and building necessary

- We need to address noise. E-p and likely e-A is all about noise with their low multiplicity the event itself may not contribute too much to the actual event size. In Streaming Readout triggering will only reduce the amount of data which gets archived. Unlike for a lvl1 trigger where you trade one event for another you will not increase the number of interesting events.
- If we agree on analysis-ready data from the DAQ system being a goal, then we need to build events in (near) real time and also filter them according to physics interests.

Discussion: Drop raw data and keep only the reconstructed part?

Comments from ongoing discussion with Sergey Furletov (JLAB), David Lawrence (JLAB), Chris Pinkenburg (BNL), Maurizio Ungaro (JLAB), and Torre Wenaus (BNL):

At which point can we reconstruct data online, and drop the raw data to keep only the reconstructed part?

- Why would we do that? We can store the raw data for the EIC.
- Does this make sense? Often raw data is more compact then the resulting reconstructed output. Take ALICE approach to buffer data for a long time in the O2 farm for calibration and reconstruction and then only write reconstructed data to tape? Probably not allowed by DOE regulations which I think demand saving the data permanently as soon as possible.
- DUNE does this: Their event size is O(8GB) and they reduce it to O(80MB).
- The second part of the question is when we can reconstruct online. I think we can do this (near) real time if we can do calibrations online.

Discussion: Same question concerning the event selection?

Comments from ongoing discussion with Sergey Furletov (JLAB), David Lawrence (JLAB), Chris Pinkenburg (BNL), Maurizio Ungaro (JLAB), and Torre Wenaus (BNL):

Same question concerning the event selection?

- I would give the same answer as above. If we can reconstruct in near real time, then we can also build events and select on them.

Markus's point of view: a unique machine-detector-analysis framework, with strong interconnection among the different elements.

Streaming readout and online analysis highly preferable.

David’s slides

DAQ

Ongoing discussion

Jefferson Lab

My Opinions

1. **Is software event triggering and building necessary?**
   If Jin’s numbers hold then, no.
   Should we do it? yes
   There is no technical reason for us not to and it will
   a. reduce the time it takes to start real physics analysis
   b. position us for potential detector upgrades without redesigning DAQ

2. **Could we drop raw data and keep only the reconstructed part?**
   Yes, but if the cost of keeping both is relatively small then don’t. It mitigates risk considerably if you can keep the raw data as well. It does not necessarily need to be kept online and easily available (e.g. tape vault)

3. **Same question concerning the event selection?**
   Same answer.

4. **What sort of calibrations would be necessary?**
   It is conceivable to record raw data but only for events triggered based rough calibrations. Details on detector design and event topologies would be needed to answer this.

David’s point of view: online analysis and event reconstruction is not strictly necessary, but has advantages (e.g. potential detector upgrades without redesigning DAQ). There are difficulties related to online calibrations.

Online analysis not strictly necessary, but possible.

Rosen Matev's slides

DAQ

Ongoing discussion

Takeaway

- Going "triggerless" helps if you have the processing power and storage
- Align and calibrate your detector online
  - helps with improving efficiency and reducing background
- Squeeze the offline A&C and reconstruction online
  - you are sure to have the best physics objects for analysis
  - you can be much tighter on selections
- After that, it's "easy"
  - just throw away what is not necessary from the events
  - still, make sure you've convinced yourself first it's ok
  - still, make sure your QA/QC is solid as there is no going back

Chris' slides

Storing the data at EIC

- offline event building and then tape storage runs afoul of DOE rules
- There is a grey area when data is deemed to have been "stored", but once stored, that's the data we have to preserve
- We have successfully (and correctly) argued that the PHENIX PRDF compression scheme works on data before it hits any disk storage, so that was ok
- We could not have written the files, compressed them, deleted the originals - no go
- So whatever shows up as "the originals" needs to be stored. Period.

Rosen's point of view: online analysis and event reconstruction is really critical, in particular for a new detector.
This point of view was shared by Chris, who also reminded us about DOE regulations.

Online analysis is critical.

Future steps for the studies on front-end electronics

Collaboration with detector WGs to gather informations on the future detectors

- Kind of detector, characteristics of the signals, number of channels, expected rates due to physics and background, functions requested for the read-out electronics, environment (radiation, temperature, pressure, magnetic field,...)

Summarize state of the art of read-out electronics

- Existing chips and foreseen evolutions, project of future read-out chips → 1 page for each chip or chip family, a link to a common document will be provided
- Same for detectors with integrated electronics

Contacts with physics WGs to get information about foreseen experimental conditions

- Physics and background rates, particle multiplicities, event sizes,...


APV25-S1

Chip functionalities

→ analog preamplification
→ analog bufferisation
→ amplification/shaping

General characteristics

Front-end chip with analog pipeline, for Silicon and gaseous detector, 128 channels, 8x7.1mm², bare chip to be bonded, 0.25µm CMOS, ~300mW, ~25€/chip, obsolete, no evolution foreseen (AFAIK)

Analog amplification, bufferization characteristics

CR-RC shaping amplifier, noise 246e⁻ + 36e⁻/pF, peaking time 100-250ns, gain range +/- 20% around 100mV/mip (Si detector), sampling rate ~40MHz, 192 analog cells per channel, analog signal dynamics ?, both polarities (integrated inverter), sample selection but no channel selection, 1 analog multiplexed output channel which send the 128 channels, 20 and 40MHz output sample rate

Self-triggering characteristics

none

Signal discrimination and digitization characteristics

none

Bufferization and data treatment characteristics

none

https://docs.google.com/document/d/1qm_0vTbEdkj9GDMo77Yy7XdmE5Xuc5fP_gIgSFhRQGM/edit?usp=sharing

Future steps for the studies on DAQ systems

Produce a glossary of terms to make sure we have a common ground in the discussion (streaming DAQ vs triggerless DAQ vs continuous readout)

- The document was prepared starting from the input of EIC eR&D 21 (streaming readout) consortium
- Live document, you are invited to contribute!

Glossary

Front-end electronics (FEE): The electronics which interfaces with the detector, typically amplifiying, shaping and converting the analog signal from the detector via an analog-to-digital (ADC), charge-to-digital (QDC), or time-to-digital (TDC) converter into the digital domain. FEE typically includes data bufferization and logics for data transfer to the downstream element in the read-out & DAQ chain. Zero and noise suppression is also performed in FEE in some of the FEE chips.

Triggered readout: A data acquisition system in which hardware produces an electrical signal according to a trigger criterion based on a subset of detector information available quickly. The signal is used to control the conversion of detector signals into the digital domain, or to trigger the read-out of a data-window from a continuously filled buffer.

(Andrea): **Triggered readout:** A data acquisition system in which some data from a subset of detectors ("trigger data") is sent to a dedicated subsystem to produce a trigger decision. The trigger decision is based on a partial elaboration of the "trigger data". "Trigger primitives" are reconstructed and analyzed to assess whenever all the data from the detector has to be stored for later analysis. In this case, a proper signal is sent back to all the readout elements to control the conversion of detector signals into the digital domain, or to trigger the read-out of a data-window from a continuously filled buffer. A key aspect of a triggered readout system is the fixed latency between the physical event time (FE -> Trigger system) and trigger time (Trigger system -> FE).

Second-level / high-level trigger: In triggered systems, higher-level triggers are often used to reduce deadtime (via a fast clear) or data amount (by dropping the so-far recorded data for that event). Each level in such a system typically has different time constraints and complexity limits. For example: a certain time frame could not be forwarded to the tracker if certain conditions are not met. In certain, complex, triggered setups, the later stages can resemble a streaming system, where a stream of events flows through a network of analysis nodes, and data selection criteria either accept or drop the event. The main remaining difference for this part is then that the data is organized and tagged by an event number instead of time stamps.

Pipelined/buffered readout: A triggered readout system where event data is stored on the front ends and read out asynchronously by the backend.

Damien N
19:42 17 mar
Aggiungi: *"amplifiying, shaping and"*

Andrea Celenta...
12:51 19 mar
Risolvi
I don't agree on this. The trigger may be obtained from a software system: some portions of the detector may stream data to a CPU farm, where a trigger decision is made, and then the trigger is sent back to all the other sub-detectors. This requires real-time operation of the software system.

Andrea Celenta...
13:03 19 mar
Risolvi
Trigger should here be filtering

https://docs.google.com/document/d/10pz1Z6gaQN4b06u7FxKmBEMlFgL-7BsL_PqnxyISgus/edit?usp=sharing

Future steps for the studies on DAQ systems

Produce a glossary of terms to make sure we have a common ground in the discussion (streaming DAQ vs triggerless DAQ vs continuous readout)

- The document was prepared starting from the input of EIC eR&D 21 (streaming readout) consortium
- Live document, you are invited to contribute!

Summarize DAQ proposals and ideas for EIC using a common template, to allow apples-to-apples comparison

System name: FELIX-based streaming DAQ

FELIX-based streaming DAQ filled in the template for each system.

Contact: Jin Huang <jhuang@bnl.gov>, Martin Purschke <purschke@bnl.gov>

General properties of the system

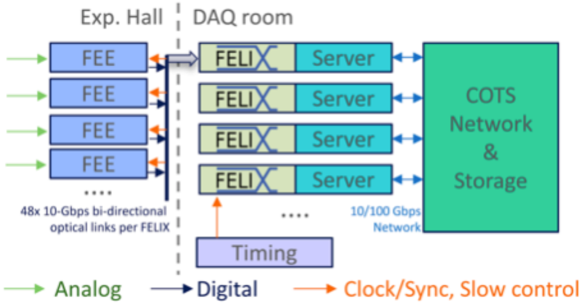


Figure A: FELIX-based EIC DAQ concept. All FEEs are synchronized to the beam interaction clock via the FELIX interface [DOI: 10.1109/TIM.2019.2947972]. The digitized data are streamed and buffered on the FELIX FPGA and server prior sent to network storage servers in the counting house.

- Triggered / streaming / mixed
  - Streaming by default
  - Trigger-throttled streaming in calibration mode and when the background rate is far higher than the collision data rate
- Number and type of filtering (streaming) / triggering (trigger) layers
  - 1. FEE: zero-suppression or self-triggering at the front-end FPGA or ASIC (e.g. SAMPA, FPHX, ALPIDE). Latency ~ O(1-10 μs)
  - 2. FELIX->Computing: Streaming time-frames (O(100 μs)-wide) of hit data to online buffer disks. Further noise filtering and compression if needed. Latency ~ O(1 s)