Streaming Readout for sPHENIX and EIC

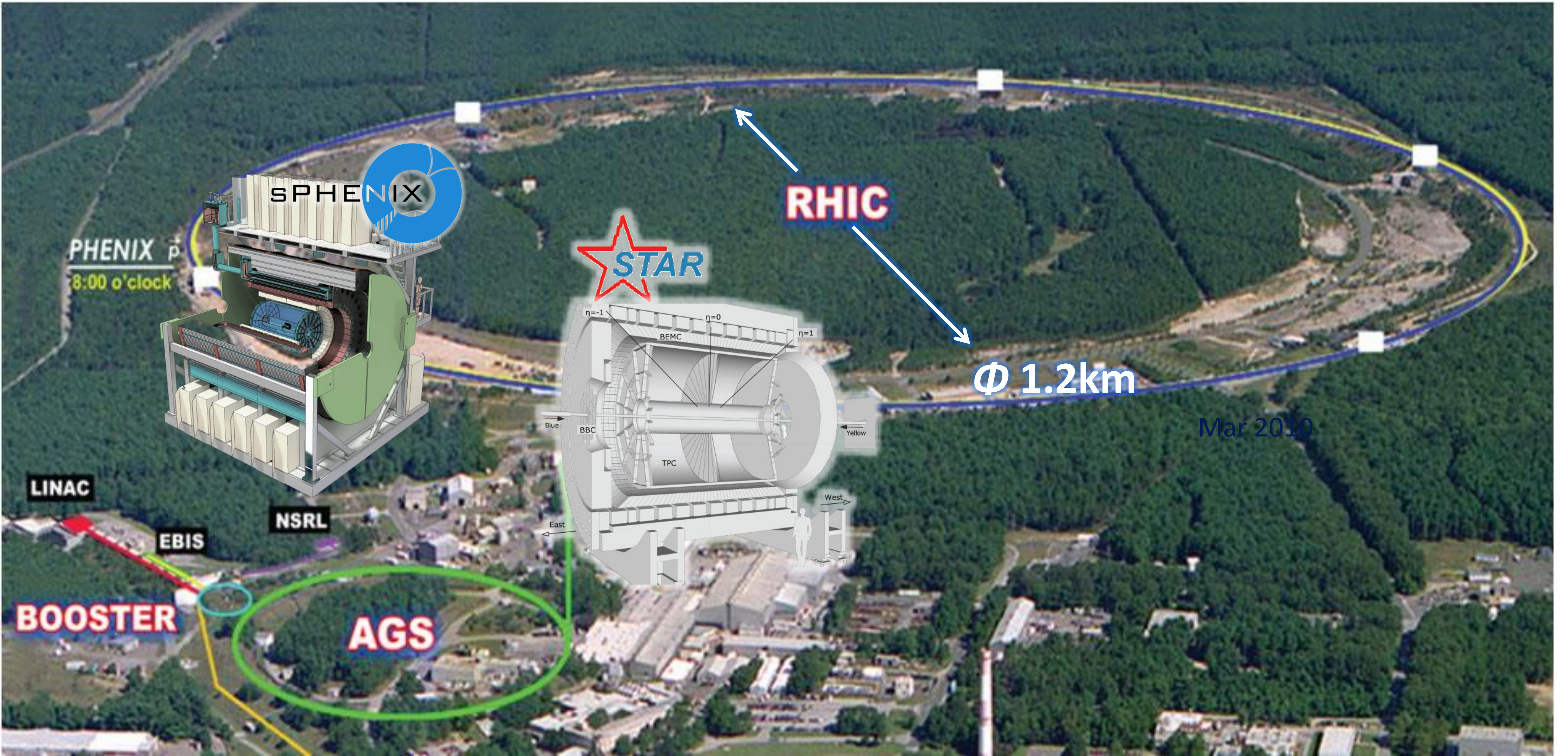
Jin Huang

Brookhaven National Lab

Thanks to the inputs from many colleagues!

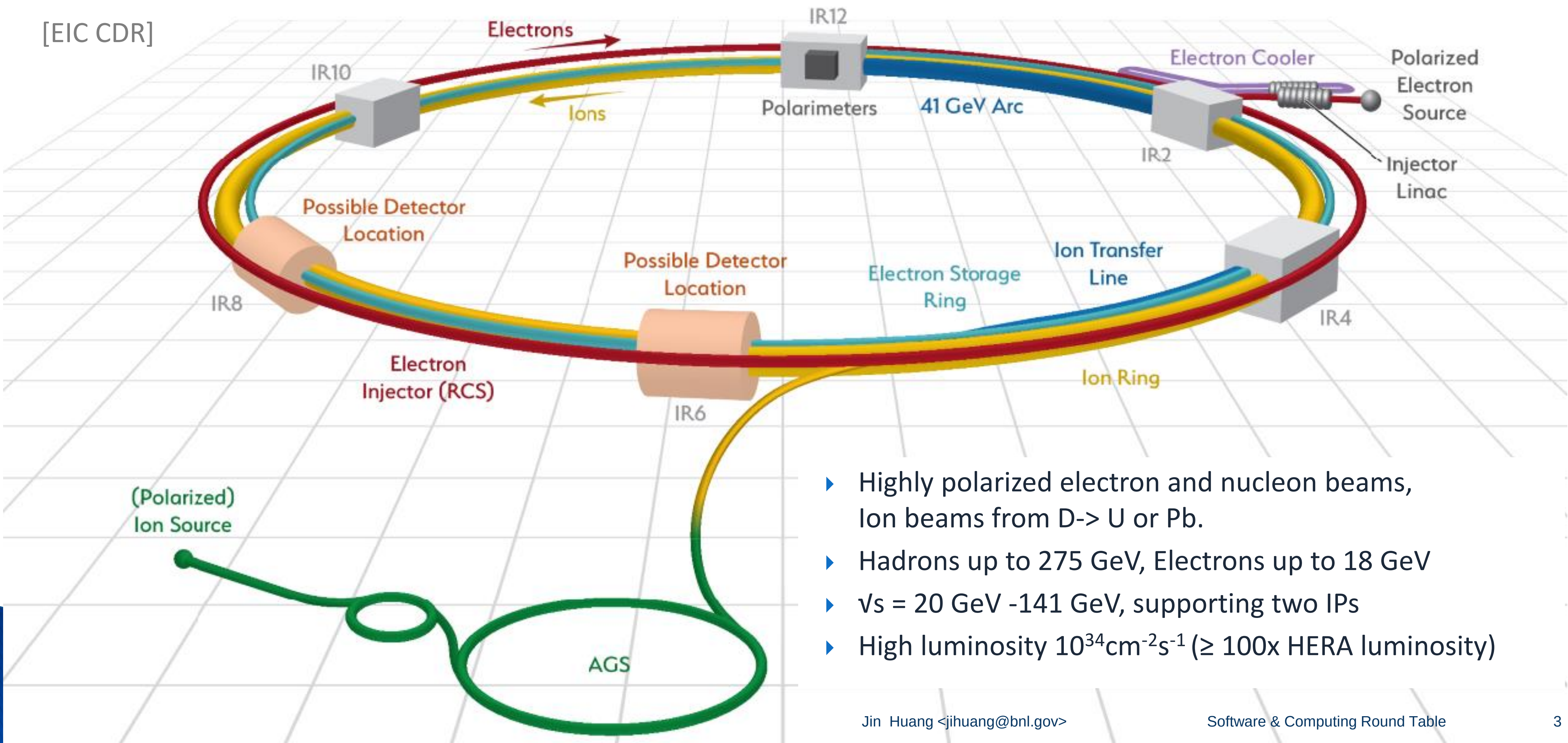
Also with reference to talks at Streaming Readout VIII and AI4EIC workshops

Relativistic Heavy Ion Collider



RHIC transition to the EIC

[EIC CDR]

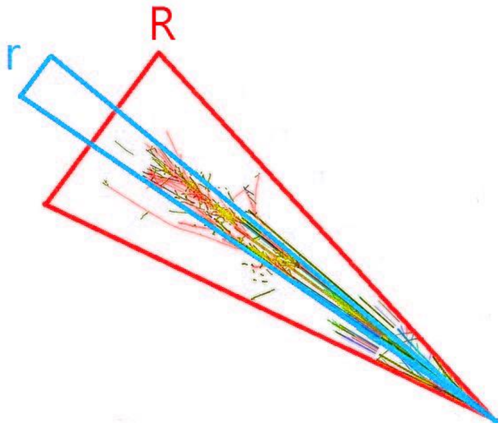


- Highly polarized electron and nucleon beams, Ion beams from D-> U or Pb.
- Hadrons up to 275 GeV, Electrons up to 18 GeV
- √s = 20 GeV -141 GeV, supporting two IPs
- High luminosity $10^{34}cm^{-2}s^{-1}$ (≥ 100x HERA luminosity)

Core physics programs

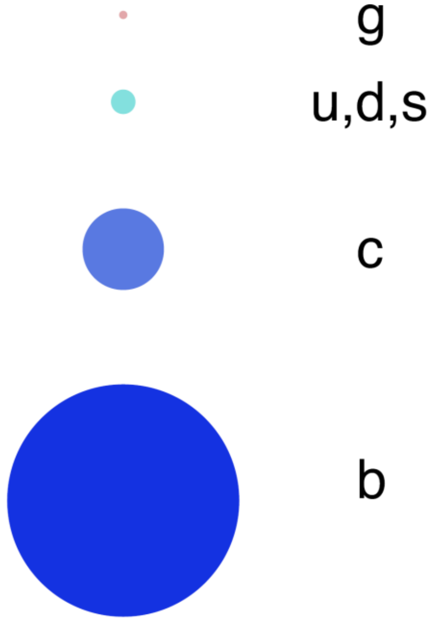
Jet cor. & substructure

Vary momentum/angular size of probe



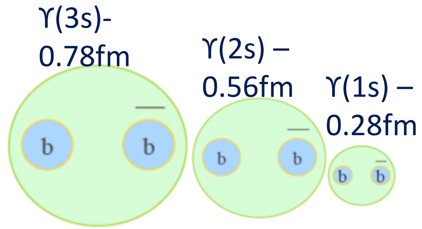
Parton energy loss

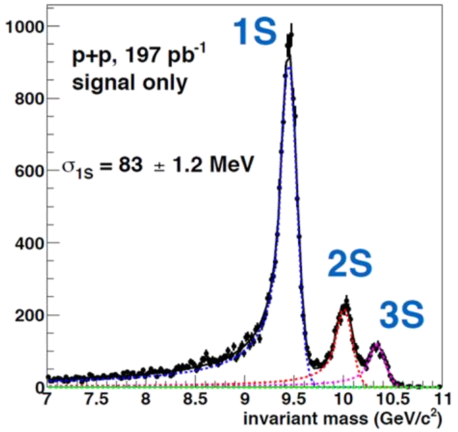
Vary mass/momentum of probe



Upsilon spectroscopy

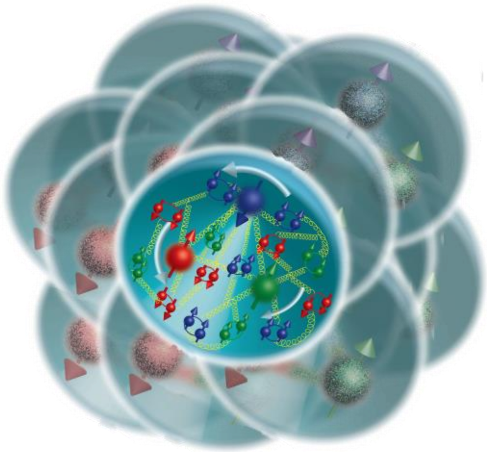
Vary size of the probe



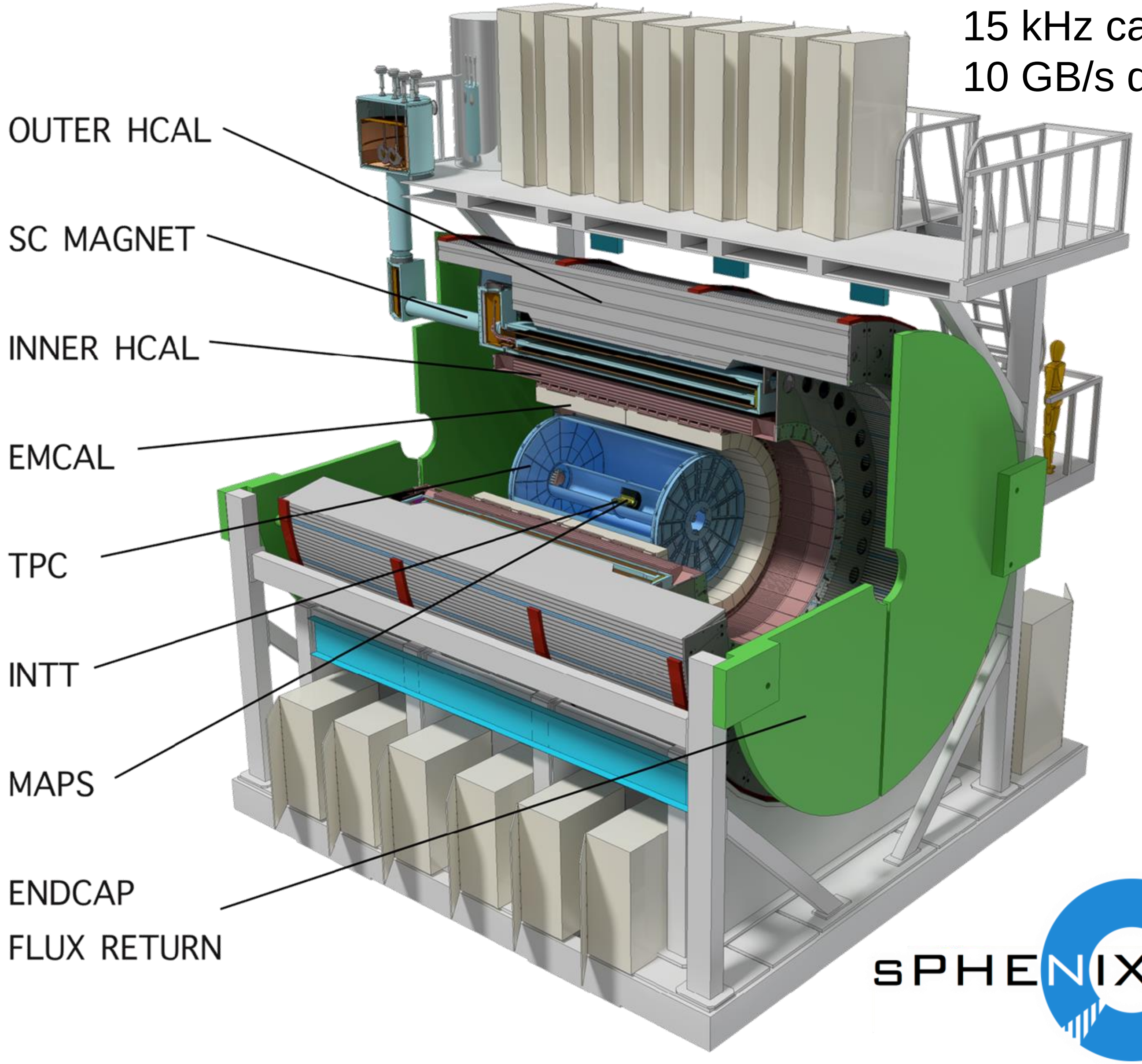


Cold QCD

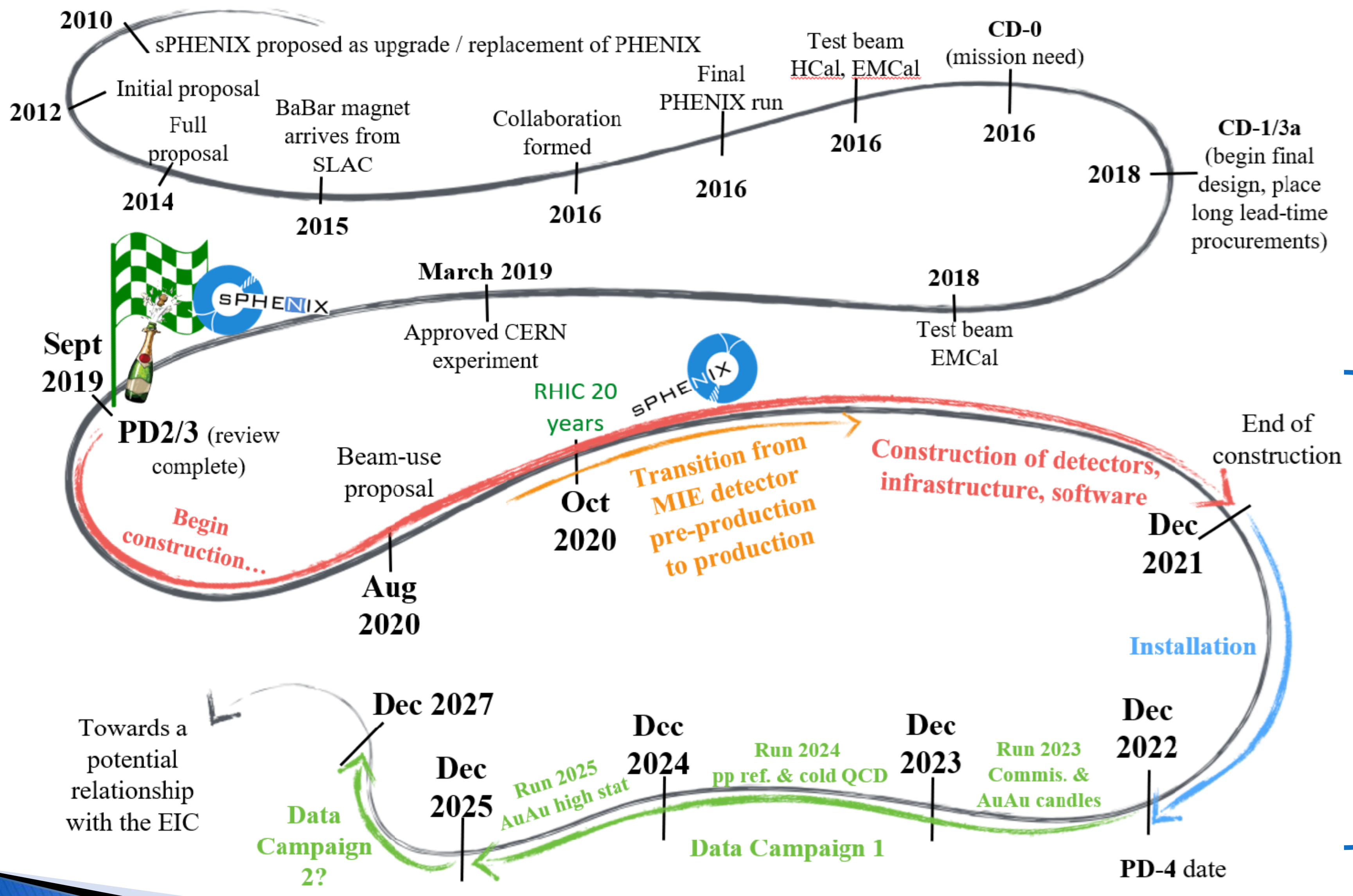
Vary temperature of QCD matter



15 kHz calo trigger + 10% streaming DAQ
10 GB/s data logging



sPHENIX Detector



~1 years !
From now to first data

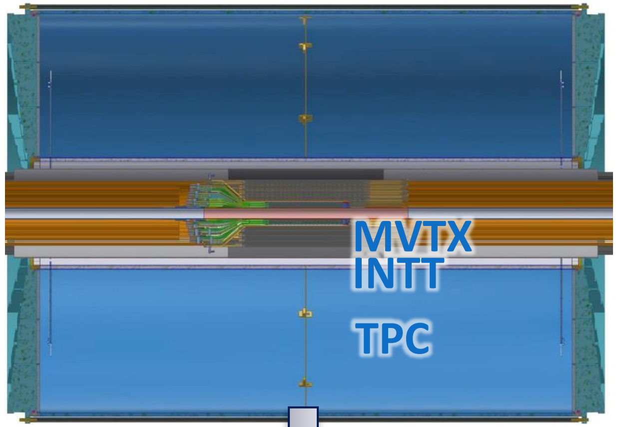

sPHENIX magnet installation
RHIC IP8 Hall, Oct 7, 2021
← 13/32 sPHENIX hadronic calorimeter sectors installed

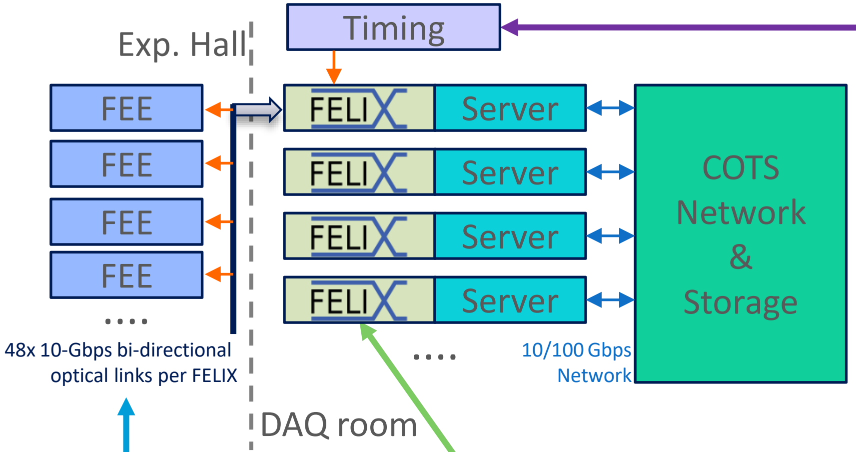
Related streaming readout electronics

Associated test projects

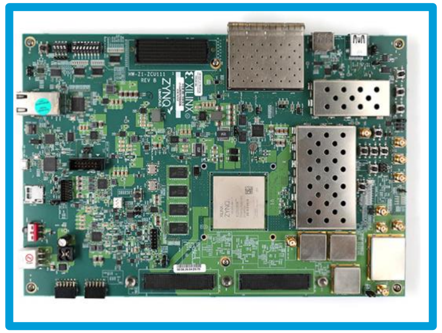
Precision timing digitizer
DRS4GIO (SBIR/LDRD)

sPHENIX streaming DAQ for tracker

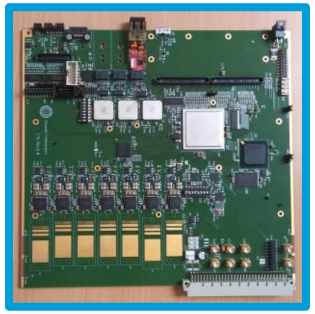




Global Timing Module (NSLS II/sPHENIX)
Receiving from RHIC RF low glitter clock source



High density multiplexer+ ADC
RFSoC Digitizer (LDRD)



MVTX RU, 200M ch

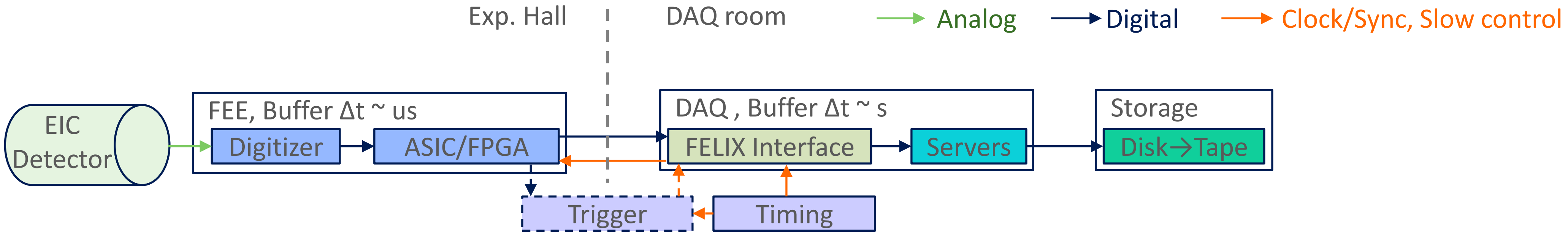
INTT ROC, 400k ch
ALPIDE (ALICE/sPHENIX), FPHX (PHENIX)

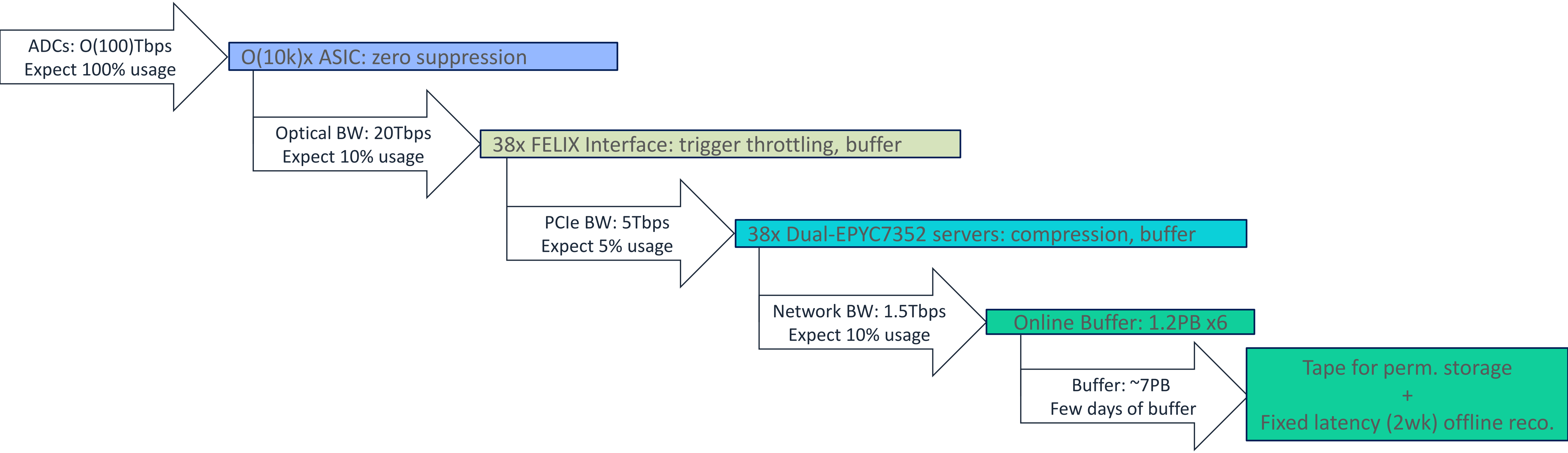
TPC FEE, 160k ch
SAMPAv5 (ALICE/sPHENIX)

BNL-712 / FELIX v2 x38 (ATLAS/sPHENIX)

FELIX Ref: 10.1109/tim.2019.2947972
Similar role as PICe40 in LHCb [R. Aaij's talk]

sPHENIX Streaming data flow





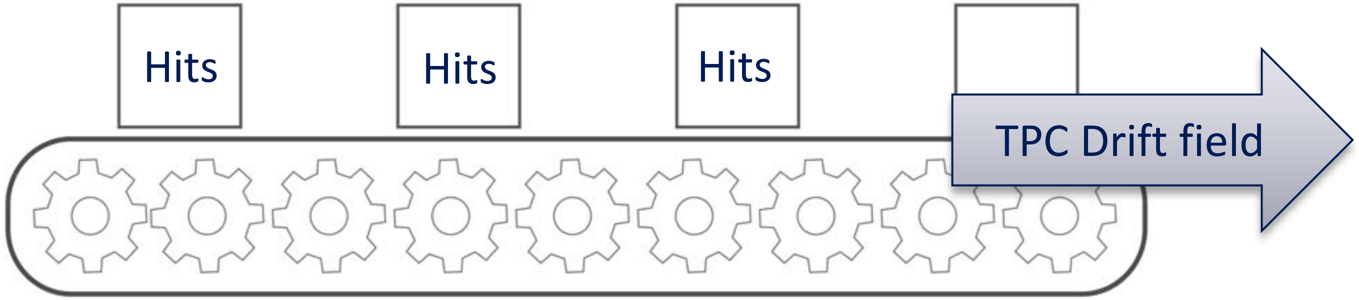
Streaming readout status at sPHENIX

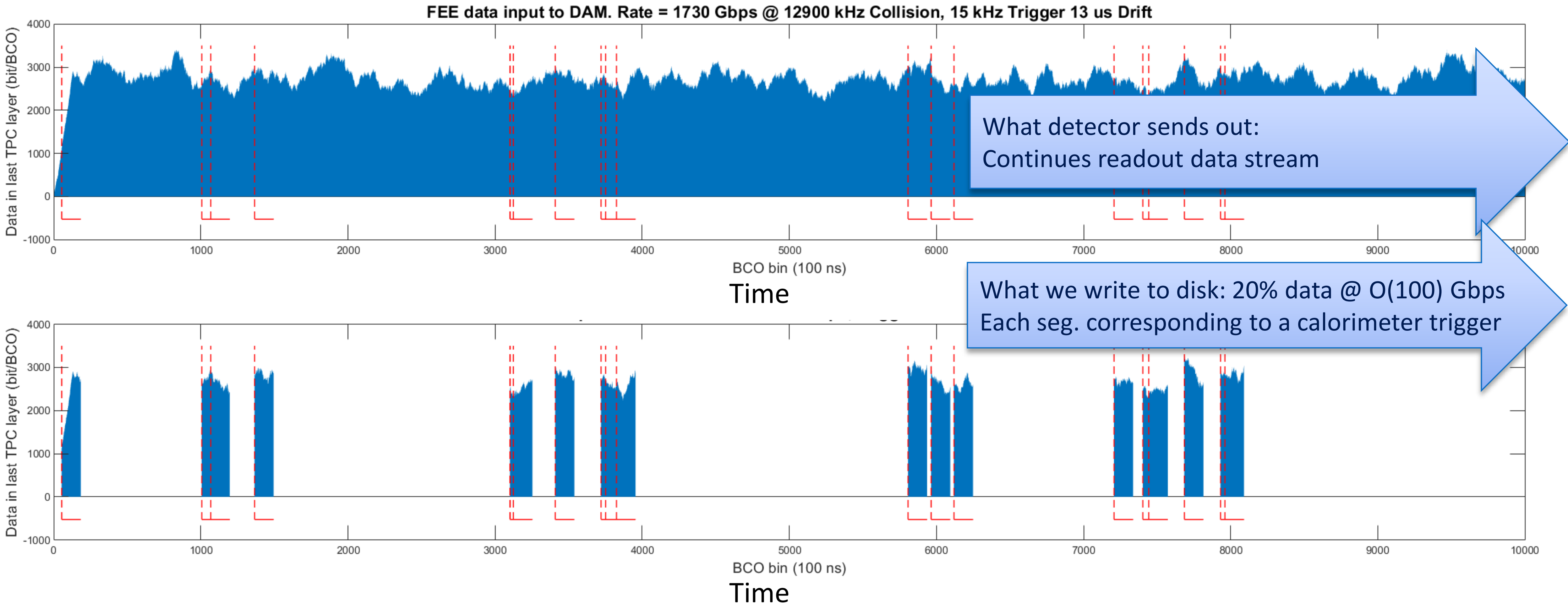
- All three sPHENIX tracking detector uses streaming readout
- Developed plan to take streaming data for heavy flavor physics program (next slides), commended by RHIC PAC.
- Completed construction of sPHENIX FELIX DAQ interface (~50) and procurement of DAQ servers, network infrastructure and online disk buffers
- Data taking start in 2023!

RHIC PAC 2020 report

> We commend sPHENIX for developing the continuous streaming readout option for the detector, which increases the amount of data that can be collected in Run-24 by orders of magnitude. In particular in the sector of open heavy flavor, this technique will give access to a set of qualitatively novel measurements that would otherwise not be accessible. Given the tight timeline for completing the RHIC physics program before construction of the EIC begins, this is a tremendous and highly welcome achievement.

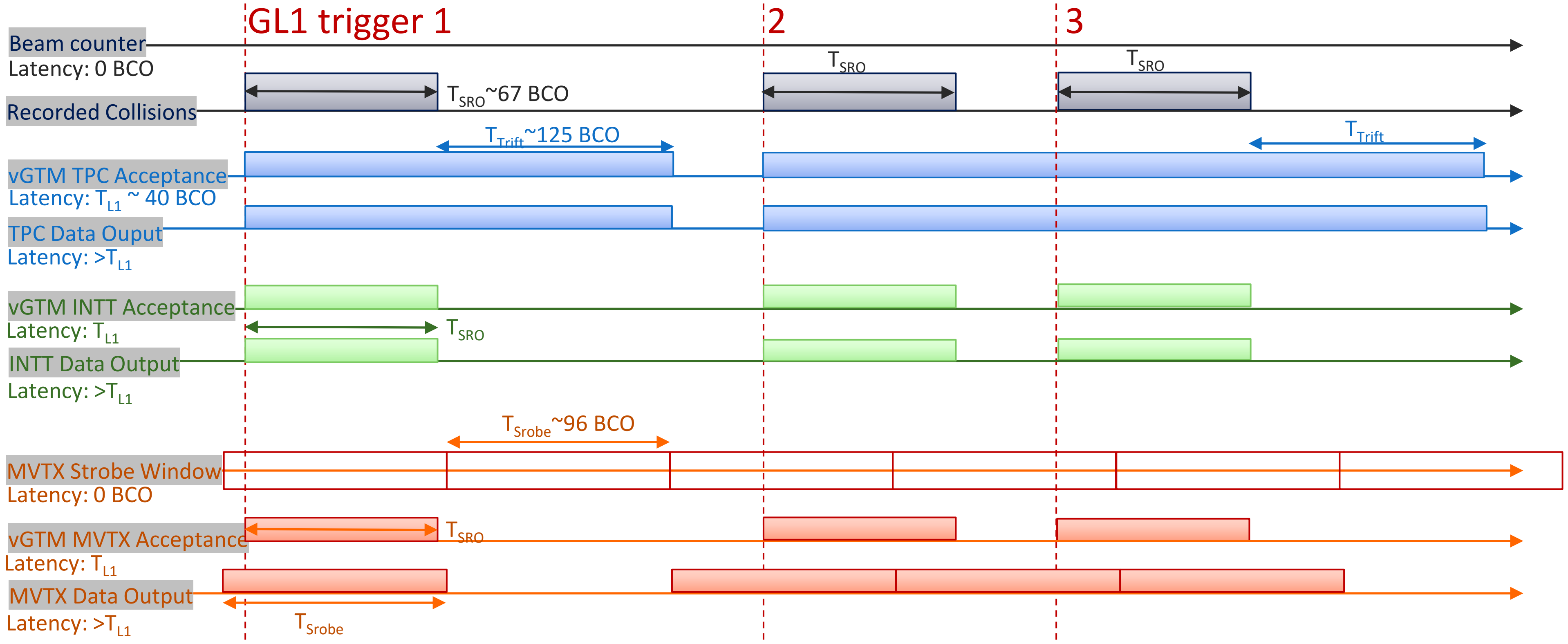
TPC data stream in sPHENIX triggered DAQ





SRO-Mode1-Simple [Recommended]

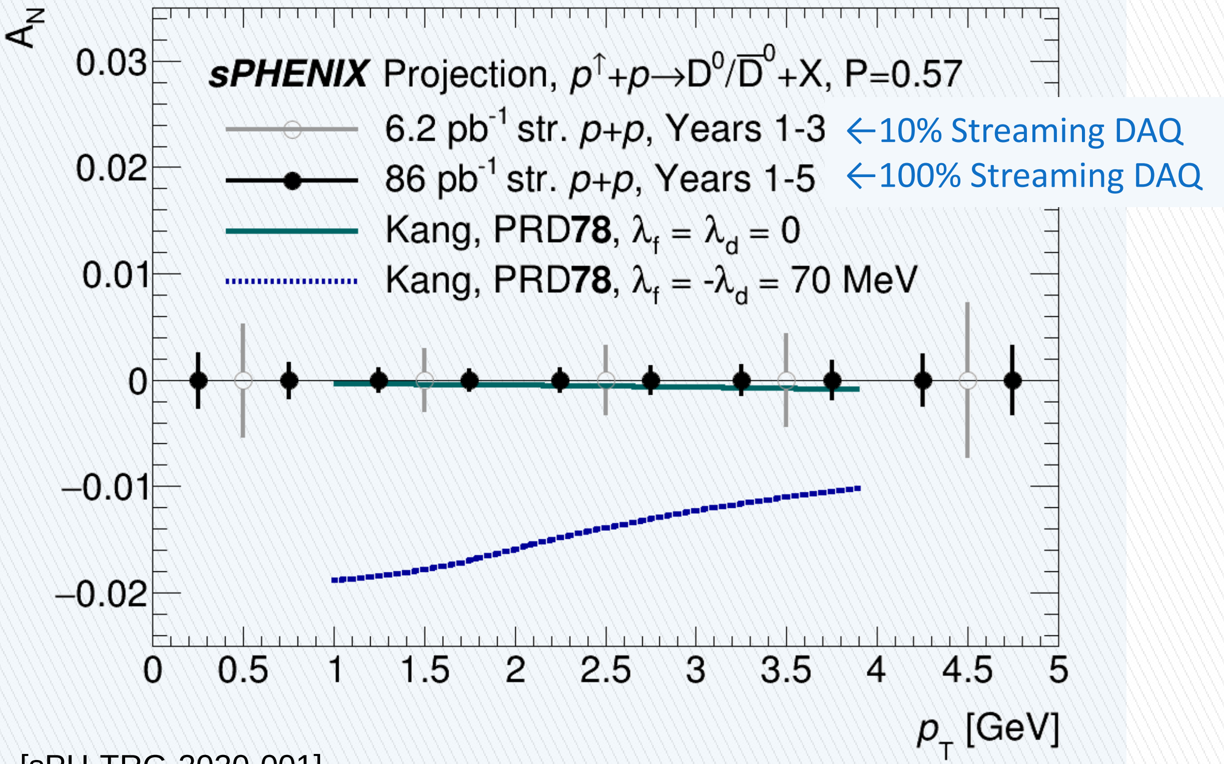
Simply prolong L1-Acceptance signal to each subsystem, from 1 BCO to T_{SRO}~67 beam crossings (~7us or 10% SRO data)
→ x500 increase of hard-to-trigger p+p sample
→ at cost only 50% increase in data vol. (by piggy back on long TPC readout window of 13us)



Streaming-DAQ enabled scientific connection: e.g. gluon dynamics via heavy flavor A_N

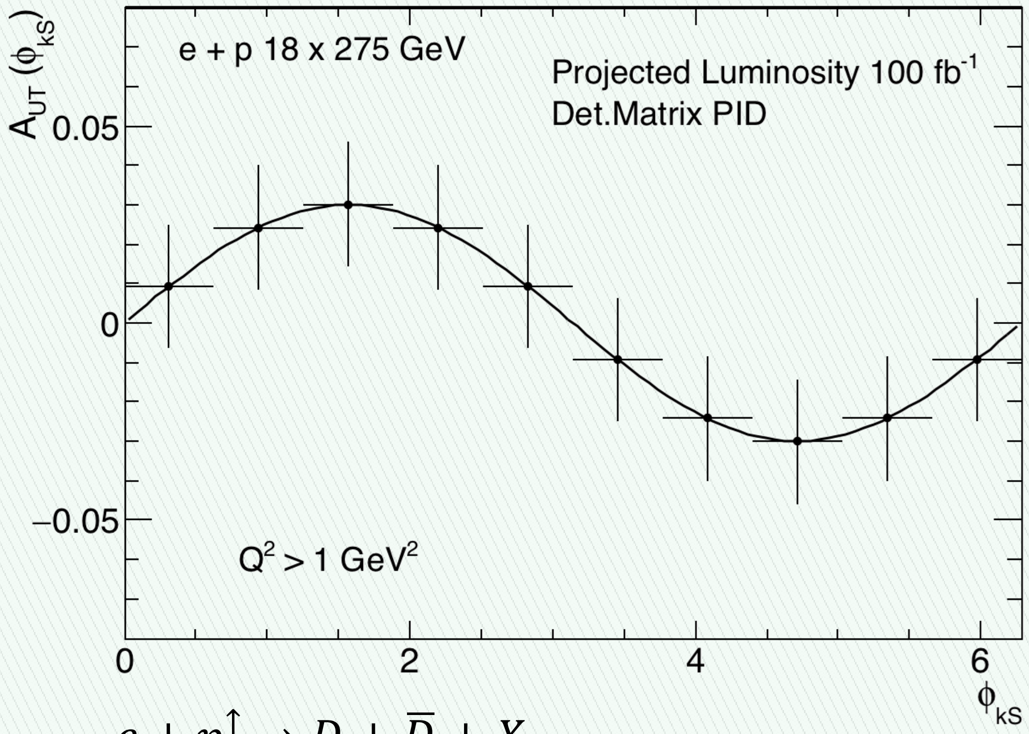
Universality test on gluon Sievers

sPHENIX D^0 trans. spin asymmetry, $A_N \rightarrow$ Gluon Sievers via tri-g cor.



[sPH-TRG-2020-001]
Model: 10.1103/PhysRevD.78.114013

EIC SIDIS D^0 transverse spin asymmetry → Gluon Sievers

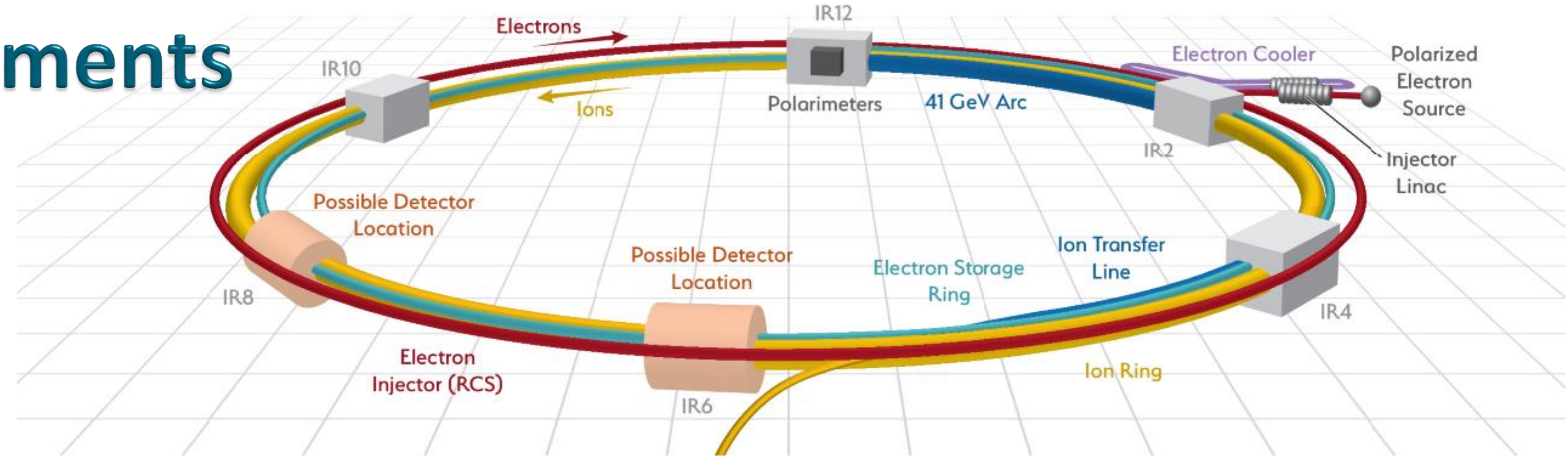


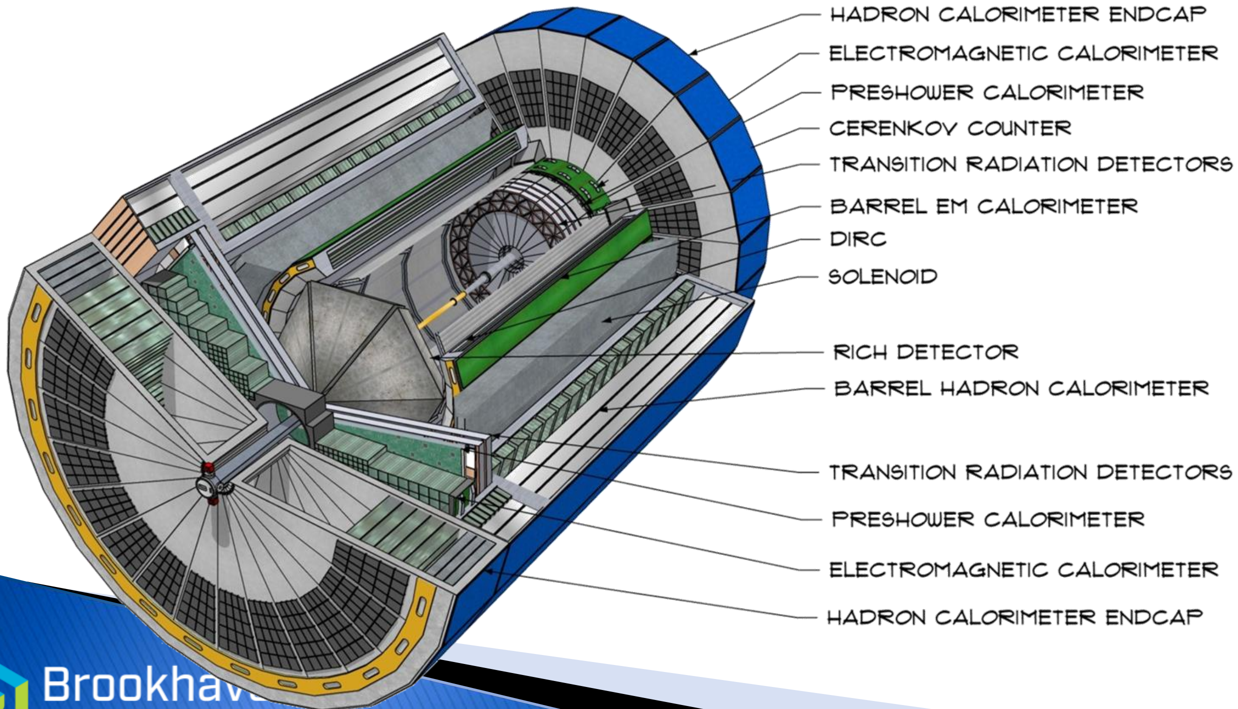
$$e + p^{\uparrow} \rightarrow D + \bar{D} + X$$

[CNFS HF@EIC workshop, Nov 4-6, 2020]

Brookhaven National Laboratory

EIC experiments





Generic EIC Detector model [EIC YR]

See also: proposal detectors in development

- ATHENA: athena-eic.org
- CORE: eic.jlab.org/core
- ECCE: ecce-eic.org

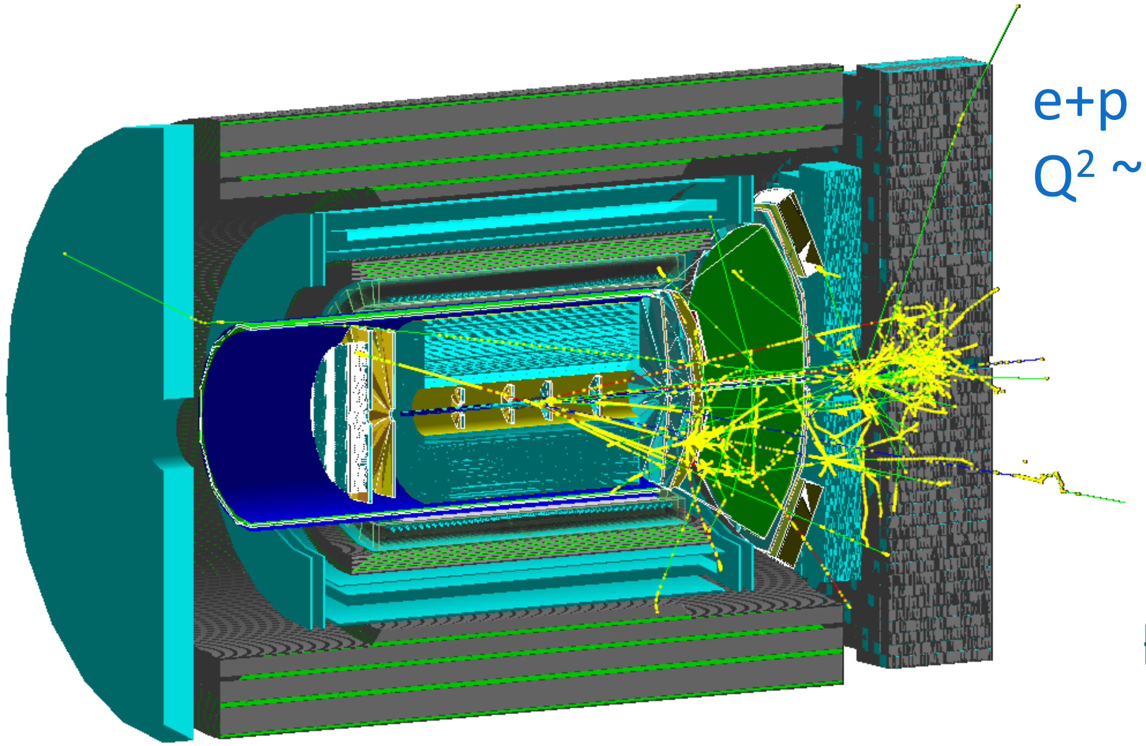
EIC: unique collider
→ unique real-time system challenges

	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	100 ns	25 ns
Peak x-N luminosity	10^{34} cm^{-2} s^{-1}	10^{32} cm^{-2} s^{-1}	10^{34} → 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
- But events are precious and have diverse topology → hard to trigger on all process
- Background and systematic control is crucial → avoiding a trigger bias

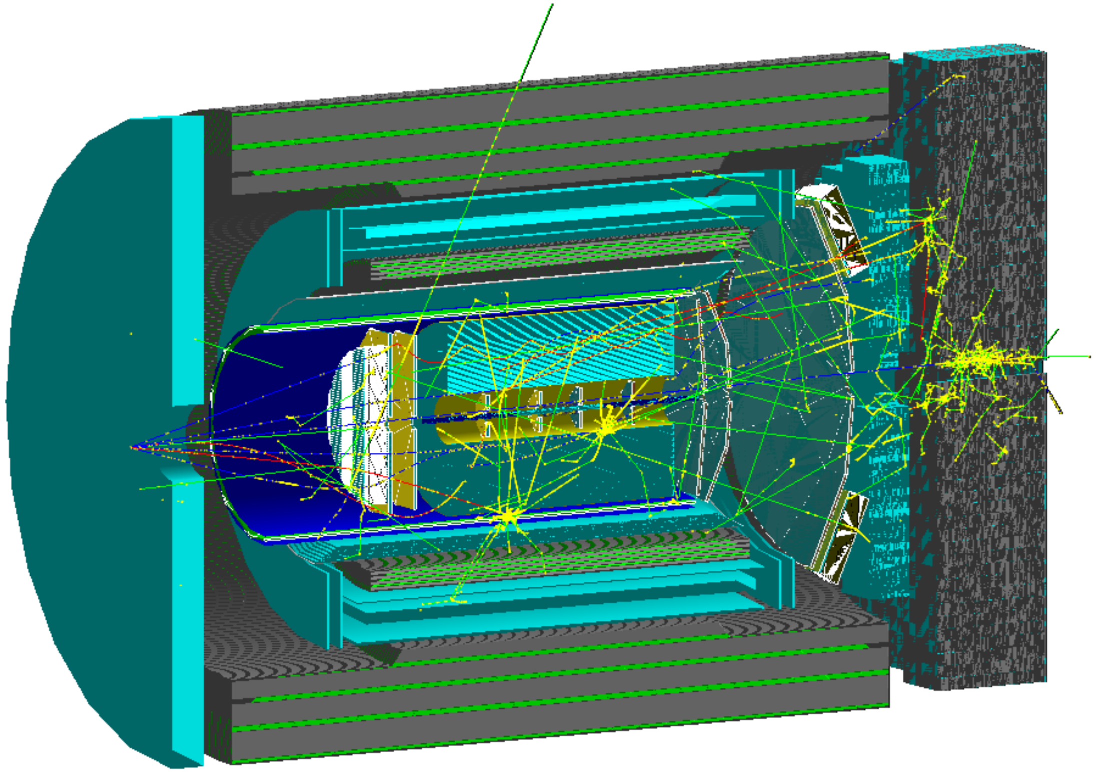
EIC DAQ in Fun4All-EIC simulation

Refs: EIC CDR, sPH-cQCD-2018-001: https://indico.bnl.gov/event/5283/

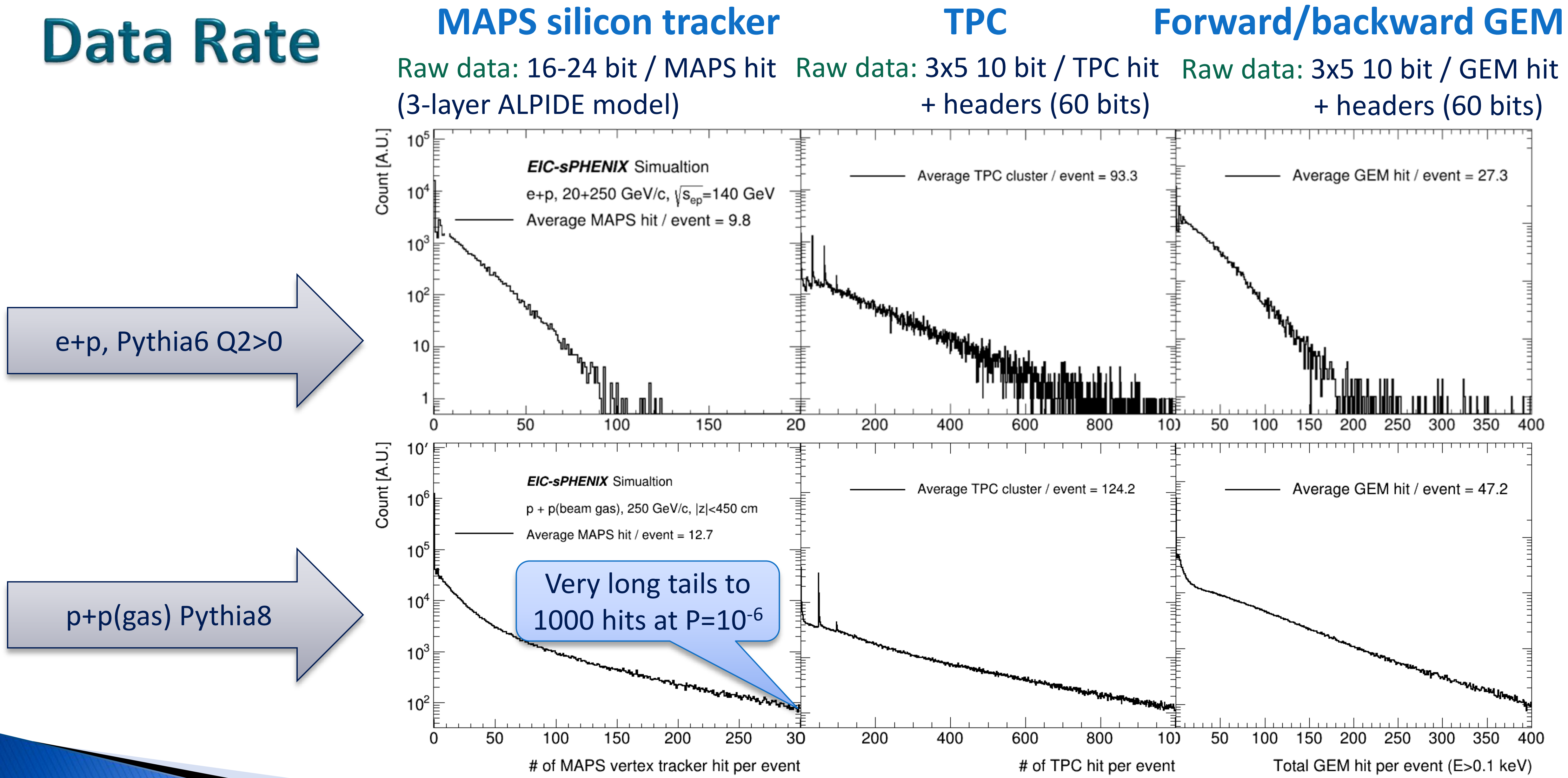


e+p DIS 18+275 GeV/c
$Q^2 \sim 100\ (\text{GeV}/c)^2$

Beam gas event
p + p(gas), 275 GeV/c
at z=-4 m



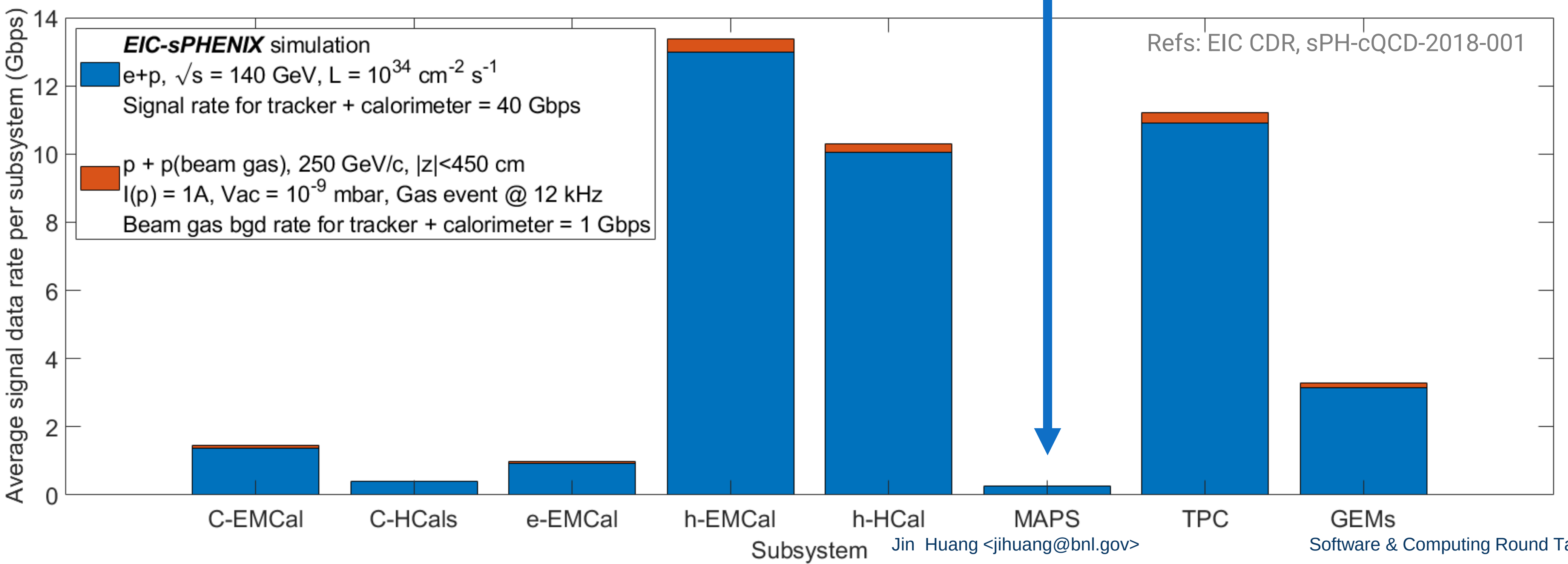
Data Rate



Refs: ECI CDR, sPH-cQCD-2018-001: https://indico.bnl.gov/event/5283/

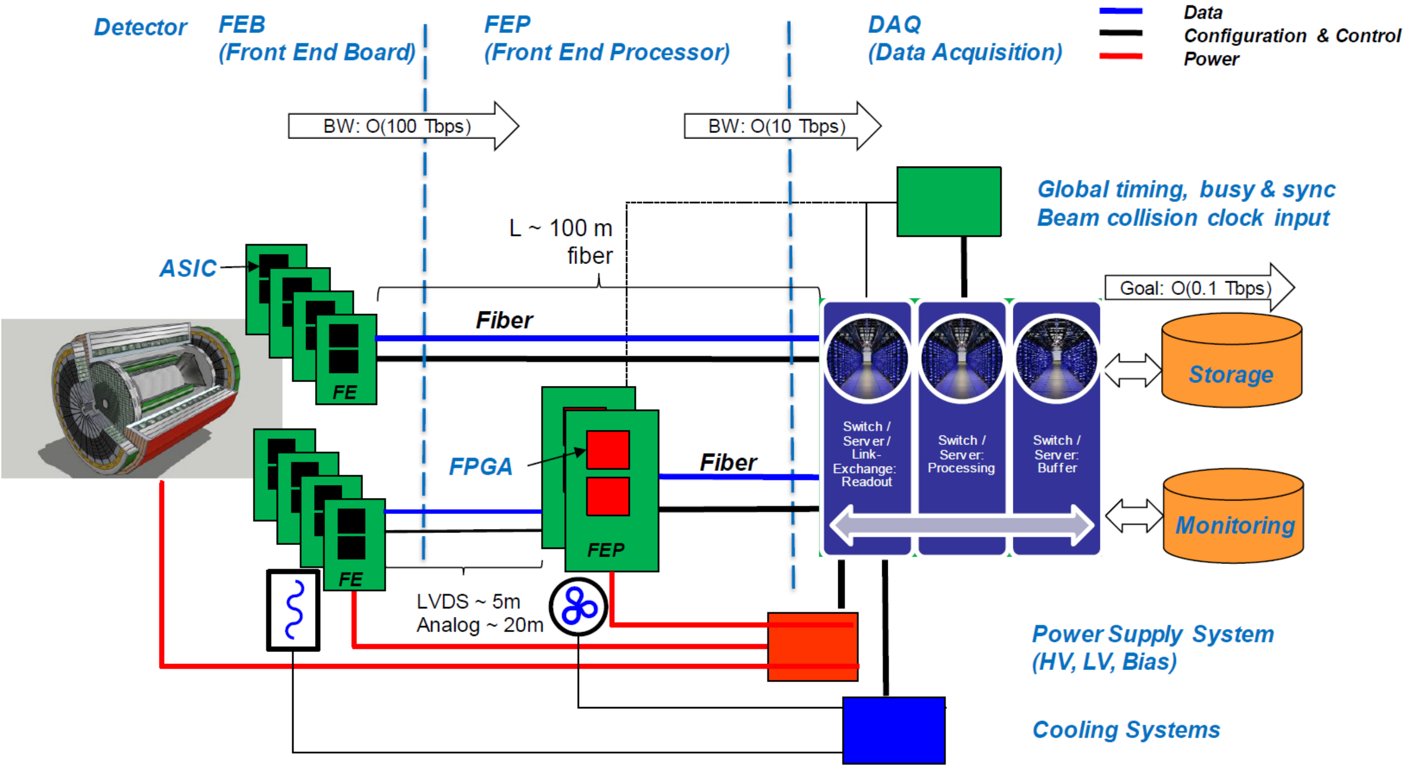
Signal data rate -> DAQ strategy

- What we want to record: total collision signal ~ 100 Gbps @ 10^{34} cm^{-2} s^{-1}
  - Assumption: sPHENIX data format, 100% noise, Less than sPHENIX peak disk rate. 10^{-4} comparing to LHC collision
- Therefore, we could choose to stream out all EIC collisions data
  - In addition, DAQ may need to filter out excessive beam background and electronics noise, if they become dominant.
  - Very different from LHC, where it is necessary to filter out uninteresting p+p collisions (CMS/ATLAS/LHCb) or highly compress collision data (ALICE)



Strategy for an EIC real-time system

- **EIC streaming DAQ**
  - → Triggerless readout front-end (buffer length : µs)
  - → DAQ interface to commodity computing (e.g. FELIX/CRU). Background filter if excessive background rate
  - → Disk/tape storage of streaming time-framed zero-suppressed raw data (buffer length : s)
  - → Online monitoring and calibration (latency : minutes)
  - → Final Collision event tagging in offline production (latency : days+)

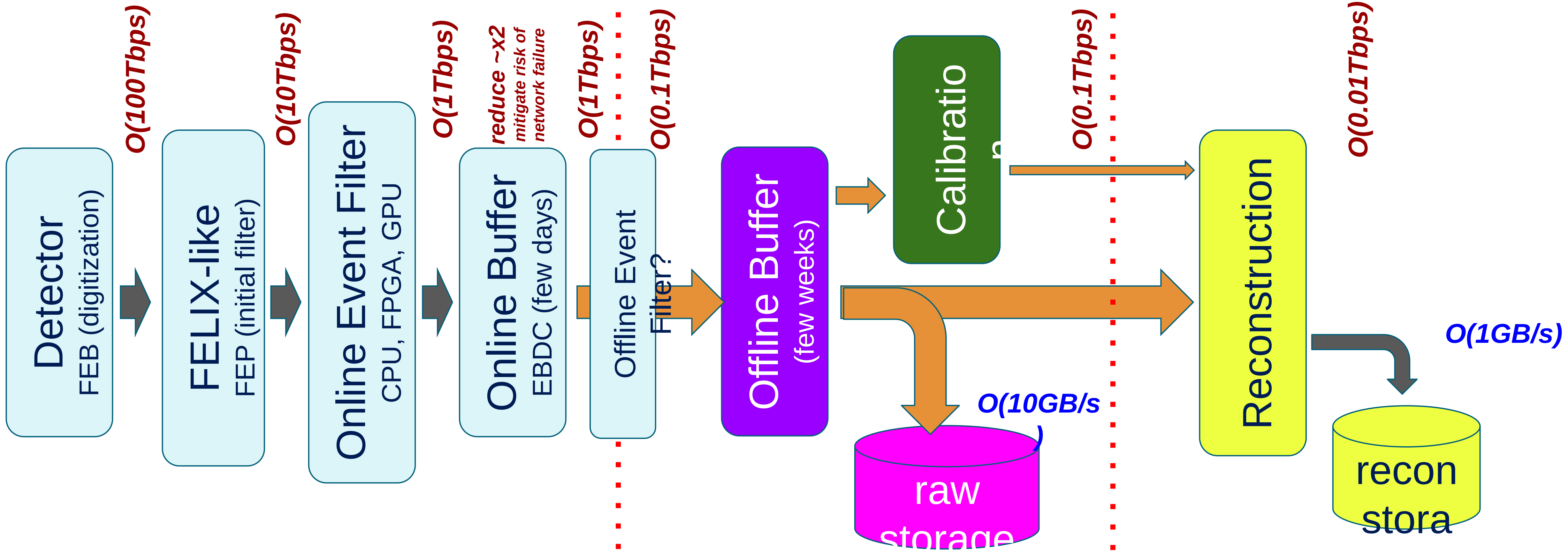


Ref: EIC-CDR

Blured boundary with offline computing

Countesy: David Lawrence
ECCE computing model [link]

See also: last talk M. Battaglieri



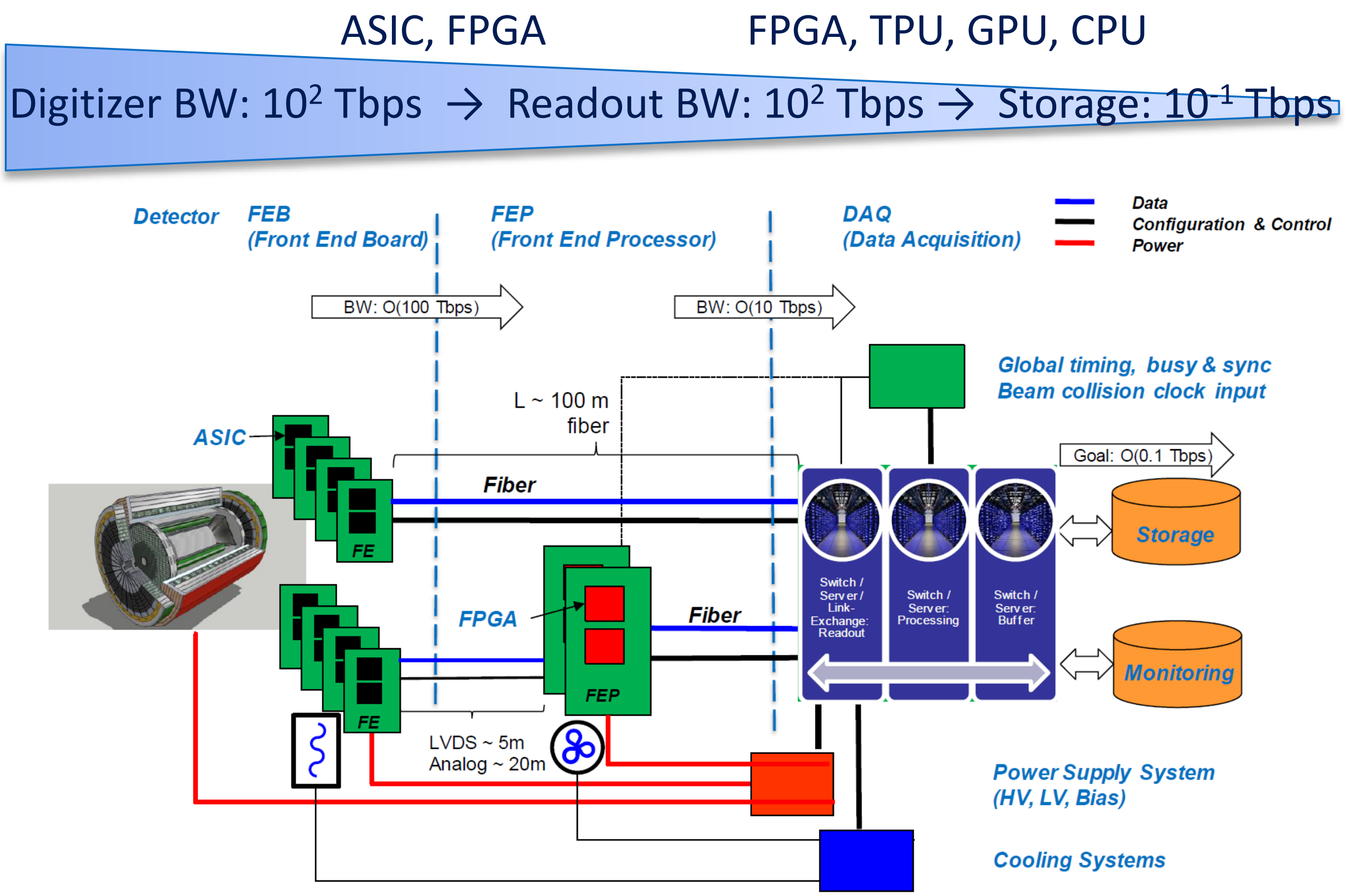
Experimental Hall and Counting House (Project Funds)

Data Center(s): SDCC [,JLab, …] (Operations Funds)

HTC Compute Facilities SDCC ,JLab, … (Operations)

Real-time computing for streaming data pipeline

- Despite low signal rate, the raw data rate can be filled with noises and background
  - Need low background & low noise detector & electronics design
- An essential job of EIC real-time computing: reliable streaming data reduction to fit permanent storage (next topics)
- And more traditional roles for online/offline server farm:
  - Online monitoring/fault det.
  - Calibration
  - Production → Initial analysis pass



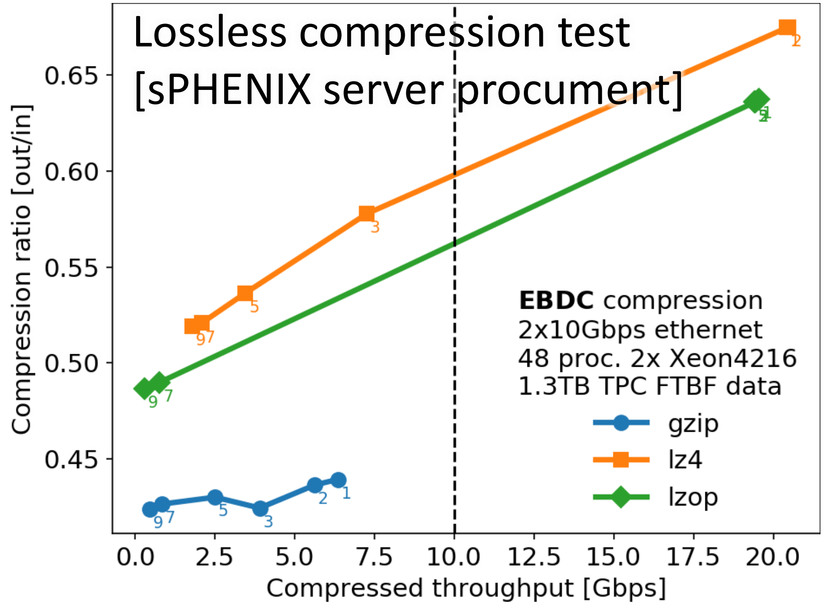
[EIC CDR]

Online computing for streaming data – trigger throttling

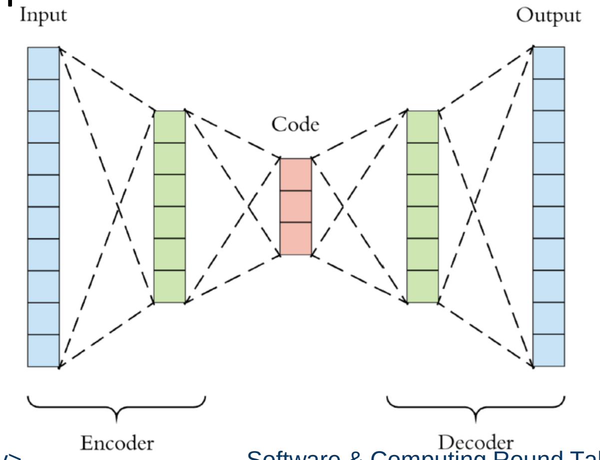
- At the beginning of the EIC operation, background & noise rate could be unpredictable and high
- A contingency method: throttling streaming data with triggering
  - Immediately reduce streaming data by orders of magnitudes
  - Widely used hardware producing trigger, fix latency or HLT (Aaji’s talk)
  - Has physics loss, added systematic uncertainty for hardware trigger efficiency
- Can utilize ML to produce more complex triggering on FPGA
  - PID trigger, e.g. ref: S. Furletov @ streaming workshop VIII [link]
  - Tracking-event topology trigger: D. Yu @ AI4EIC workshop [link]

Online computing for streaming data - compression

- Lossless compression
  - Compress by ~1/2
  - Well established fast compression algorithm
- Lossy compression
  - Opportunity for unsupervised machine learning based on data, e.g.
  - Auto-encoder on ASIC for HGCal @ CMS [link]
  - Bicephalous Convolutional Neural Encoder for zero-suppressed data (next)

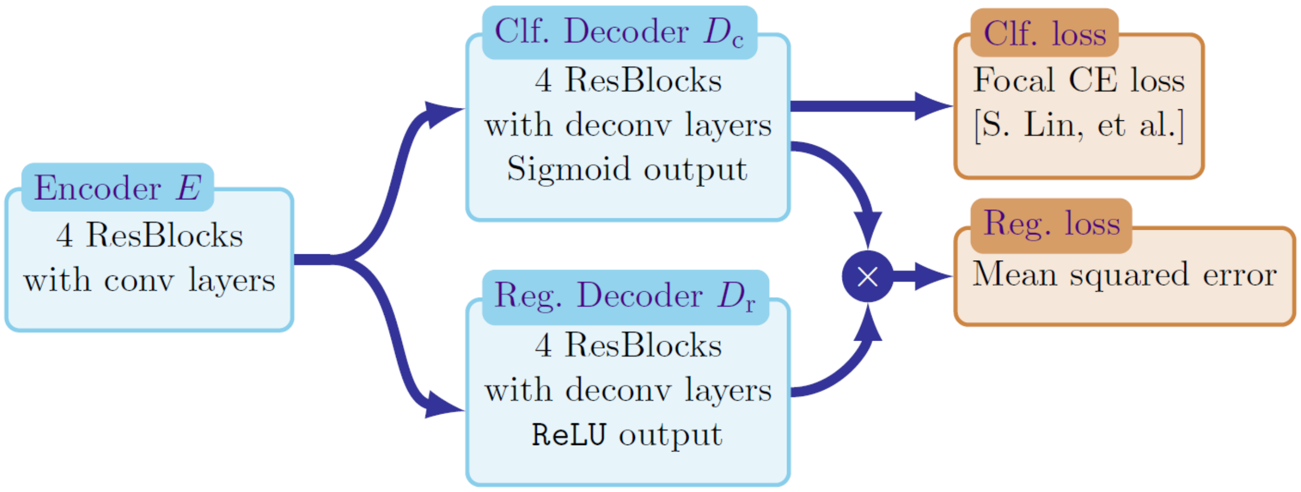


Simple auto-encode neural network

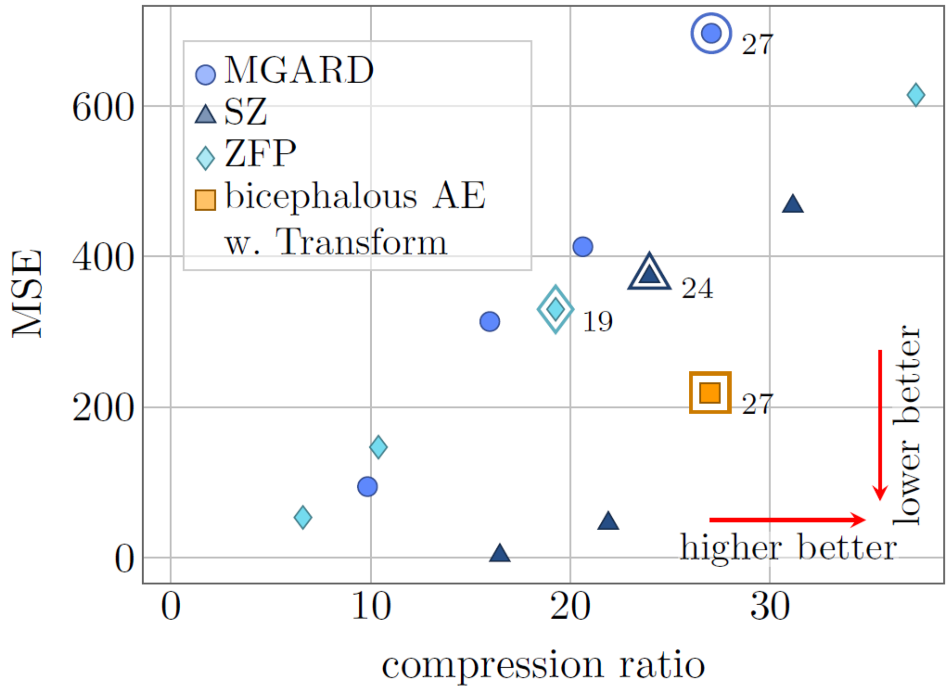


Bicephalous Convolutional Auto-Encoder for zero-suppressed data

- Some detector ADC data is challenging for Auto-Encoder, e.g. features such as zero-suppression cut off
- A dual-output auto encoder is designed to output both a region of interest and decompressed ADC. Possibility for further noise filtering
- Ref: Y. Huang @ AI4EIC workshop [link]



Compression comparison with published compressor tested on busiest sPHENIX TPC timeframes

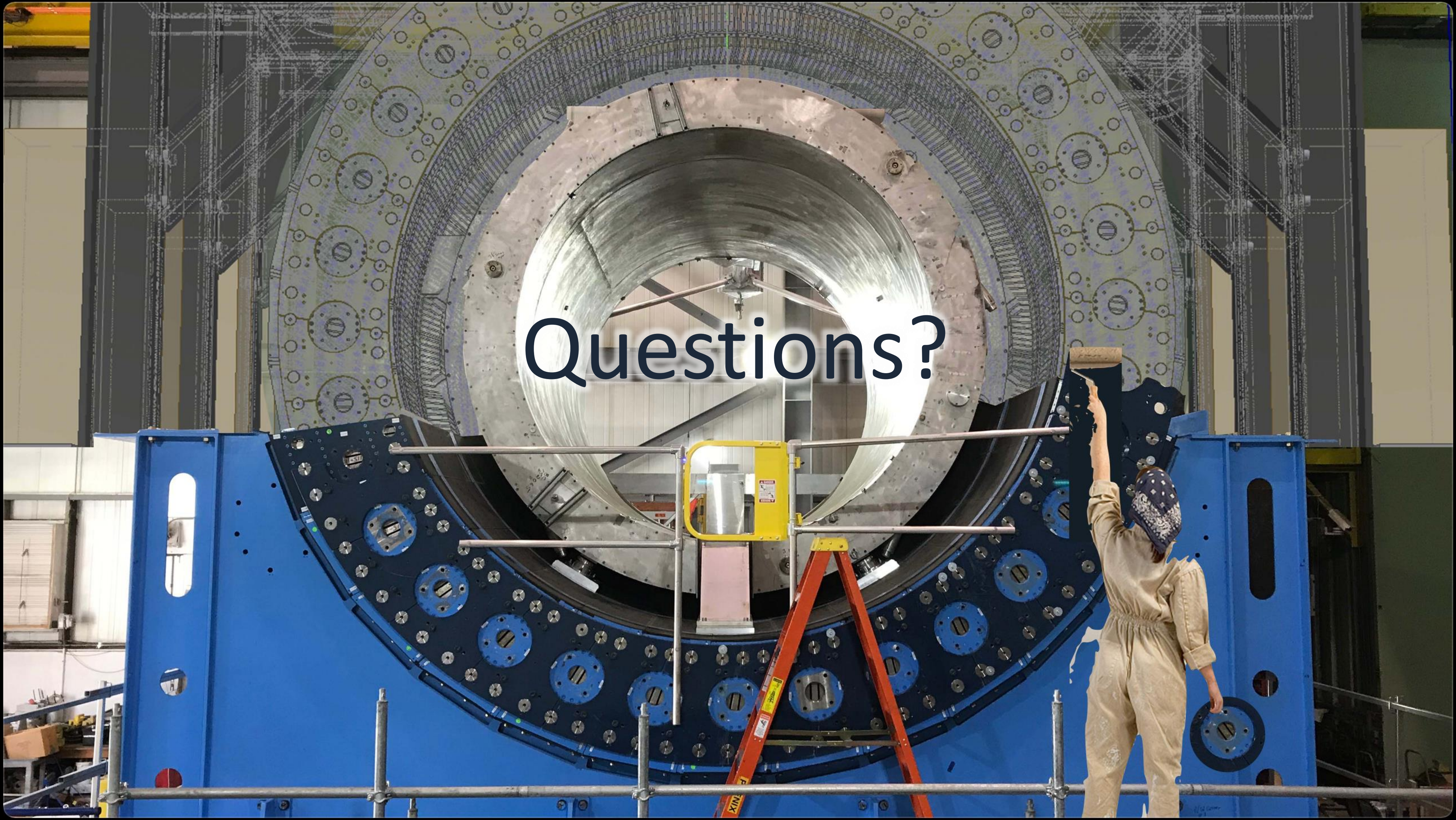


Online computing for streaming data - feature building

- Another effective way of suppressing background is feature building, e.g.
- Clustering on calorimeter
  - Effective in suppressing single tower noise
  - e.g. CLAS12 test as in M. Battaglieri's talk
- Tracklet building on tracker:
  - Effective in suppressing isolated noise, such as the synchrotron background
  - e.g. ALICE TPC streaming data [arXiv:1910.12214]. D. Yu @ AI4EIC workshop [link]
- ADC timeseries -> amplitude extraction
  - e.g. Specialized filters: C. Crawford @ Streaming readout VIII [link];
  - Neural network on ASIC: S. Miryala @ Streaming readout VIII [link]

Summary

- sPHENIX under construction for start of data taking in 2023
  - Recent development on streaming data (on top of streaming readout) for sPHENXI tracker show significant increase of physics capability
  - Strong link with EIC both technical implementation and physics program
- EIC is a unique collider with low x-section, high luminosity and stringent requirement on systematic control.
  - Streaming readout fit well, wide adoption in community [YR, CDR]
  - EIC collision signal data rate is low, but noise/background rate uncertain
  - EIC online computing: key development on reliable noise/background filtering



Remix credit: Dave Morrison

Extra information

EIC x-sec : further quantification [Courtesy E. Aschenauer]

- Inelastic e+p scattering x-sec:
  - For a luminosity of 10^{34} $cm^{-2}s^{-1}$ 50ub corresponds to **500 kHz**
- Elastic e+p cross-section:
  - For EIC central barrel, elastic cross section is **small** comparing to the inclusive QCD processes
- Beam gas interaction:
  - Beam proton – beam gas fix target inelastic interactions. The pp elastic cross section is smaller (~7 mb)
  - For a vacuum of 10^{-9} mbar in the detector volume (10m) this gives a rate of **14 kHz**

Beam [GeV]	HERA	5 x 50	10 x 100	18 x 275
$Q^2>10^{-9}$ GeV	65.6	**29.9**	**41.4**	**54.3 ub**
$Q^2>1$ GeV	1.29	0.45	0.65	0.94 ub

Beam [GeV]	HERA	5 x 50	10 x 100	18 x 275
σ [y_{Exp} >-4]	5 pb	**5 ub**	**0.7 ub**	**0.06 ub**
σ [y_{Exp} >-6]	11 ub	420 ub	100 ub	29 ub

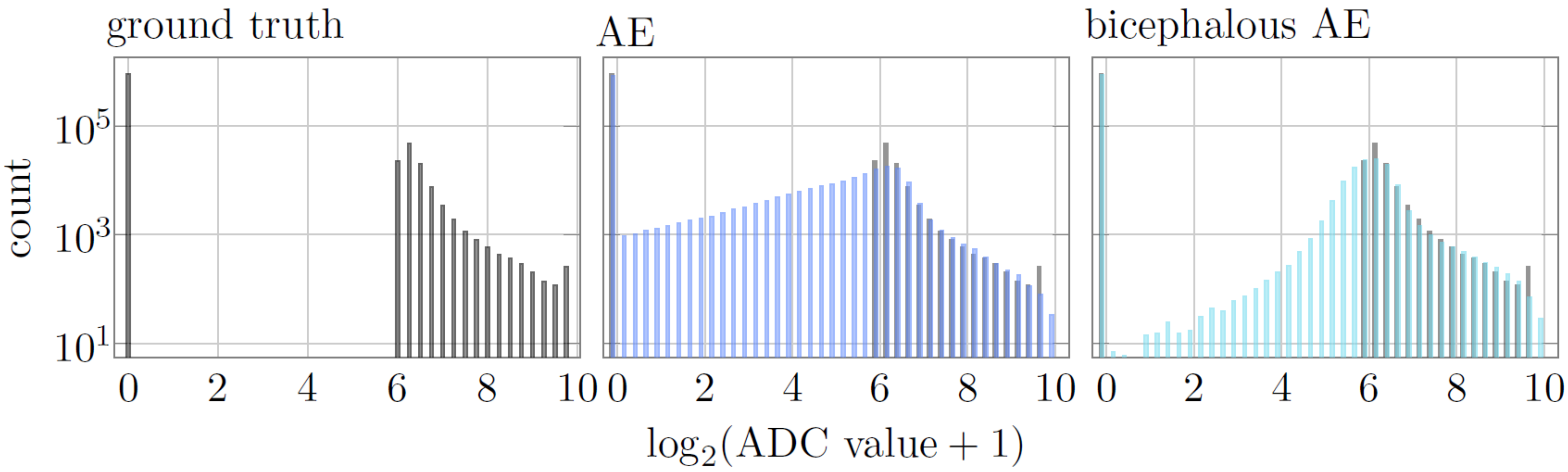
E_p:	50 GeV	100 GeV	275 GeV	920 GeV
	38.4 mb	**38.4 mb**	**39.4 mb**	41.8 mb

Brookhaven National Laboratory

Results I: AE v.s. Bicephalous AE

Compression ratio is 1 : 27

(1 : 3 for SAMPA ASIC for this busiest event)



Result III. Ablation Study

ground truth | AE | bicephalous AE | bicephalous AE w. transform