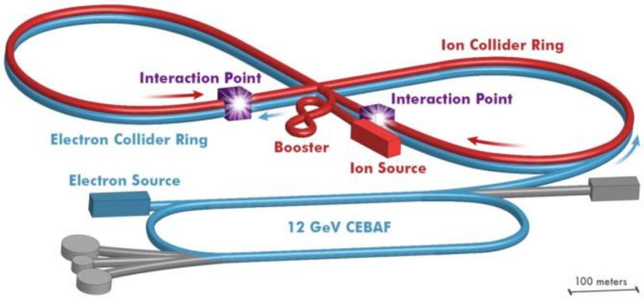
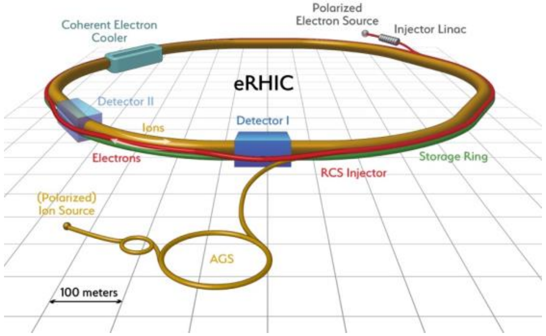
Streaming readout for a 12 GeV CEBAF and a future EIC.

CEBAF at Jefferson Lab

EIC at Jefferson Lab



EIC at Brookhaven Lab



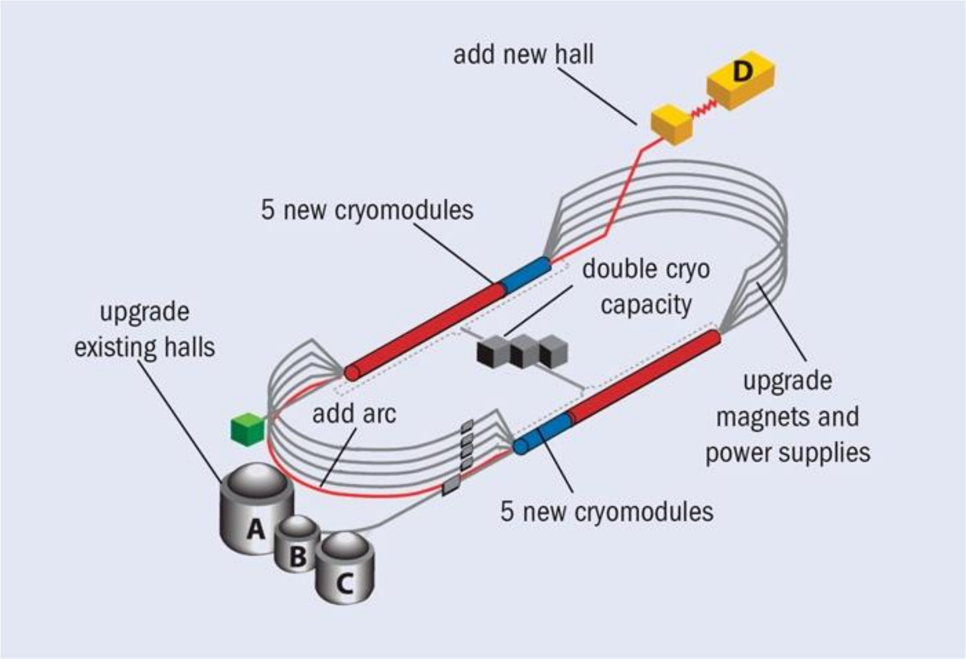
Graham Heyes

Introduction

- About Jefferson Lab.
- Current State Of The Art
- Looking forward:
  - JLab 12 GeV experiments.
  - An Electron Ion Collider.
- Trends in computing and electronics.
- What is Streaming Readout?
- Contemporary examples
- Ideas for the future
- Where do we go from here

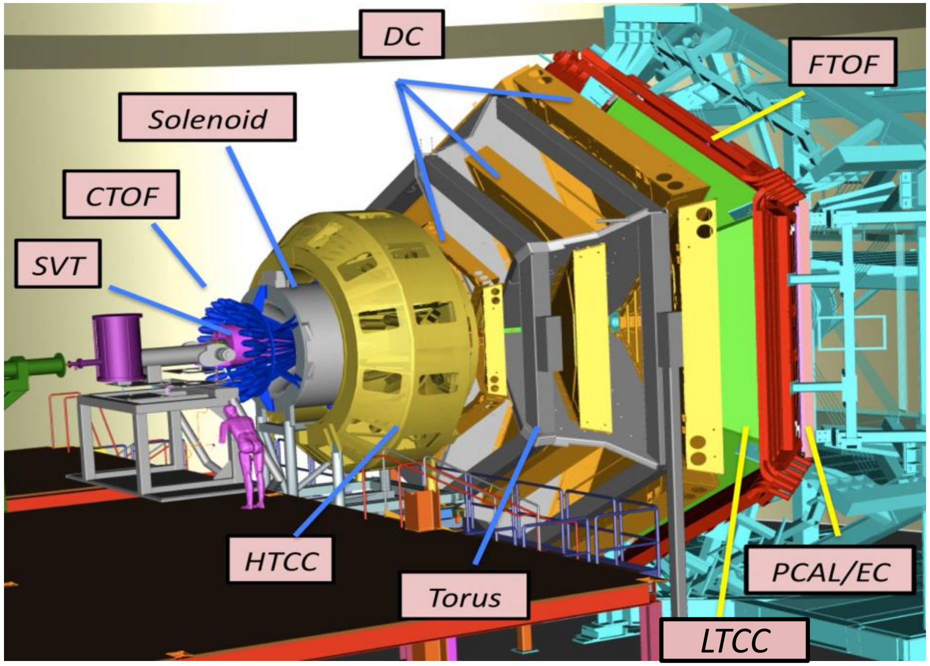
Jefferson Lab

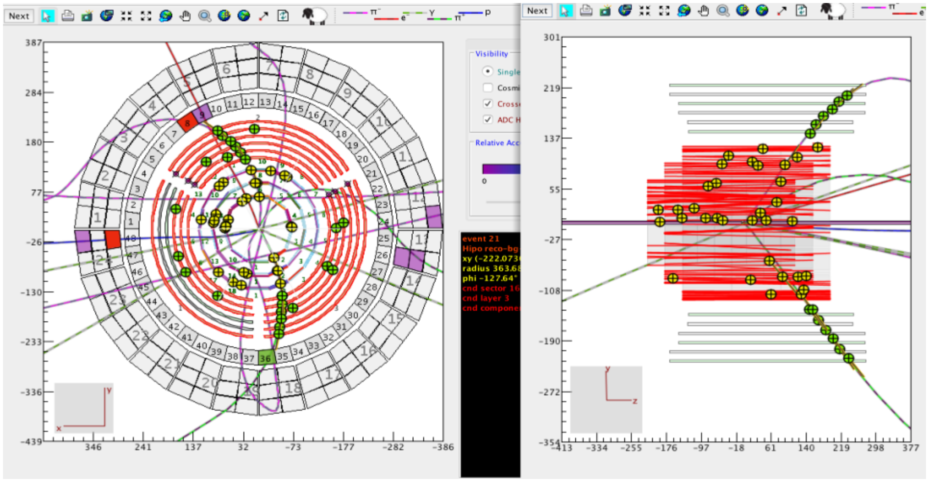
- The principle goal of Jefferson lab is nuclear physics research using the CEBAF electron accelerator.
  - Two superconducting LINACs with recirculating arcs.
  - Simultaneous beam to multiple halls.
- Each hall has equipment designed to study different, but complementary, aspects of matter in the nucleus.
- Recently we completed an upgrade of the accelerator from 6 GeV to 12 GeV, complemented by upgraded detectors and a new experiment, GLUEX, in a fourth hall.



Hall-B CLAS12 – Large Acceptance Spectrometer

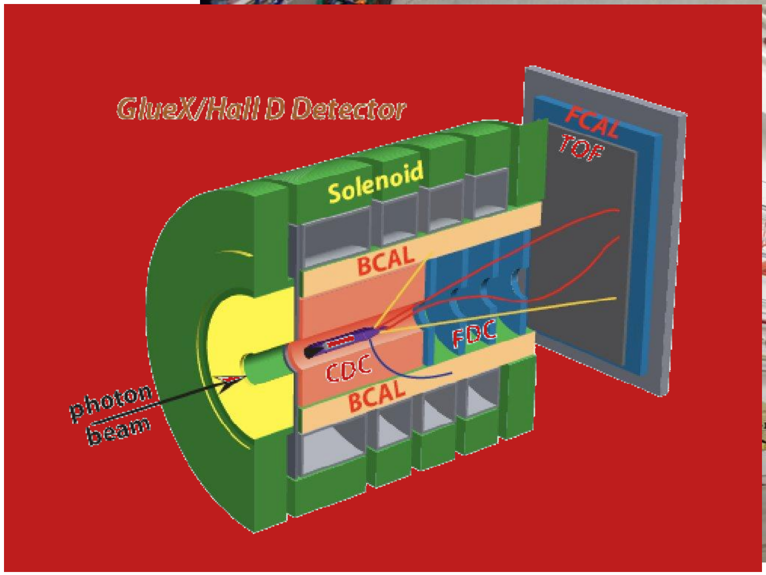
- Large number of individual detectors.
- “Run Groups” of experiments with shared beam and target requirements share data.
- Event rate 12 kHz.
- Event size 50 kB.
    - Working to lower this.
- Over 100 embedded processors.





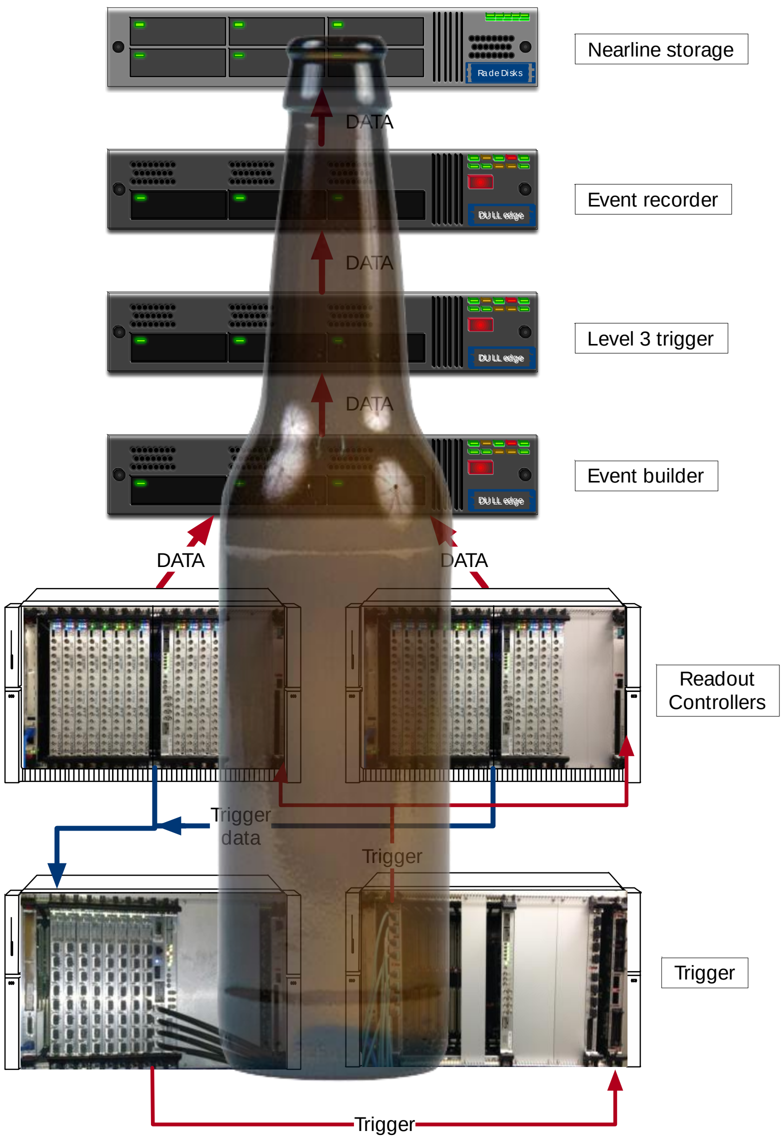
Hall-D GLUEX detector

- Electron beam converted into photon beam.
- Event size 17 kB.
- Event rate 90 kHz.
- Data rate ~1.5 GB/s
- Data is spread over 50+ VME systems.
- “Pipeline mode” readout.



Summary of SoA at JLab – pipeline mode

- Data is split into trigger and DAQ paths.
- Trigger data read over VXS serial backplane by Crate Trigger Processor and passed to global trigger.
- Trigger formed in electronics
- Blocks of data queued in "pipeline" while trigger is made.
- Readout of the DAQ data by embedded CPU over VME.
- Data sent over network to Event Builder
- Pros:
  - Data stream is a stream of events containing data from detectors.
  - Trigger filters "unwanted data".
  - Well understood way of doing things.
- Cons:
  - Have to delay prompt data until slowest data appears.
  - Relies on good understanding of trigger.
  - All parts of DAQ have to work.
    - One failure stops the pipeline.
  - Doesn't work well when events overlap in time.
  - Major bottlenecks – worked around via parallel multi-stage event builder.



Real world example

Intel CPU Read Out Controller (ROC) running Linux

Individual ADC channel

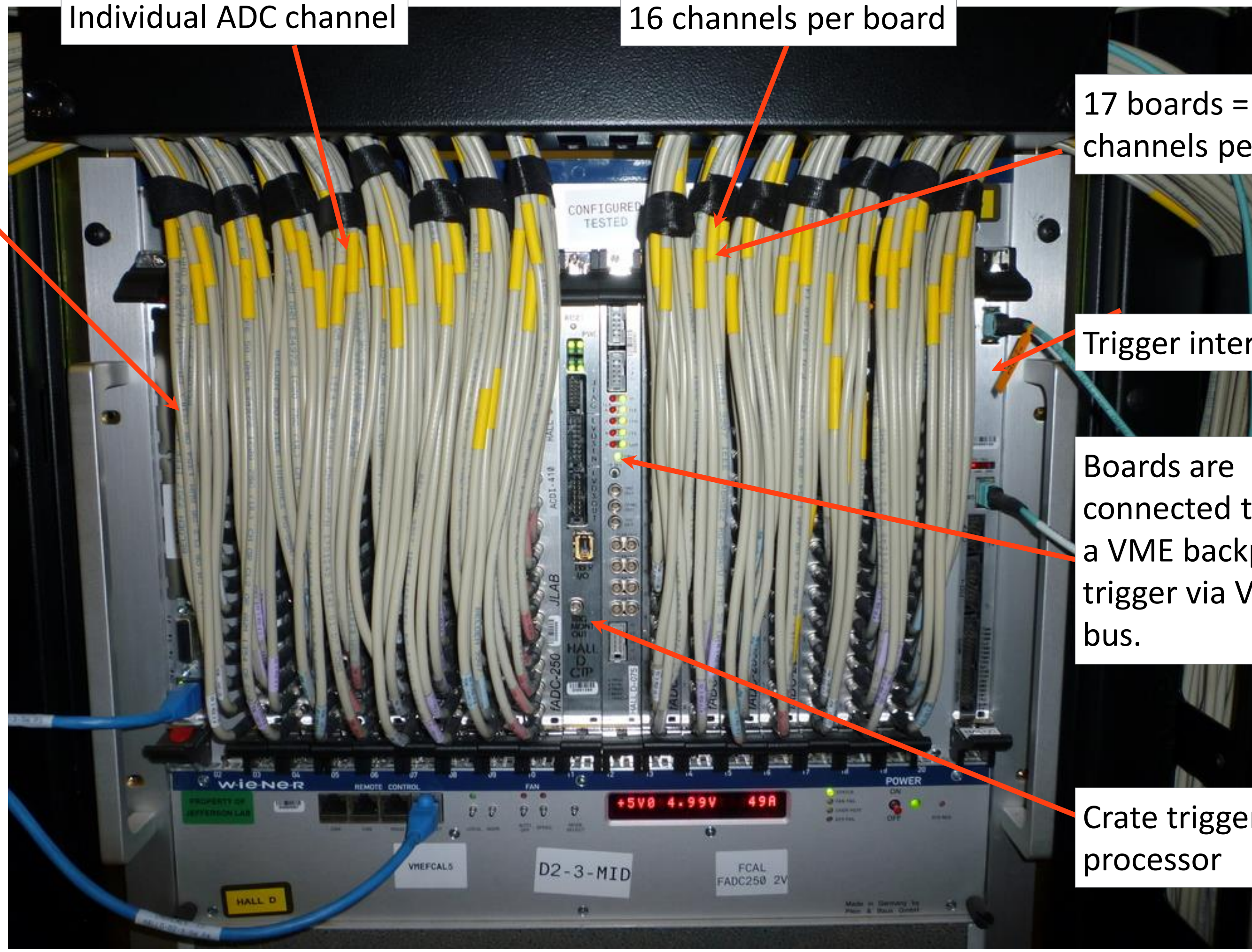
16 channels per board

17 boards = 272 channels per crate

Trigger interface

Boards are connected to CPU via a VME backplane and trigger via VXS serial bus.

Crate trigger processor

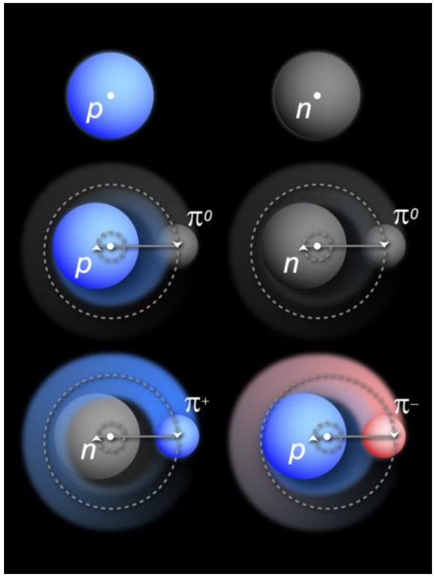


Not cheap.
VXS crate +
ADCs +
CPU +
Trigger interface +
Crate trigger processor +
Trigger Signal Distribution
~$90k

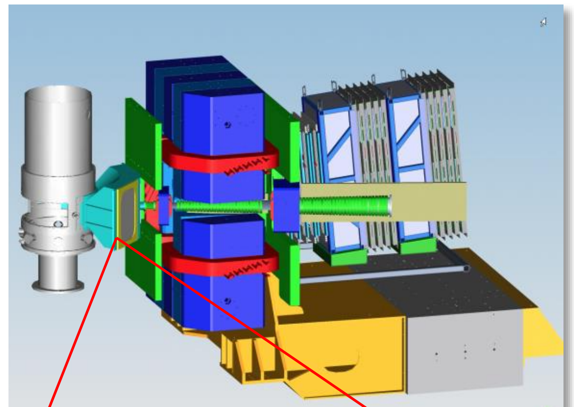
Looking forward

- New experiments are being proposed
  - Detectors that do not play well together due to timing.
    - Traditional trigger and event builder strategies are not ideal.
  - Detectors with peculiar topologies.
    - Detectors split or segmented in a way that makes forming a trigger hard.
  - High event and/or data rates.
    - Particles from more than one event in a detector at the same time – need to disentangle.
- The data acquisition from these experiments does not fit well with current techniques.

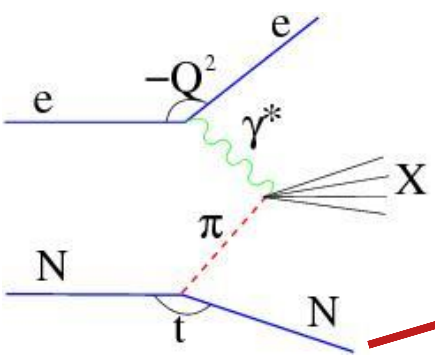
Tagged Deep Inelastic Scattering (TDIS):
Probing the Mesonic Structure in the Nucleon

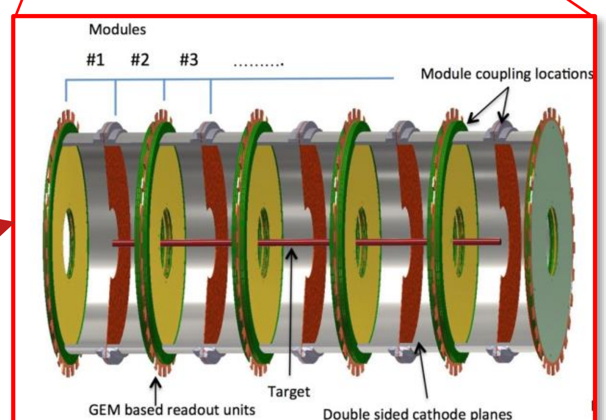


<u>TDIS Scientific Goal:</u> Access Elusive Partonic Structure of Mesons by Using Mesons in Nucleons as "Target"



Scattered electrons detected in planned Hall-A Super Bigbite Spectrometer.





"Spectator" protons detected by radial Time Projection Chamber (rTPC).

TDIS - Challenges

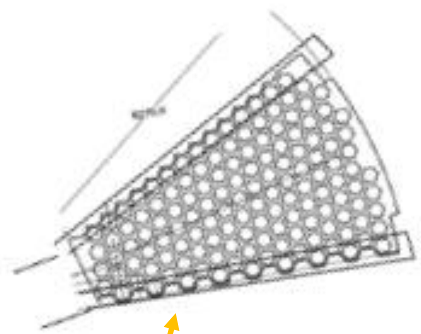
- It takes up to a microsecond for the proton track to trace itself out on the wall of the TPC.
- Super Bigbite electron detectors are much faster – large timing mismatch.
- Large flux of background electrons not associated with a proton.
- Multiple protons in the TPC at the same time.

- TPC has 25,000 pads, hit rate per pad ~800 kHz.
- Data rate from TPC up to 4 Gbyte/s total.

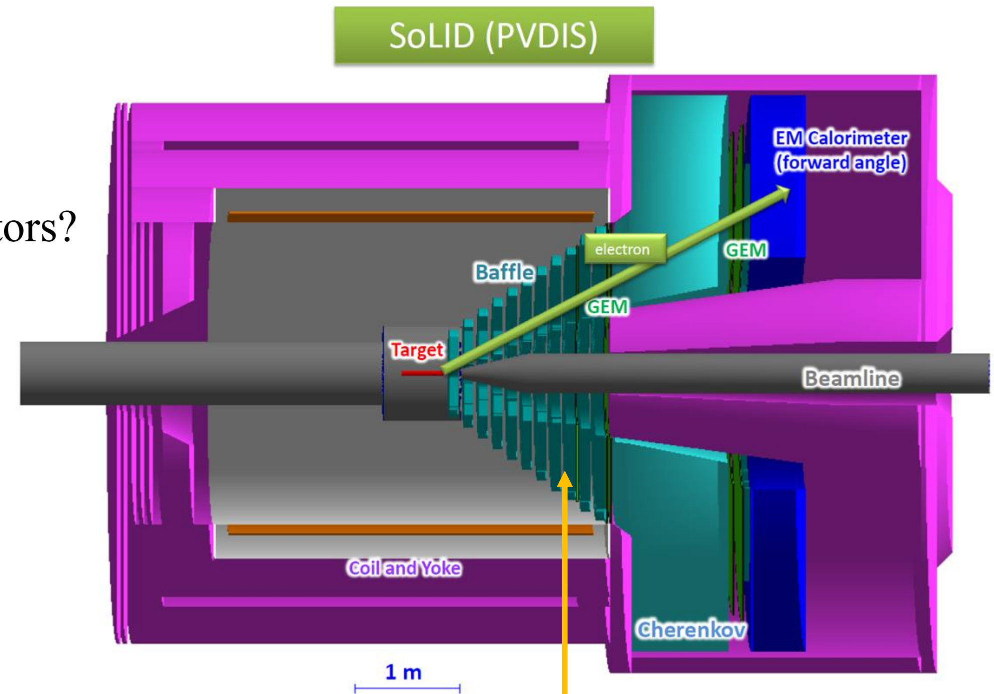
- How to read this out and match up the electrons with the protons?
  - Event building online at these rates is a bottleneck.

SoLID

- SoLID is another experiment proposed for installation hall-A at JLab.
- In the PVDIS configuration electrons are scattered of a fixed target at high luminosity.
- Spiral baffles cut background.
- The detector is split into radially 30 sectors, the single track event topology allows 30 DAQ systems to be run in parallel at rates of up to 1 Gbyte/s each 30 Gbyte/s total
- Challenges :
  - How to handle 30 Gbyte/s affordably?
  - Hits at sector edges span two sectors?
  - How to integrate GEM with other detectors?

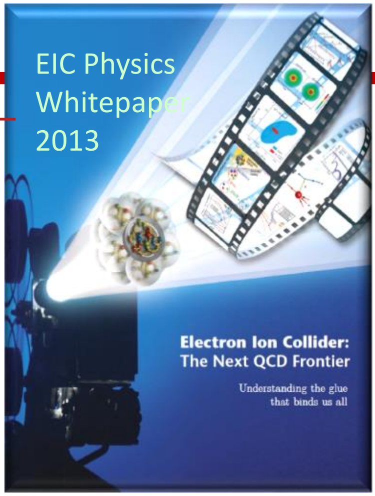


GEM detectors are segmented into 30 sectors.



Baffles constrain momentum

Electron-Ion Collider: The Next QCD Frontier

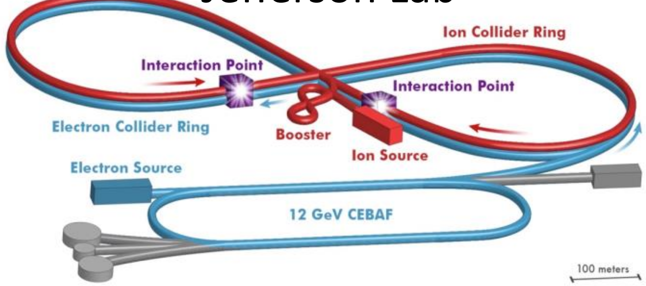


Electron-Ion Collider (EIC) parameters

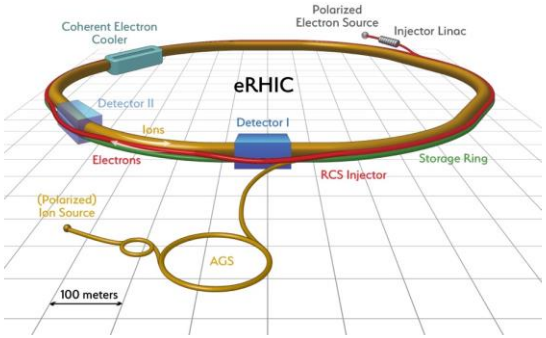
- Highly polarized ($\sim 70\%$) electron and nucleon beams
- Ion beams from deuteron to the heaviest nuclei (uranium or lead)
- Variable center of mass energies from $\sim 20 - \sim 100$ GeV, upgradable to ~ 150 GeV
- High collision luminosity $\sim 10^{33-34}$ cm^{-2}s^{-1}
- Possibilities of having more than one interaction region

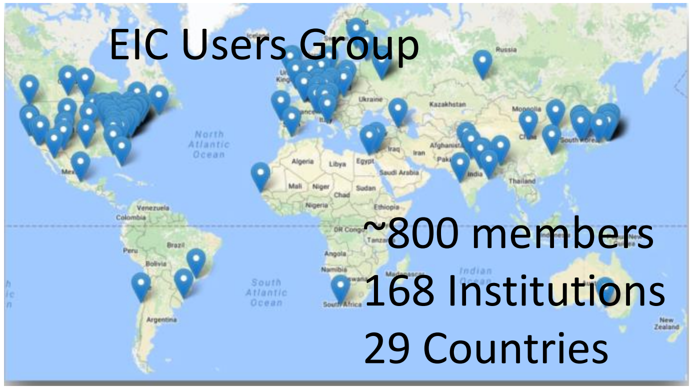
Realization Proposals

Jefferson Lab



Brookhaven Lab





EIC Users Group

~800 members
168 Institutions
29 Countries

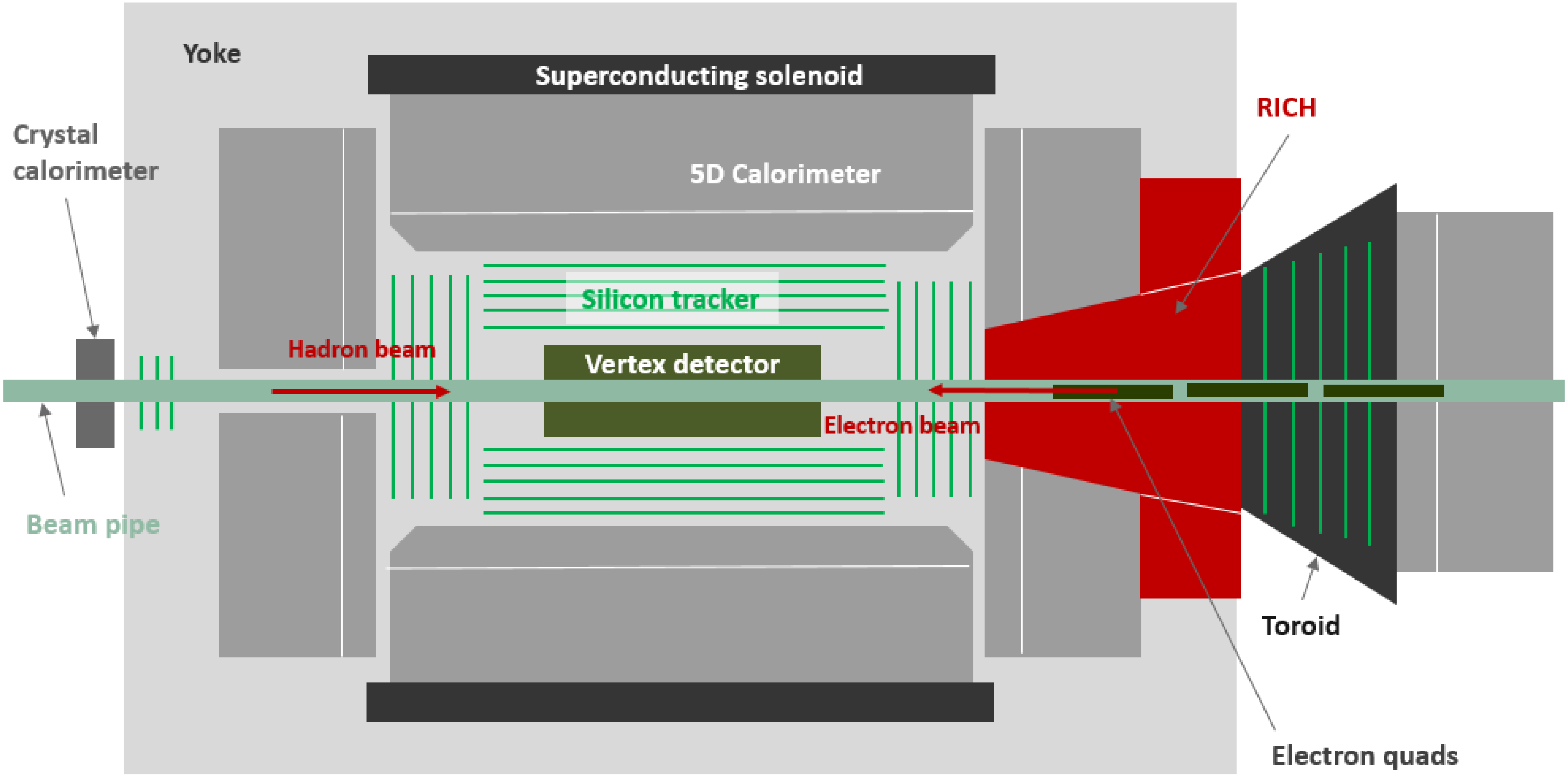
NATIONAL ACADEMY OF SCIENCES

NAS report on EIC science imminent

Next Step: DOE CD0 (2018 or 19?)

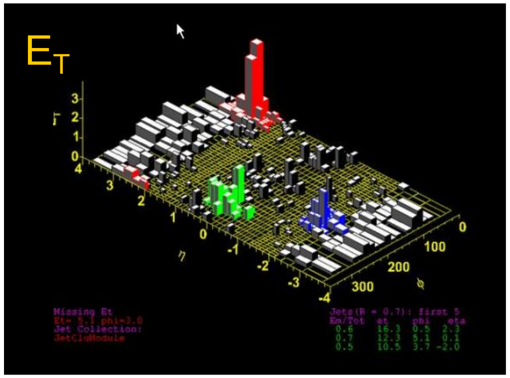
Jefferson Lab

TOPSiDE – 5D Concept



Area of silicon ~1,400 m^2 or \$14M @ \$1/cm^2

(Compare: CMS HGCAL upgrade ~ 600 m^2)



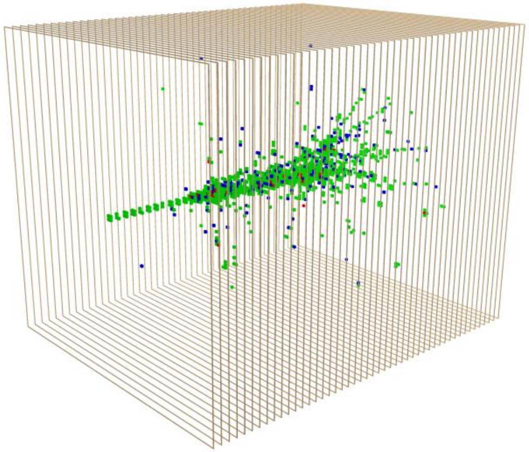
A) Imaging Calorimetry

Replace

Tower structure with very fine granularity (lateral and longitudinally)
Few 1,000 channels -> few 10,000,000 channels
Option to reduce resolution on single channels to low-bit depth

Technologies developed in past decade

Silicon sensors with 1 x 1 cm^2, 0.5 x 0.5 cm^2 and 0.16 cm^2 pixels
Scintillator strips (4.5 x 0.5 cm^2) or scintillator pads (3 x 3 cm^2)
Resistive Plate Chambers with 1 x 1 cm^2 pads
Micromegas and GEMs with 1 x 1 cm^2 pads

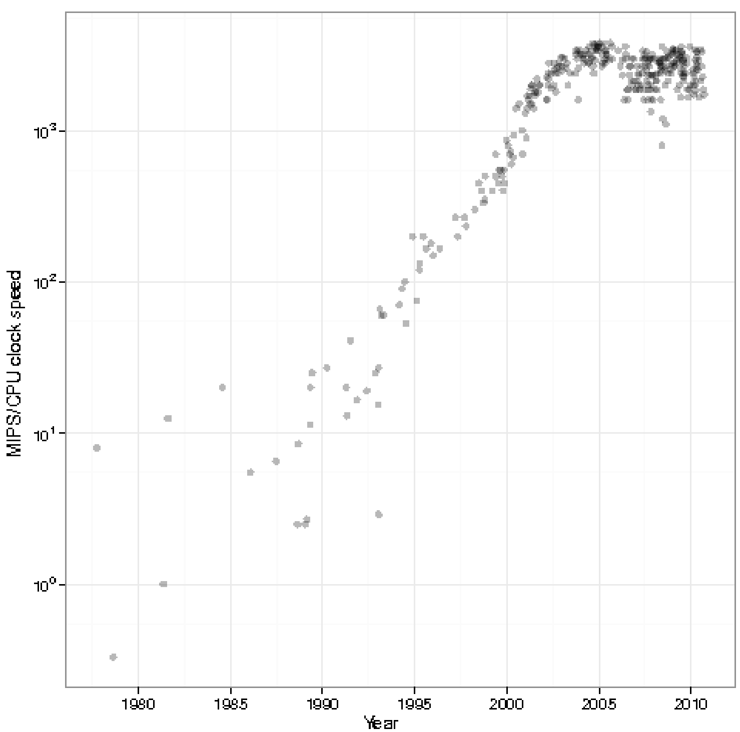


These technologies have been (mostly) validated

Trends - computing

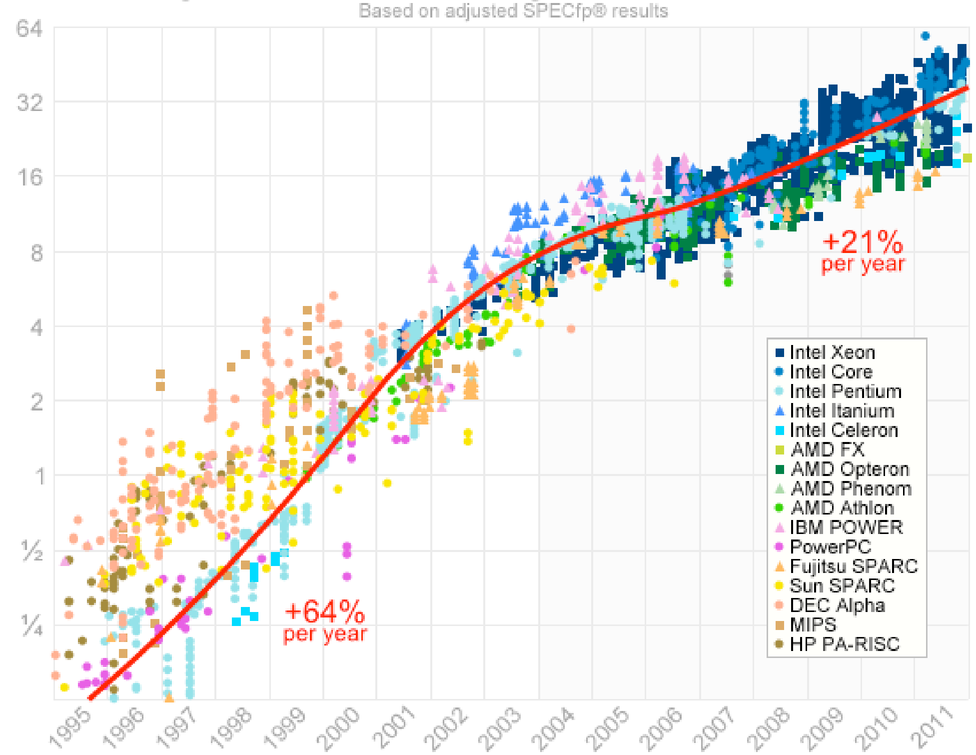
- Famous Moore's law - transistor count doubles every two years.
- In fact MIPS/clock speed plateaued in ~2005.
  - Matters because online we rely heavily on integer performance.
- Remaining gains rely on a much slow growth in clock speed since ~2000.
- Current advances by squeezing more cores per chip and improved architectures – vector processing units etc.

MIPS / clock speed



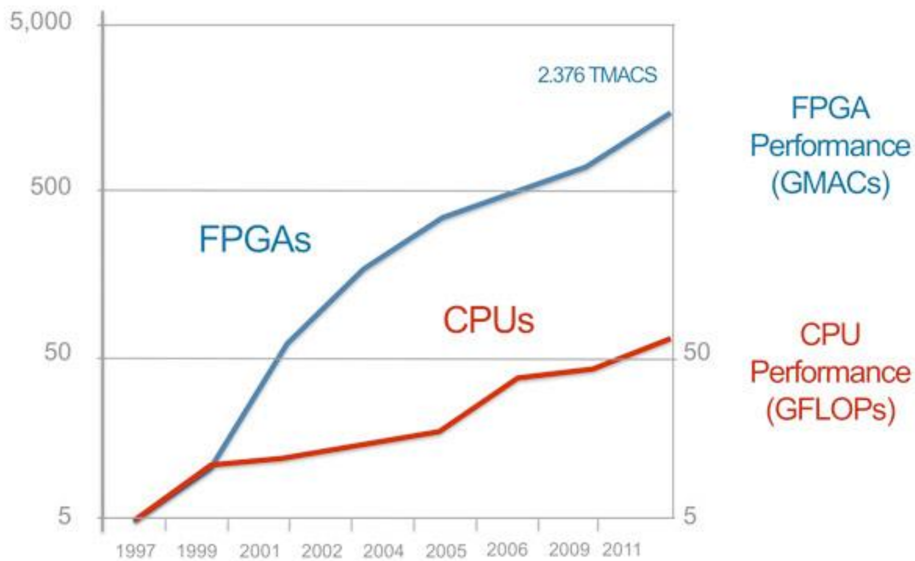
Single-Threaded Floating-Point Performance

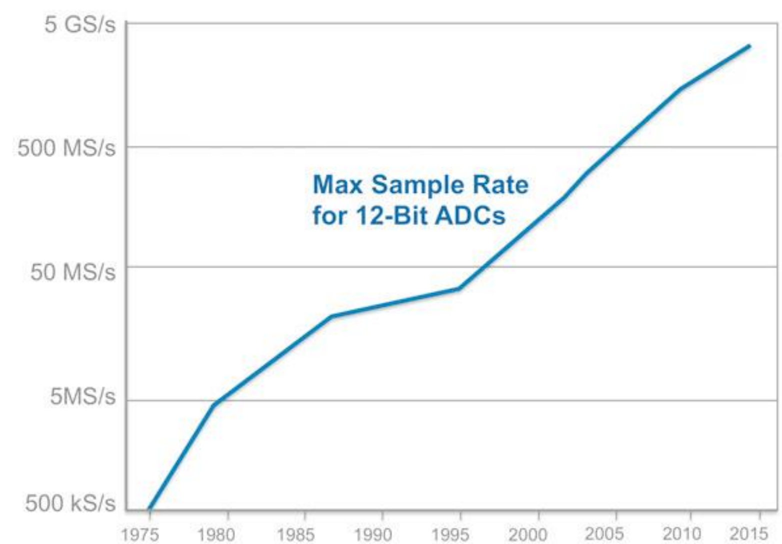
Based on adjusted SPECfp® results



Trends - electronics

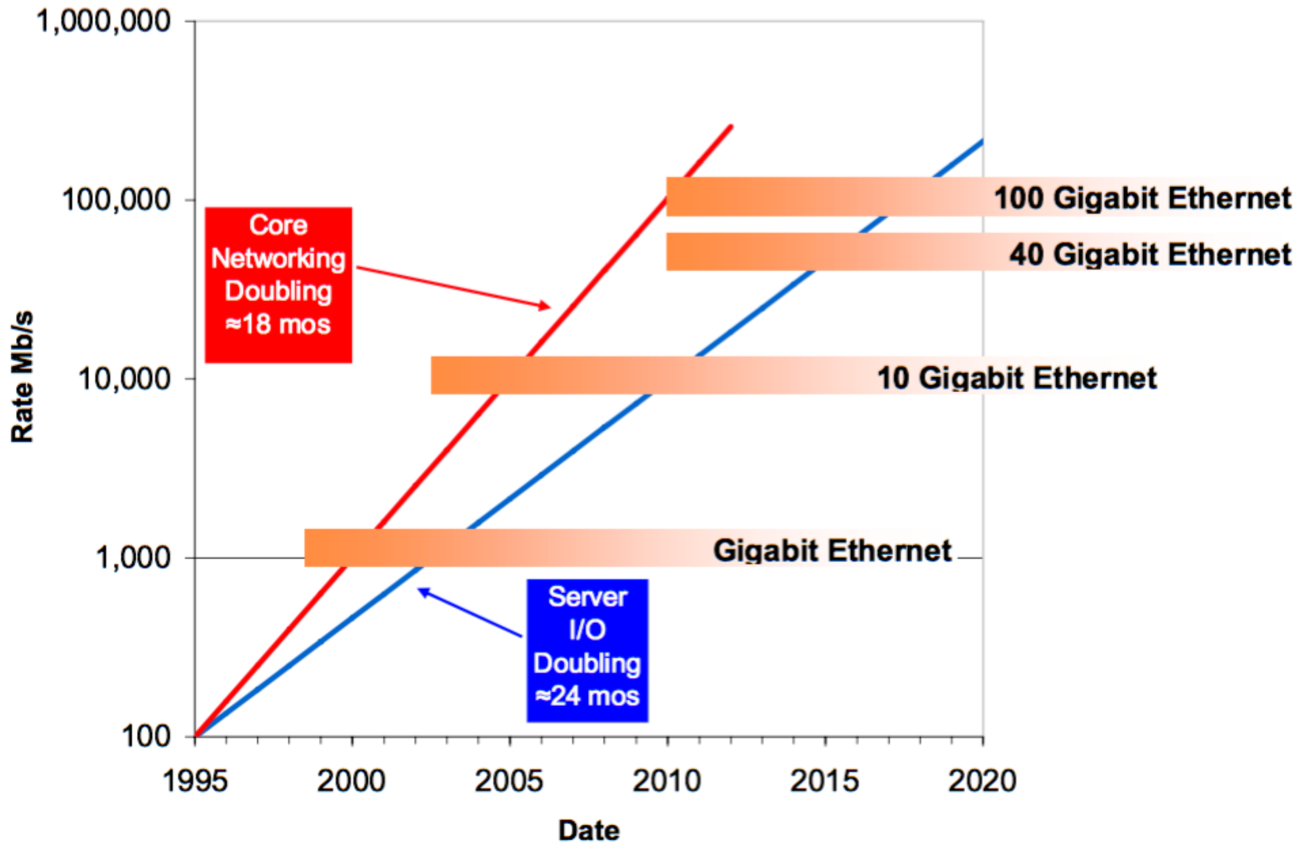
- Already at JLab the number of channels per crate is limited by VMEbus bandwidth.
  - Move to other industry standards like MicroTCA?
  - Crateless serial fiber networks.
- Current trend is to push some functionality currently performed in software running on embedded processors into firmware on custom electronics. Drivers are :
  - Slow increase in single thread CPU performance.
  - Availability of “large”, easily programmable, affordable components.
- Delay between when technology is developed and when it becomes affordable for use in custom electronics means that there is room for growth over the next ten years.





Trends - data transport

- What matters is not only what technology exists but also whether it marketable and affordable.

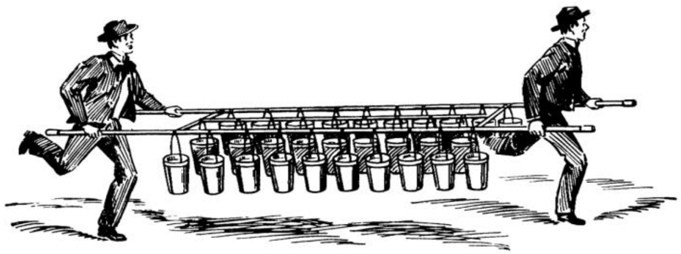


What would we expect to be happening?

- If we extrapolate current trends:
  - CPUs are becoming more powerful but the performance that matters for online systems is achieved mainly through doing more in parallel rather than improvements in per-core performance.
  - FPGA performance, affordability and usability are still improving.
  - IO and serial network bandwidth still seem to be growing exponentially.
- It is time to revisit the ideas that have dominated nuclear physics detector readout for the past twenty or thirty years.
  - Move things that we moved from hardware to software 25 years ago back into hardware (or at least firmware).
  - Reevaluate the use of busses – serial links may be more cost effective.
  - Reevaluate data flow.

Streaming mode

- In traditional triggered readout:
  - Data is digitized into buffers and a trigger, per event, starts readout.
  - Parts of events are transported through the DAQ to an event builder where they are assembled into events.
  - At each stage the flow of data is controlled by “back pressure”.
  - Data is organized sequentially by event.
  - A variation increases bandwidth by moving blocks of events.



- In a Streaming readout:
  - Data is read continuously from all channels.
  - Validation checks at source reject noise and suppress empty channels.
  - The data then flows unimpeded in parallel channels to storage or a local compute resource.
  - Data flow is controlled at source.
  - Data is organized in multiple dimensions by channel and time.

Streaming advantages

- The lack of a trigger means that:
  - Potentially useful physics is not discarded.
  - Run groups of experiments in parallel.
  - The system is simplified.
  - Readout speed is independent of detector response time.
  - Flow control at data source not via backpressure.
- Parallel timestamped streams mean:
  - System is robust against minor hardware or firmware glitches.
  - Can use different analysis techniques such as looking for hit patterns rather than reconstructing events.
- Requires robust and accurate time stamp generation and distribution.
  - Is still a simpler task than an online trigger.

Jefferson Lab

What does streaming readout solve?

- TDIS
  - High data rate from TPC handled as parallel streams – no one stream handles the full rate.
  - Electron data is it’s own stream and is matched up with proton candidates offline – removes high rate event building issue.
- SoLID
  - High rates from GEMs handled in parallel – no event building.
  - Sector edge effects and correlation with other detectors is handled in software offline.
- EIC
  - All of the above.
  - New data analysis modes – for example looking statistically at hit distributions instead of event-by-event analysis.
    - Analogy – crystallography looks at diffraction patterns rather than ray tracing every x-ray through a crystal.

State of the art – sPHENIX (Hi Martin)

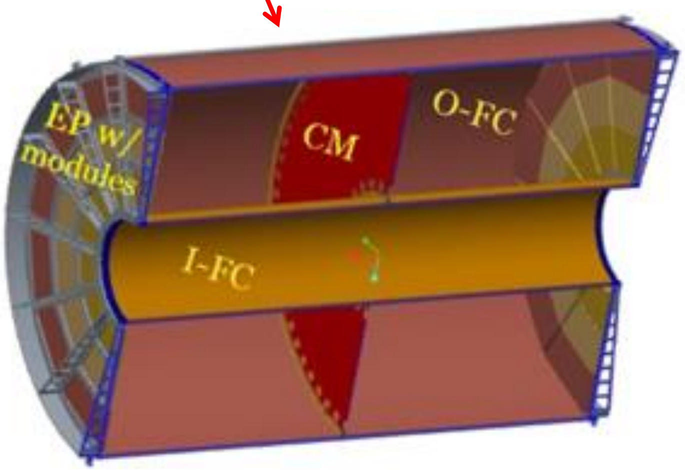
Streaming Readout Detectors in sPHENIX

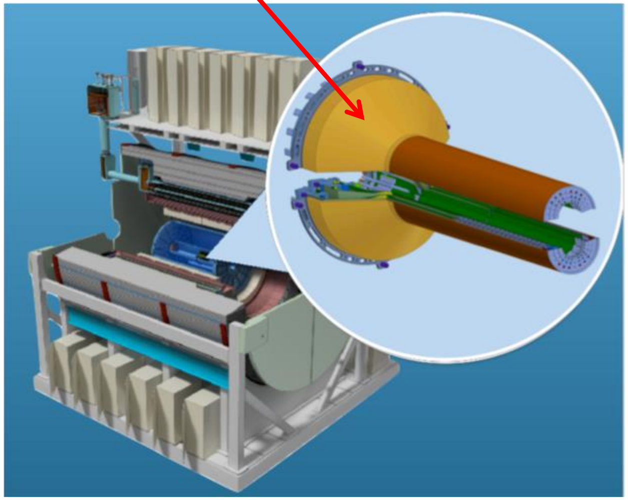
TPC

MVTX

No final design yet. Could be triggered or streaming readout. Time will tell.

I will mostly concentrate on the TPC.





This is a next-gen "gate-less" TPC design

State of the art – sPHENIX (Hi Martin)

TPC Readout Hardware rundown

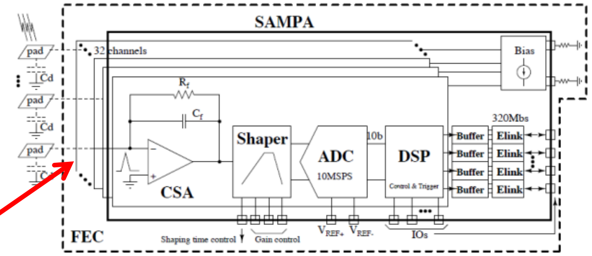


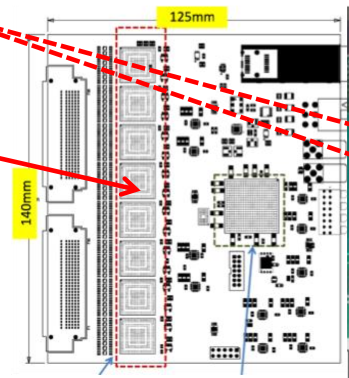
Figure 6.4: Schematic of the SAMPA ASIC for the GEM TPC readout, showing the main building blocks.

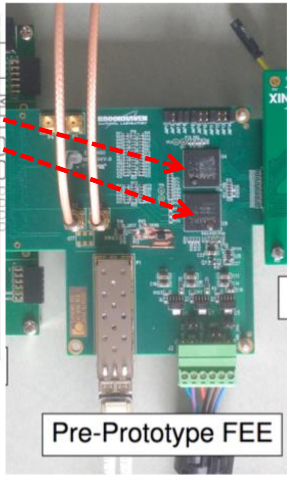
ALICE SAMPA ASIC – 32 channels digitizer, 10Ms/s, 10Bit

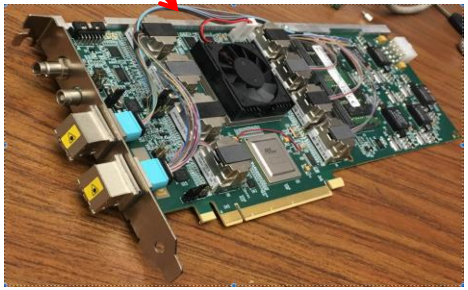
Front-end board (FEE) with 8 SAMPAs
256 channels (~ 500 pcs)

FEEs are read out with ATLAS FELIX cards (Data Aggregation Module, DAM)-> Data end up in a PC

That PC is called EBDC
("Event Buffering and Data Compressor")



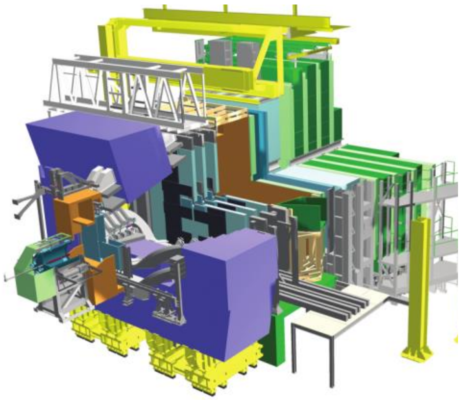




State of the art - LHCb

JINST 8 (2013) P04022

Real-Time Processing



Simple feature-building, e.g. in FPGAs, required to reduce the data rate.*

1 TB/s
post zero suppression

50 GB/s

1 MHz

Heavy use of machine learning:
V.Gligorov, MW, JINST 8 (2012) P02013.

Real-time reconstruction for all charged particles with $p_T > 0.5$GeV.

8 GB/s

Data buffered on 10 PB of disk.

Real-time calibration & alignment.

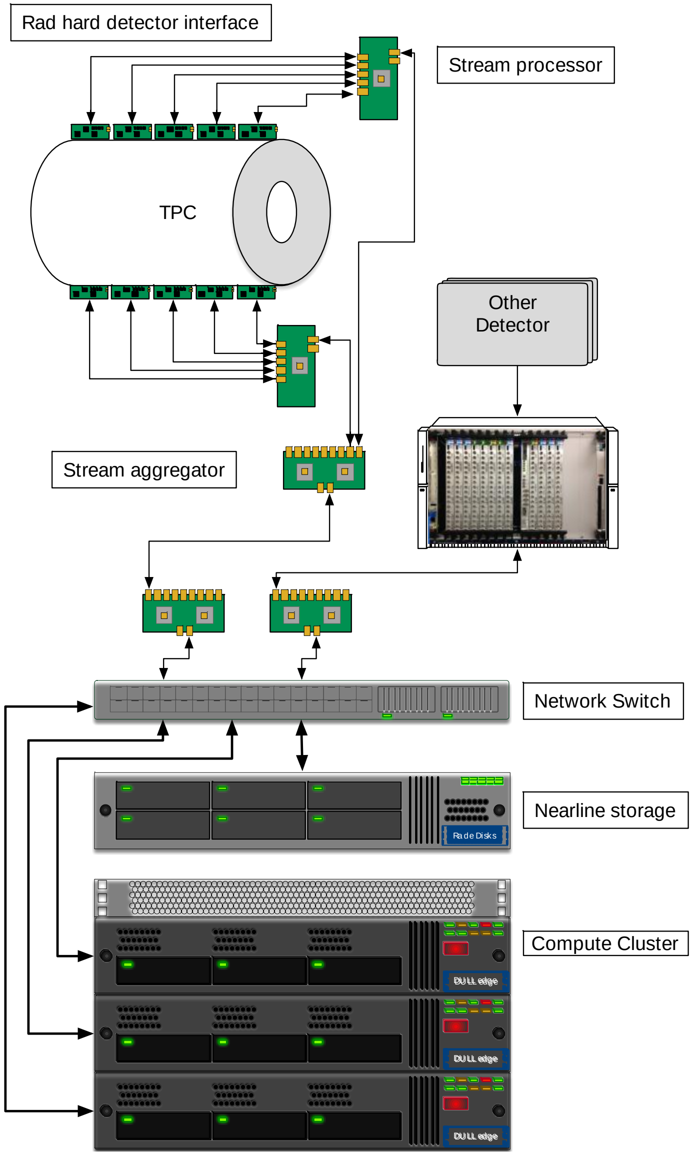
Full real-time reconstruction for all particles available to select events.

0.7 GB/s (mix of full + partial events)

*LHCb will move to a triggerless-readout system for LHC Run 3 (2021-2023), and process 5 TB/s in real time on the CPU farm.

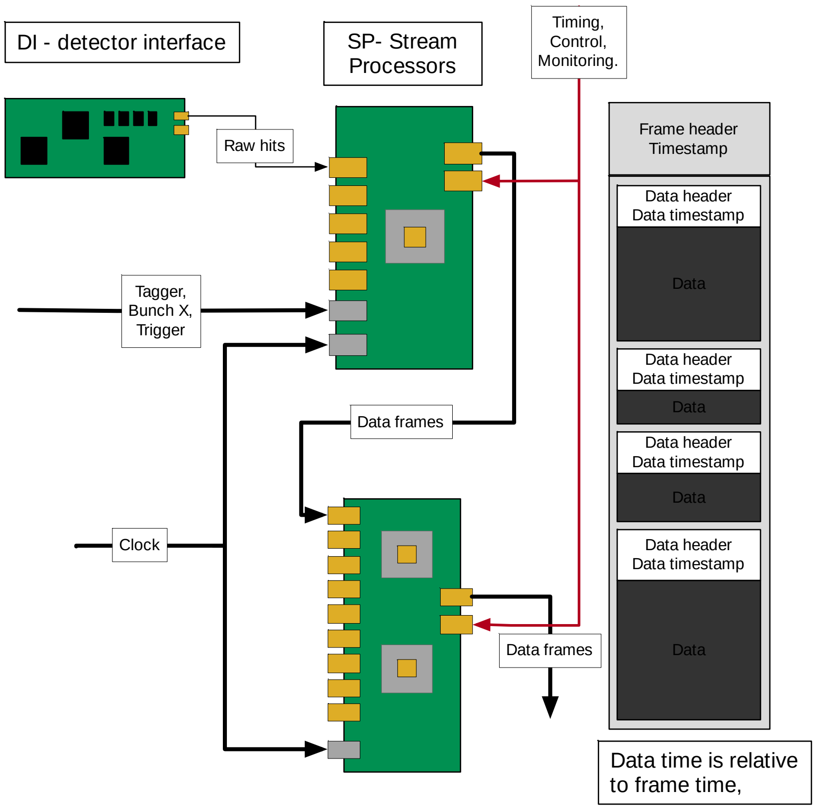
Generic stream based readout

- Rad hard detector interfaces convert detector signals into serial stream.
- Stream processor – “smart” FPGA and maybe CPU equipped electronics processes the raw stream to reduce data rate.
- Stream aggregator - since there will be many streams need to reduce the number of cables
  - Note : this is NOT event building, this is pushing more than one stream down a single fiber.
- Modularity allows growth from small to large systems.
  - If the fiber protocol is a network standard like Ethernet we connect directly to a switch.
  - If not then we will provide a PCI card to put in the storage node(s).



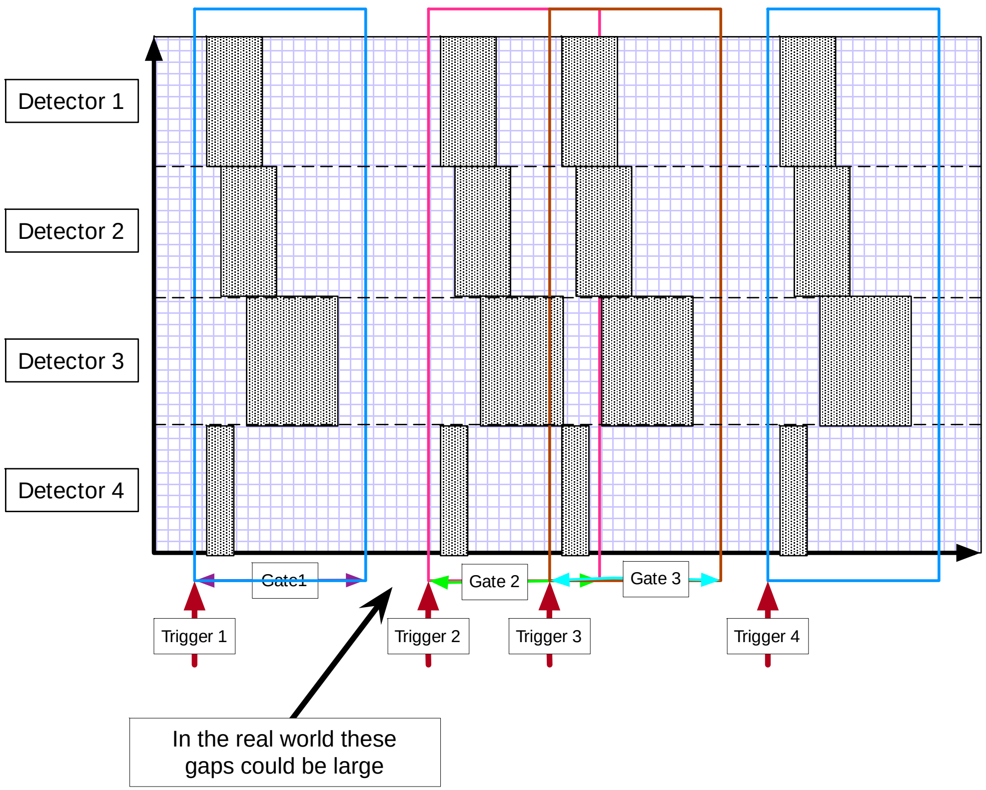
Concept- hardware

- Raw data from the detector is continuously streamed
- A very basic trigger is be used to reject time periods unassociated with events.
- The SP can reject noise and “zero suppress”.
- Data packets are packed into an efficient standard format and sent over a standard hardware interface.
  - If all components use the same hardware interface and format we have a “plug and play” system!!
- Individual components should be cheap!
  - One of out VXS crates with support boards such as CPU and trigger cards costs ~$90k!
- Stream via a CoTS standard like Ethernet?



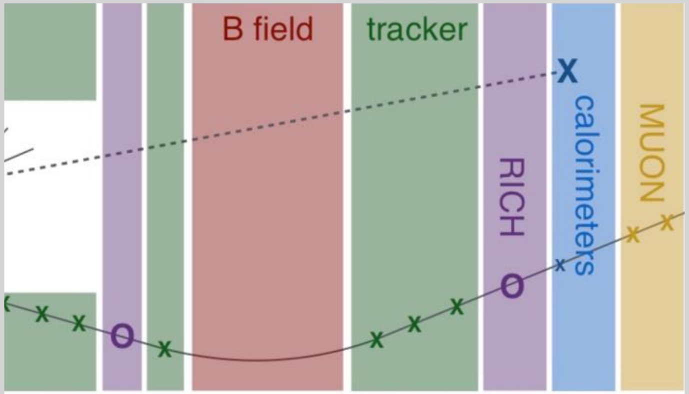
Concept - software

- Packets containing data are stored indexed by time and channel.
- How to store and process data in this form? We are used to 1D files of events, this is a 2D structure indexed by detector and time.
- Reconstruction software requests data for a particular channel based on a time window where the data associated with an interaction is likely to be found.
- In this example triggers 2 and 3 close together.
  - The request for trigger 3 gets part of trigger 2 data from detector 3.
  - In this case it is resolved because that data appears too early in the region of interest to be from the trigger 3 event.
- Note that this does not represent a layout in storage. Areas where there is no data would be compressed.



Integrated readout and detector design

- In many experiments someone builds a detector and someone else has to figure out how to read it out and integrate it into the DAQ.
- With a streaming system this is, in principle, simpler because each detector generates one, or more independent data streams in a standard format – plug and play.
- It is still easy to fall into the trap of adding bottlenecks by having many parallel streams but the bulk of the data generated in only a few very high bandwidth streams.
- Avoid generating data that is costly to process:
  - Tracking in weird magnetic fields.
  - Data packed in a convoluted way.
  - Segmented detectors that have a lot of overlap between sectors.



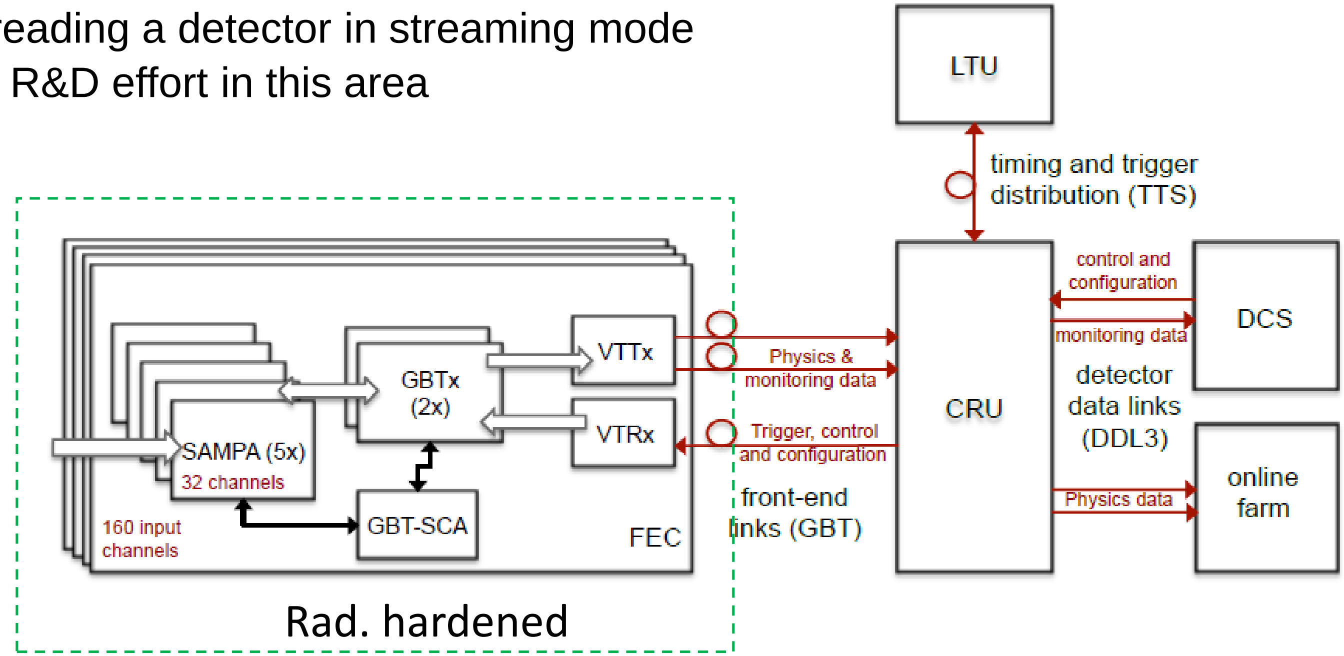
Here tracking is done outside the field to reduce processing cost.

Where do we go from here

- JLab has several proposals, TDIS and SoLID are examples, that would benefit from streaming mode readout.
- The technology, both hardware and software is available and affordable to allow us to move to this mode.
  - Proven by LHC readout and adopted at other labs.
  - Simpler and more robust than the current readout mode.
  - Other advantages over pipeline readout already mentioned.
- We will develop a test stand to investigate techniques and technologies.
- We will develop a way to integrate legacy hardware – CLAS12?
- The goal is to use this mode for TDIS.
- If pilot system is successful:
  - Investigate moving CLAS12 and GlueX to this mode.
  - Use for SoLID and future JLab experiments.
    - I hope we may see new or modified proposals based on this.
  - Use for EIC?

JLab TDIS TPC test stand using SAMPA

- **Fast Track test stand**– use as many components of the ALICE TPC readout/control chain as possible
- Target test stand operational July 1st 2018
  - Validate use of SAMPA for TDIS
  - Experience reading a detector in streaming mode
  - Guide future R&D effort in this area



FEC – Front End Card

CRU – Common Readout Unit

DCS – Detector Control System

LTU – Local Trigger Unit

JLab – DAQ and Electronics R&D

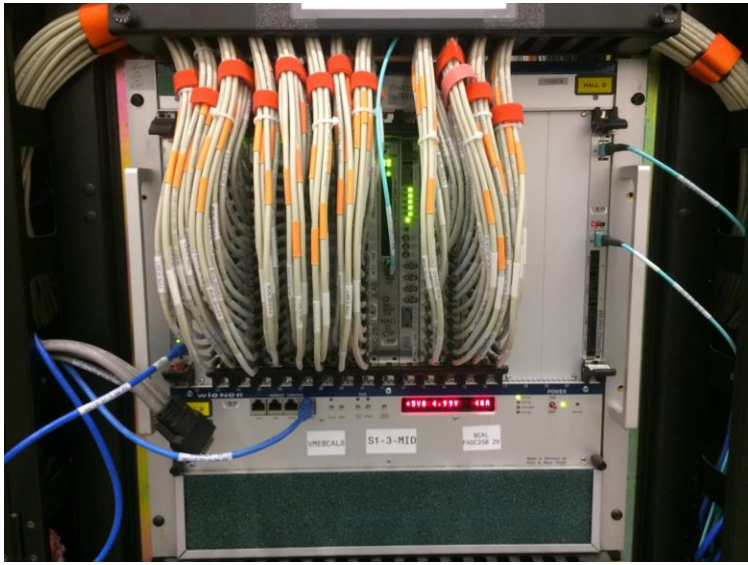
- The VTP board is able to read 250 MHz fADC via the VXS serial lines and stream the data out over the front panel transceivers.

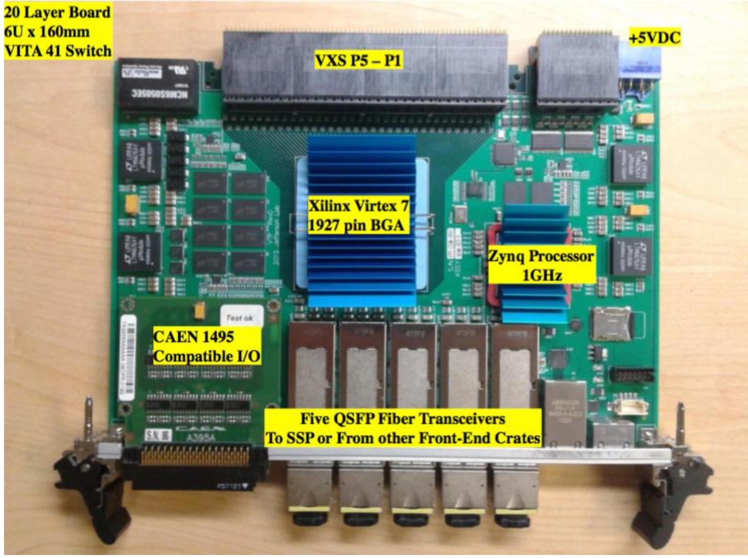
a) Read a crate in this mode.
   - Zero suppression logic
   - Flow control
   - What comes out on the front panel fiber?
   - How to interface with DAQ?

b) Read out some detectors in streaming mode!

c) Integrate with TDIS test-stand.
   - Can we replace the CRU etc?

d) Can we repackage to remove need for expensive VXS crates etc.
   - Cheap system for university groups!





EIC focus group

- Two streaming readout workshops, last one January 2018
- Virtual meetings regularly and also discuss via a mailing list :
  https://mailman.mit.edu/mailman/listinfo/eic_streaming_readout
- Subgroups looking at :
  - Pilot projects
  - Electronics
  - Software
  - Protocols
  - Detectors

EIC Streaming Readout Consortium

- *Catholic University of America - S. Ali, V. Berdnikov, T. Horn, I. Pegg, R. Trotta*
- *INFN - M. Battaglieri (Co-PI), A. Celentano*
- *MIT - J.C. Bernauer* (Co-PI), D.K. Hasell, R. Milner*
  ** Also Stony Brook University, Stony Brook, NY*
- *Jefferson Lab - C. Cuevas, M. Diefenthaler, R. Ent, G. Heyes, B. Raydo, R. Yoshida*
- Proposal for streaming readout R&D targeted at EIC
  - Deliverable is a document/publication containing:
    - Reports on the relevant aspects of the performance of the prototypes studied.
    - A list and definition of streaming-readout parameters relevant to the EIC.
    - Initial estimates of some of the parameters with current technology as well as extrapolations to the time period of EIC detector construction phase.
    - A list and definitions of relevant parameters for detector technologies (e.g. TPC, Crystal Calorimeter..etc.) when considering streaming readout.
    - Initial estimates of some of these parameters.
  - Projects
    - TPC in streaming-readout at JLab
    - Crateless-Streaming at JLab
    - FEE and Circuit Designs for Streaming Readout at MIT
    - Streaming readout for an EIC Calorimeter at INFN Genoa and CUA
    - Multilayered Architecture for Streaming Readout at Stony Brook
    - Growing list…

Summary

- There are clear advantages to be gained by transitioning to a streaming mode of readout for 12 GeV CEBAF era detectors at Jefferson Lab.
- Our plan is to begin with hardware already installed for detectors such as CLAS12 and gain experience.
- In parallel we are looking forward to:
  - Proposals such as TDIS and SoLID that we can instrument for streaming mode readout from the start.
  - EIC focused R&D and collaborations to generate standards for such systems and provide input for EIC detector design.
- We hope that by forward looking design we can influence new physics proposals by expanding the view of what is possible and enabling different techniques for data analysis.