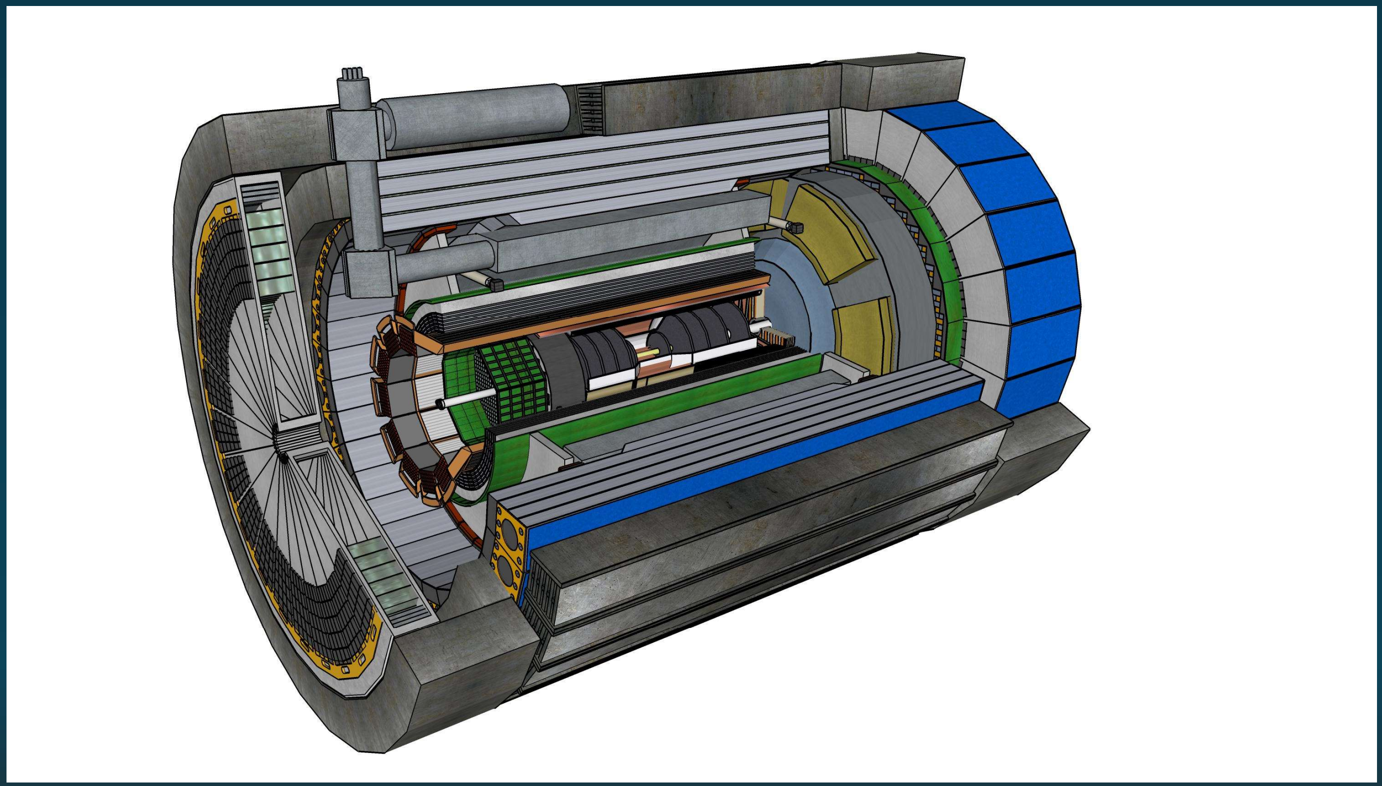
Streaming Readout for the ePIC DAQ

ePIC Electronics & DAQ WG Conveners: Fernando Barbosa, Jin Huang, Jeff Landgraf



Outline

- Introduction to the EIC and to the ePIC detector
- Streaming in the context of the ePIC DAQ
- Overview of the ePIC DAQ system components
- Channel counts and data volume expectations
- Status of the development
  - ASICs
  - DAQ system components
  - DAQ computing
- Connection to ePIC Computing
- Summary

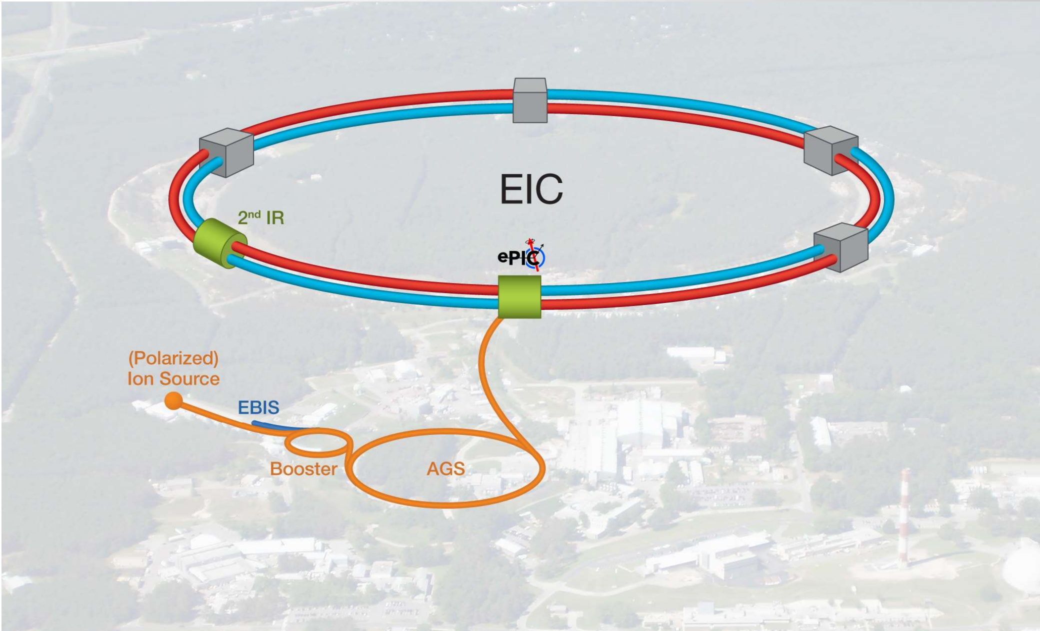
The Electron Ion Collider

- To be built at Brookhaven National Laboratory on Long Island, NY in partnership with JLAB
- CD-3A review just held, completion scheduled for early 2030s

- Collides electron/ion beams and polarized electron/proton beams
  - Electron beam energy up to 18 GeV
  - Proton beam energy 41-275 GeV
  - Ion beam energy 41-110 GeV / nucleon

- 1160 Bunches arriving at 98.5MHz
  - Maximum interaction rates ~500kHz
  - Expect low multiplicity events

- Key DAQ requirements
  - High Rates: ~500kHz + backgrounds
  - Low Noise
  - Resolve collisions to bunch to understand Polarization

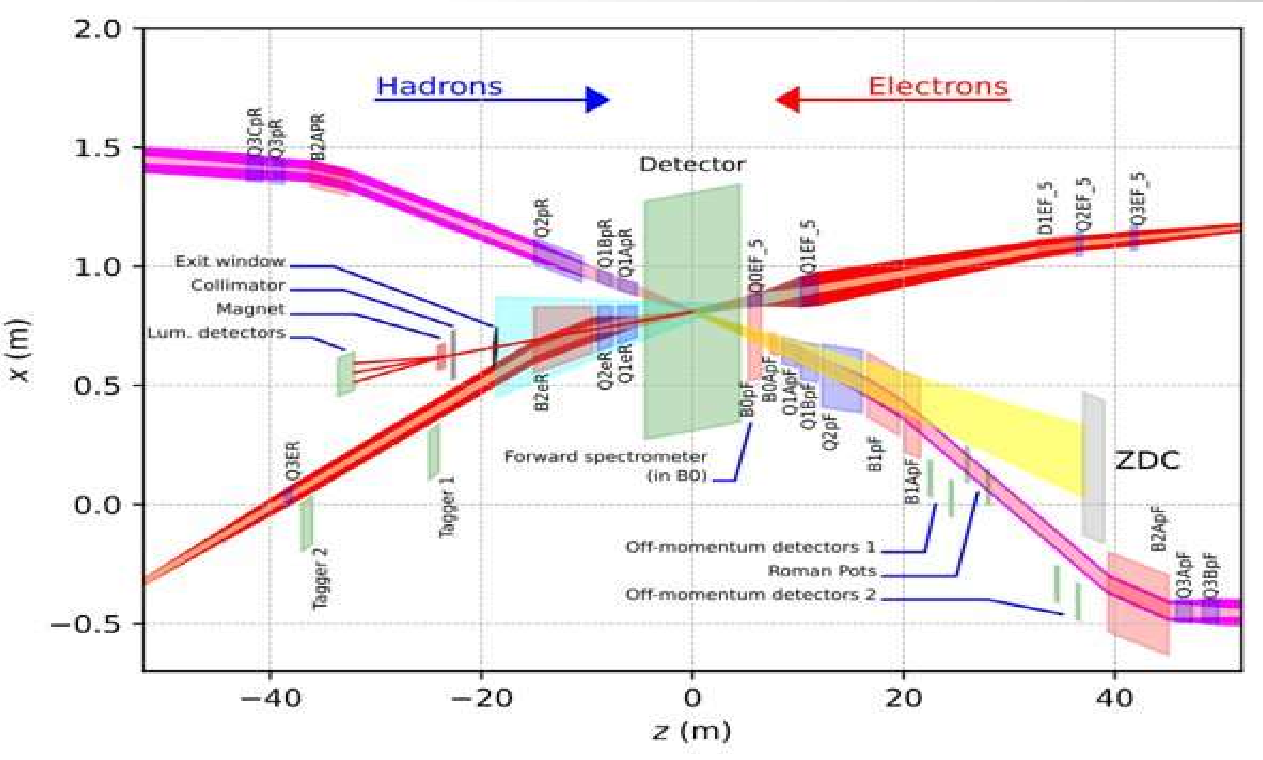


The ePIC Detector

- ePIC is a large international collaboration and works in partnership with the EIC project (6.10 L2 Control Account)
  - 171 institutions, 24 countries, 500+ participants

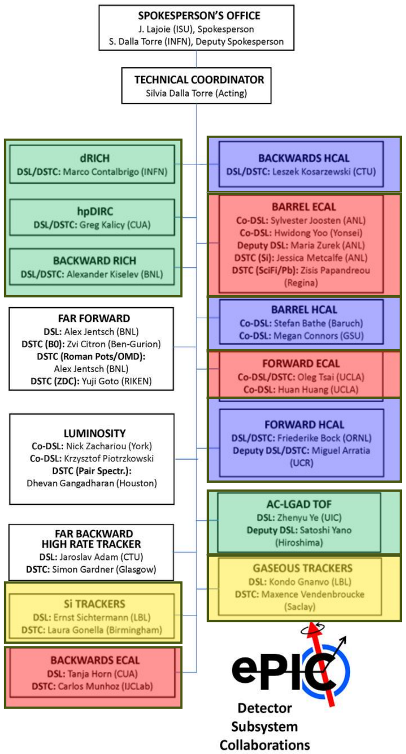
- 24 Detectors with multiple technologies
  - MAPS (~16 Billion Pixel)
  - AC-LGAD (~10 Million Channels)
  - SiPM (~1M channels)
  - HRPPDs (~70k Channels)
  - MCP-PMT (~70k channels)

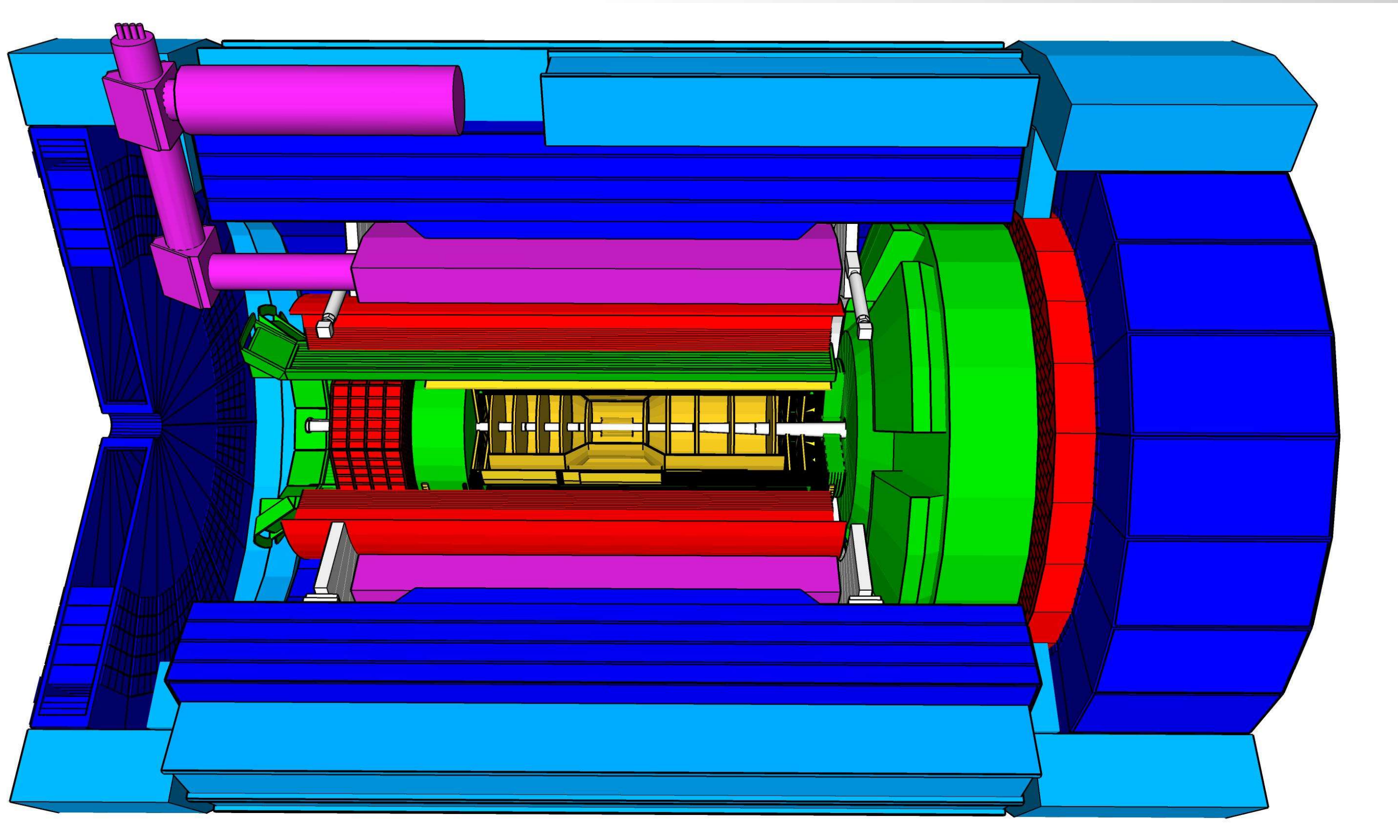
- Detector spans about 80m
  - Central detector ~4m
  - Far backward detectors
  - Far forward detectors



The ePIC Detector

- Central detector has 4π, but asymmetric coverage for
  - Hadronic and EM Calorimetry
  - Tracking
  - PID





Streaming in ePIC

ePIC DAQ

Definition of streaming is "No L0 trigger"

- No system wide deadtime in normal operation
  - Detectors ***will*** have minimum double hit times for channels or modules, as well as throughput limitations.
- Collaboration should have the full ability to make data selection cuts on the widest possible criteria
  - Full flexibility for event selection
  - As full flexibility for data selection as possible
  - As full background characterization as possible
- But subject to an overall throughput budget of ~100Gb/sec

ePIC Streaming will include

- Capabilities for software triggering
- Capabilities for hardware triggering
- Capability for flow control
- Zero suppression & aggregation within data packets

ePIC Computing

Definition of streaming is "Process data as it arrives"

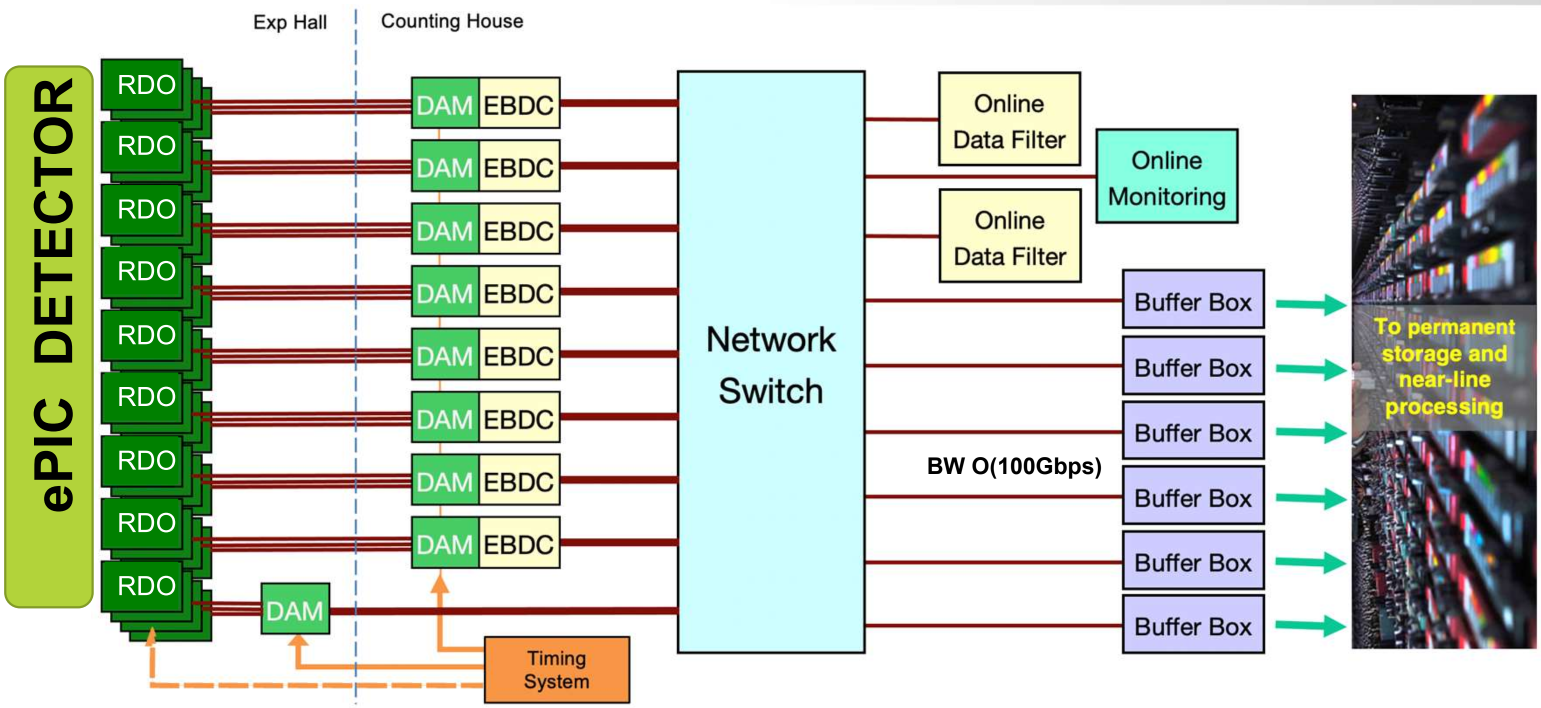
- Physics event selection and tagging
- Fast Analysis (~3 weeks not months or years) using automation of calibration and reconstruction.
  - Fast understanding of operations needs
  - Fast understanding of calibration
  - Fast publication
- Distributed analysis
- Efficient use of diverse architectures (eg. Support for GPU)
- Efficient use of diverse software (eg. AI)
- Incorporate worldwide computing facility contributions

Overlap between DAQ and computing

- Automation of Calibrations
- QA and monitoring
- Consistent schemes and language for managing data and metadata
- Event Selection and Tagging

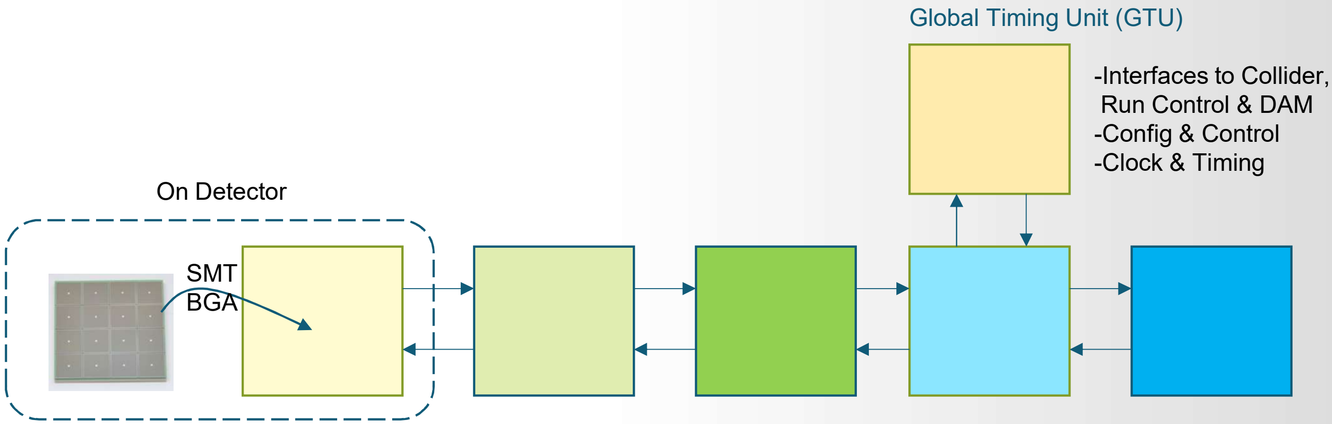
EIC Streaming DAQ/Computing Architecture

Bunch Crossing ~ 10.2 ns/98.5 MHz
Interaction Rate ~ 2 us/500 kHz
Low occupancy



GTU/Distributed clock
(jitter ~5ps)

ePIC Readout Chain

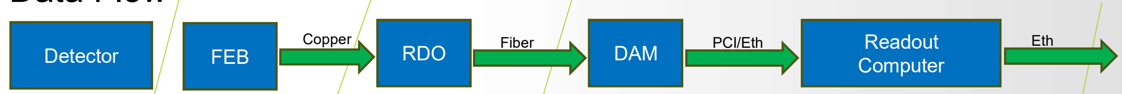


Name	**Sensor**	**Adapter**	**Front End Board (FEB)**	**Readout Board (RDO)**	**Data Aggregation Module (DAM)**	**Computing**
Design	Detector Specific	Detector Specific	Detector Specific	Few Variants	Common	Common
Function	Multi-Channel Sensor	-HV/Bias Distribution -HV divider -Interconnect Routing	-Amplification -Shaping -Digitization -Zero Suppression -Bias Control & Monitoring	-Communication -Aggregation -Formatting -Data Readout -Config & Control -Clock & Timing	-Computing Interface -Aggregation -Software Trigger -Config & Control -Clock & Timing	-Data Buffering and Sinking -Calibration Support -QA/Scalers -Collider Feedback -Event ID/Building -Software Trigger -Monitoring
Attributes	MAPS, AC-LGAD MPGD, MCP-PMT SiPM, LAPPD	-Sensor Specific -Passive	-ASIC/ADC -Discrete -Serial Link	-FPGA -Fiber Link	-Large FPGA -PCIe -Ethernet	-COTS -Ethernet

Summary of Channel Counts and Data Flow

Detector Group	Channels					RDO	Fiber	DAM	Data Volume (RDO) (Gb/s)	Data Volume (To Tape) (Gb/s)
	MAPS	AC-LGAD	SiPM/PMT	MPGD	HRPPD					
Tracking (MAPS)	16B					400	800	17	26	26
Tracking (MPGD)				202k		118	236	5	1	1
Calorimeters	500M		104k			451	1132	19	502	28
Far Forward	300M	2.6M	170k			178	492	8	15	8
Far Backward	82M		2k			50	100	4	150	1
PID (TOF)		7.8M				500	1500	17	31	1
PID Cherenkov			320k		140k	1283	2566	30	1275	32
TOTAL	16.9B	10.4M	596k	202k	140k	2980	6826	100	2,000	96

Summary of Data Flow



100% Occupancy	27.5-1760 Pb/sec

Aggregate	2.0 Tb/sec
Noise	1.6 Tb /sec
Signal from Physics + Background	400 Gb / sec

Aggregate	2.0 Tb/sec
Per RDO (Avg)	0.7 Gb/sec

Aggregate	96 Gb/sec
Collision Signal	38 Gb/sec
Synchrotron Rad	.01 Gb/sec
Electron Beam	22 Gb/sec
Hadron Beam	4 Gb/sec
Noise	32 Gb/sec

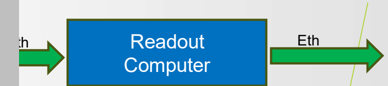
Summary of Channel Counts and Data Flow

	Fiber	DAM	Data Volume (RDO) (Gb/s)	Data Volume (To Tape) (Gb/s)
Tracl	800	17	26	26
Tracl	236	5	1	1
Calo	1132	19	502	28
Far F	492	8	15	8
Far E	100	4	150	1
PID (	1500	17	31	1
PID (	2566	30	1275	32
TOTA	6826	100	2,000	96

Significant Filtering Challenges:

- Electron Bremsstrahlung in Far Backward
  - Bremsstrahlung will produce up to 18 particles / Bunch Crossing.
  - Need to produce luminosity information
- dRICH – SiPMs requiring single photon sensitivity
  - Dark currents increase with radiation damage
  - Expect several years before annealing necessary to reduce dark currents
- After several years in a radiation environment, “noise” dominates signal at the RDO by factors up to x4
- Main Strategy for dRICH/Far Backward
  - Supply enough bandwidth to account for maximum data volume to the DAM boards
  - Apply cross-detector correlation filter in nearby detectors in DAM / Readout computers to reduce recorded data volume

Su



100% Occupancy	27.5-1760 Pb/sec

Aggregate	2.0 Tb/sec
Noise	1.6 Tb /sec
Signal from Physics + Background	400 Gb / sec

Aggregate	2.0 Tb/sec
Per RDO (Avg)	.7 Gb/sec

Aggregate	96 Gb/sec
Collision Signal	38 Gb/sec
Synchrotron Rad	.01 Gb/sec
Electron Beam	22 Gb/sec
Hadron Beam	4 Gb/sec
Noise	32 Gb/sec

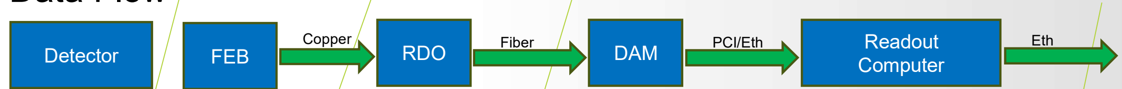
Summary of Channel Counts and Data Flow

Detector Group	Channels					RDO	Fiber	DAM	Data Volume (RDO) (Gb/s)	Data Volume (To Tape) (Gb/s)
Tracking (MAPS)								17	26	26
Tracking (MPGD)								5	1	1
Calorimeters								19	502	28
Far Forward								8	15	8
Far Backward								4	150	1
PID (TOF)								17	31	1
PID Cherenkov								30	1275	32
TOTAL	16.9B	10.4M	596k	202k	140k	2980	6826	100	2,000	96

Data Filtering

- Cluster finding accounts for the balance of the calculated noise reduction, and for the reduced signal data volume.
- Additional methods are being considered including AI/ML techniques for pattern recognition and/or data compression
- These functions will be performed in the DAM boards and the Online computing farm

Summary of Data Flow



100% Occupancy	27.5-1760 Pb/sec

Aggregate	2.0 Tb/sec
Noise	1.6 Tb /sec
Signal from Physics + Background	400 Gb / sec

Aggregate	2.0 Tb/sec
Per RDO (Avg)	.7 Gb/sec

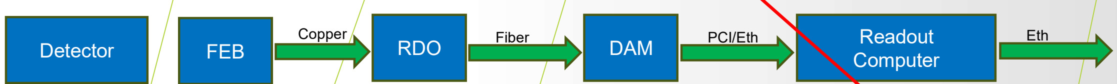
Aggregate	96 Gb/sec
Collision Signal	38 Gb/sec
Synchrotron Rad	.01 Gb/sec
Electron Beam	22 Gb/sec
Hadron Beam	4 Gb/sec
Noise	32 Gb/sec

Summary of Channel Counts and Data Flow

Detector Group	Channels					RDO	Fiber	DAM	Data Volume (RDO) (Gb/s)	Data Volume (To Tape) (Gb/s)
	MAPS	AC-LGAD	SiPM/PMT	MPGD	HRPPD					
Tracking (MAPS)	16B					400	800	17	26	26
Tracking (MPGD)				202k		118	236	5	1	1
Calorimeters	500M		104k			451	1132	19	502	28
Far Forward	300M	2.6M	170k			178	492	8	15	8
Far Backward	82M		2k			50	100	4	150	1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	00	1500	17	31	1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	83	2566	30	1275	32
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	80	6826	100	2,000	96

After online processing we expect roughly equal contributions from collision hits, beam background hits, and noise.

Summary of Data Flow



100% Occupancy	27.5-1760 Pb/sec

Aggregate	**2.0 Tb/sec**
Noise	1.6 Tb /sec
Signal from Physics + Background	400 Gb / sec

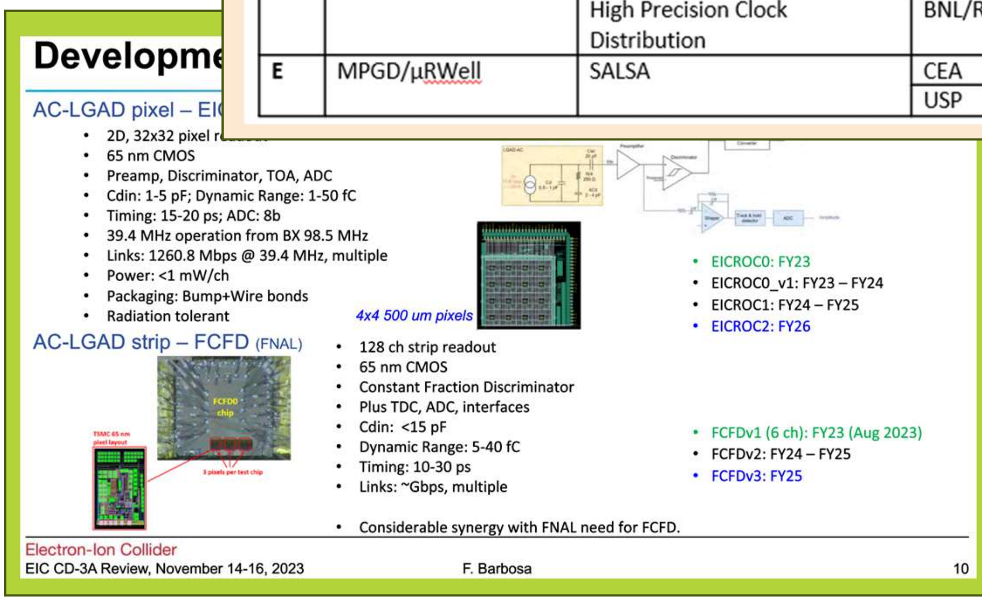
Aggregate	**2.0 Tb/sec**
Per RDO (Avg)	.7 Gb/sec

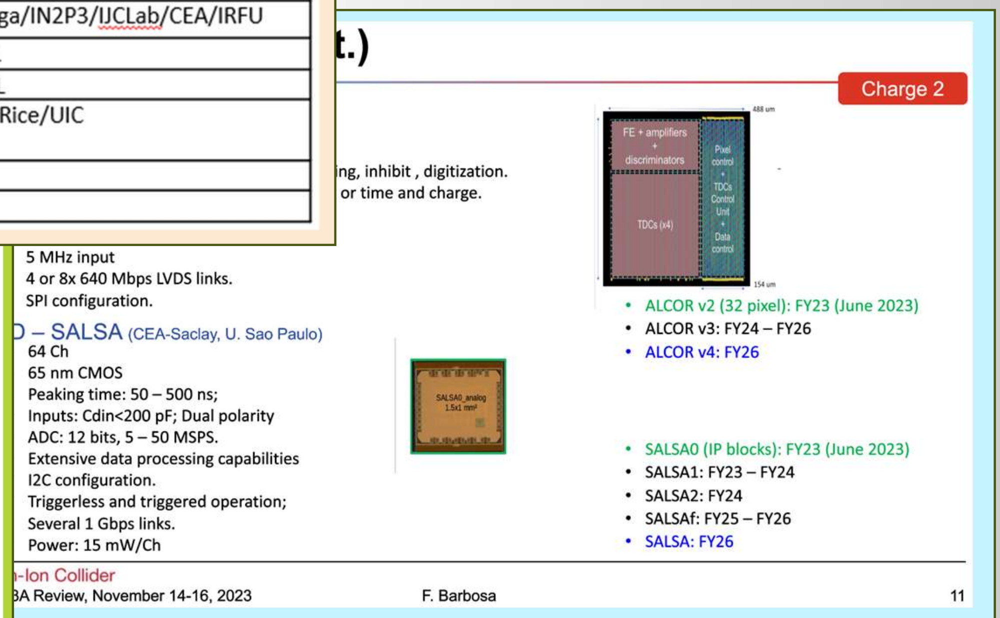
Aggregate	**96 Gb/sec**
Collision Signal	38 Gb/sec
Synchrotron Rad	.01 Gb/sec
Electron Beam	22 Gb/sec
Hadron Beam	4 Gb/sec
Noise	32 Gb/sec

System Component Status: ASIC development

- ASIC development is funded through through eRD109 (F. Barbosa Saturday 4pm)

	Detector/Technology	Discrete/ASIC	Group
A	Calorimeter	Discrete	IUCF
B	Calorimeter	HGCROC	ORNL
C	dRICH	ALCOR	INFN – BO, TO
D	AC-LGAD	EICROC	Omega/IN2P3/IJCLab/CEA/IRFU
		FCFD	FNAL
		Barrel L-M Serv. Hybrid	ORNL
		High Precision Clock Distribution	BNL/Rice/UIC
E	MPGD/µRWell	SALSA	CEA
			USP





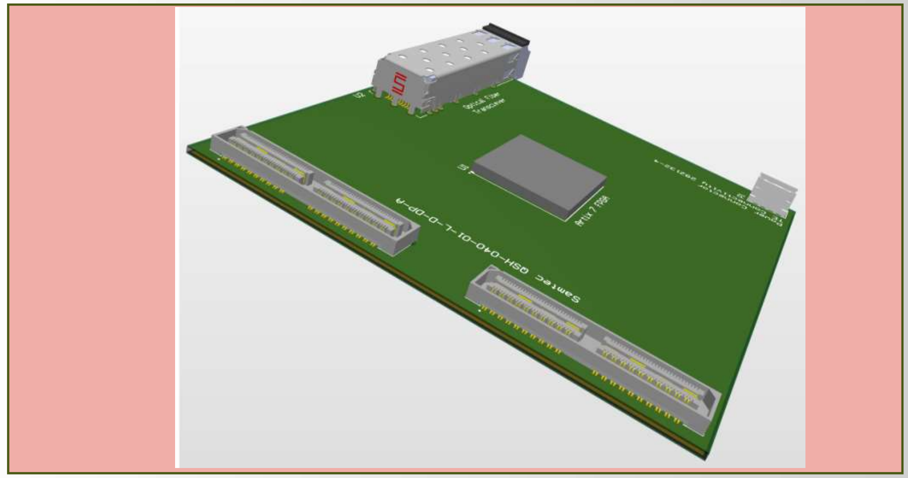
System Component Status: RDO and GTU

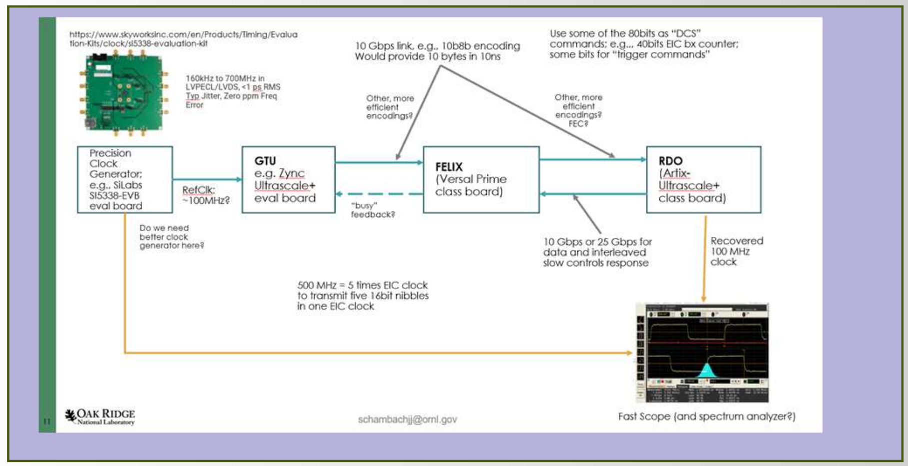
The hardware development of the RDO and GTU is currently focused on the development of the low-level timing protocol for achieving the 5ps time resolution requirement. We prefer using a reconstructed clock to avoid extra fiber

- Currently 3 options under evaluation
  - Custom protocol
  - GBT using fpga
  - Dedicated timing fiber
  - These options are being evaluated using FPGA devkits (See Jo Schambach's talk)

- Data volume studies need to be extended to the channel level in to study detailed rates needed by components

- Integration is also being discussed. Power, cooling, space, radiation needs, and cable lengths need to be fully specified and integrated for next year's TDR.
  - To this end we have provided a mocked up physical design and power estimate for the RDOs
  - Assume Artix Ultrascale+,2.5 x 2.5 inches for common components + connector space is our guidance, and 2-5 Watts power





System Component Status: DAM

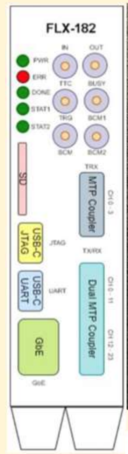
DAM Candidate

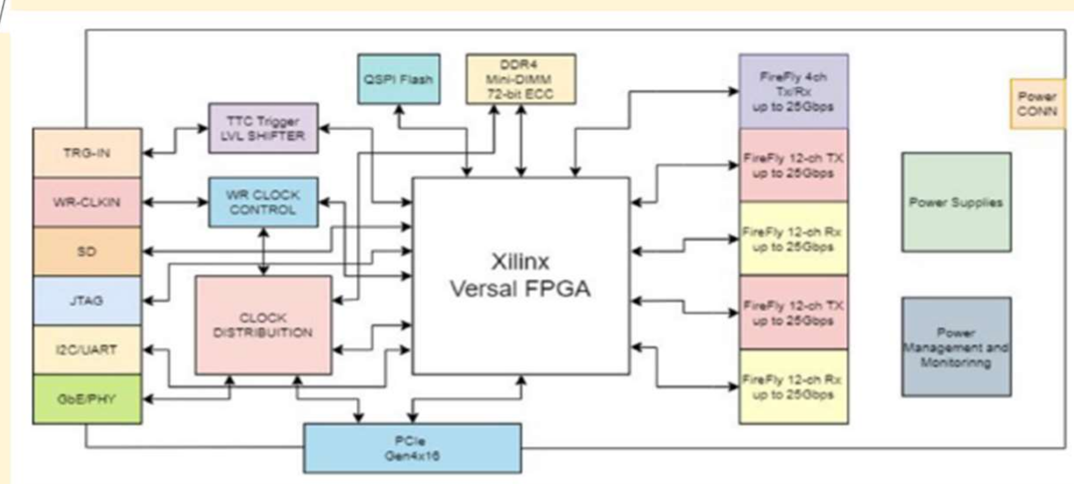
FELIX FLX-182 from Atlas / Omega group a BNL

FPGA: Xilinx Versal Prime
- PCIe Gen 4 x16: PL and CPM compatible
- 24 FireFly links with 3 possible configurations
    - 24 links@25 Gb/s
    - 24 links @10 Gb/s (CERN-B-Y12)
    - 12 links @25 Gb/s +12 links @10 Gb/s
- 4 Firefly Links with 2 possible configurations
    - LTI interface
    - 100GbE
- Electrical Signals on front panel
    - 3 input and 3 output
- 1 DDR Mini-UDIMM
- USB-JTAG/USB-UART
- Dual core ARM Cortex SoC
- Power usage is 133W external
- Can be implemented as stand alone device

We have these in hand and are getting familiar with the board

Next version has 48 link capability available in 2024

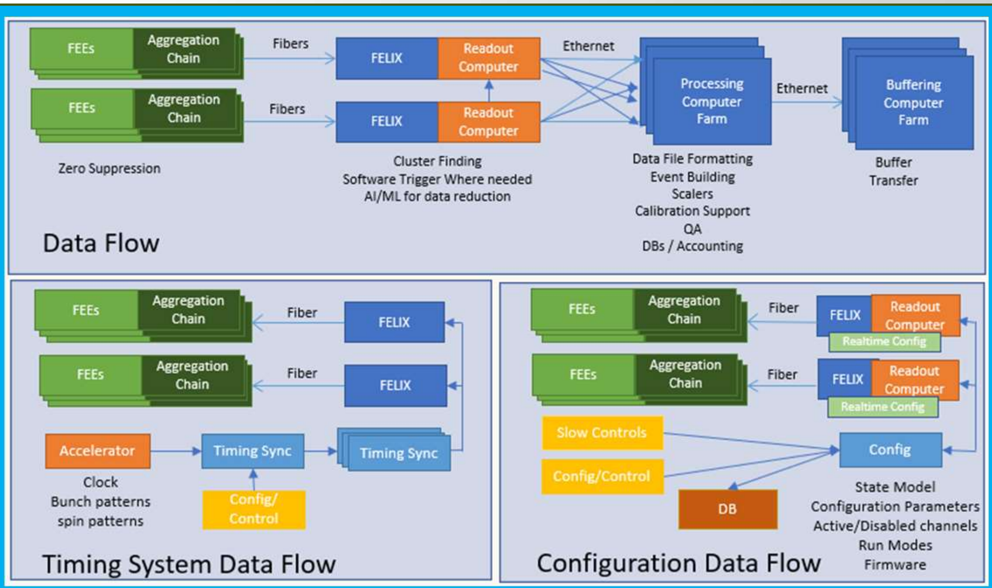




Optical Protocol Requirements

The High-Level Optical Protocol

- Defines the information content of the data flow between the GTU, RDO, and DAM boards
- Optical Protocol Requirements
  - Timing
    - <=5ps Resolution for hi-res Detectors
    - <=100ps for All detectors
    - Phase stability on power cycle
  - Configuration of RDO firmware and ASIC parameters
    - Mode for configuring RDO
    - Mode for ASIC configuration passthrough
  - Real-Time Command / Control
    - Define bunch crossing counter
    - Define time frames
    - Trigger Information
      - Mark or initiate calibration events
      - Potential debug or fallback triggered modes
    - Formatting information
    - Flow Control
      - RDOs need to mark potential dropped data
      - Need the capability of applying global busy
  - Slow Control Monitoring from ASICs/FEBs (e.g. Temperature monitoring)
  - Data Transfer (>=10Gb/s per RDO)

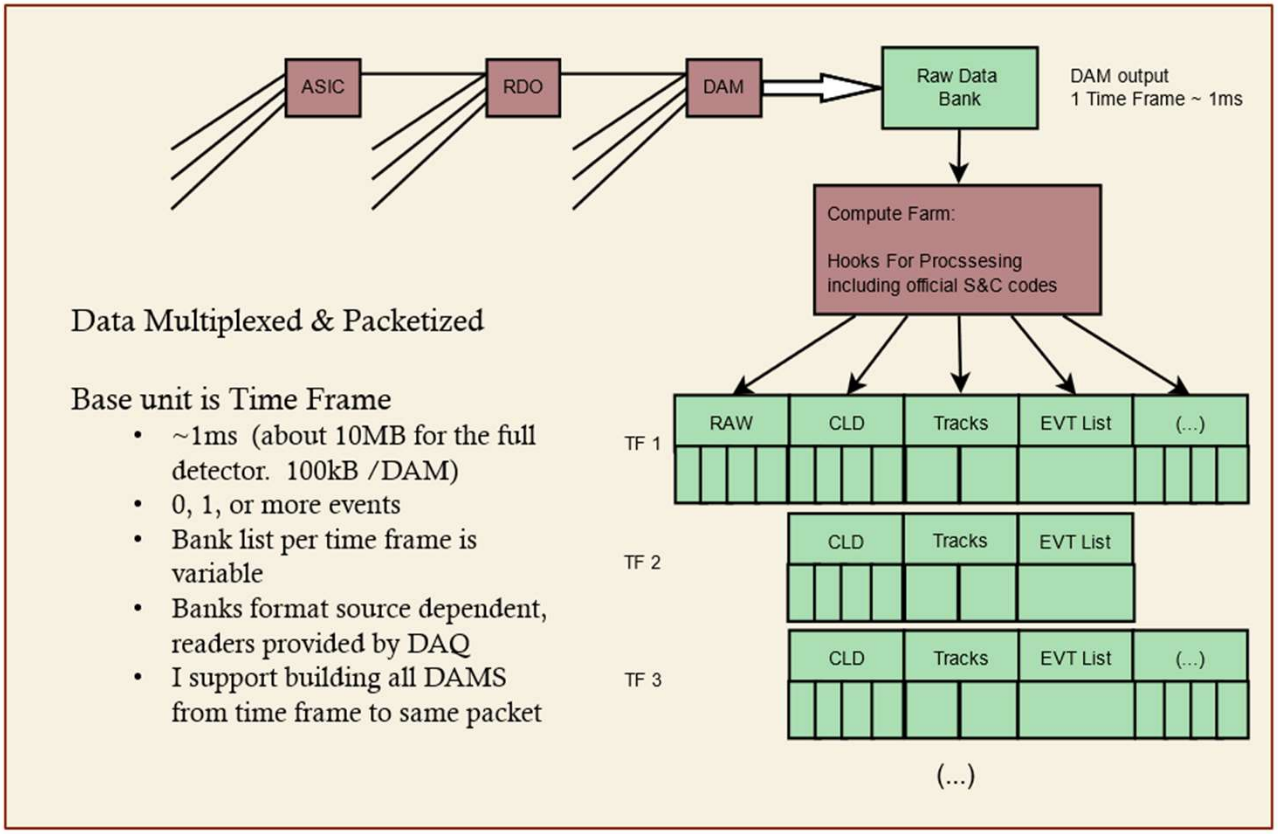


DAQ Local Computing: Time frames

- Time Frames (~1ms)
  - Up to ~500 events per Frame
  - ~10MB data per Frame
  - ~2.0MB input per DAM
  - ~100kB output per DAM

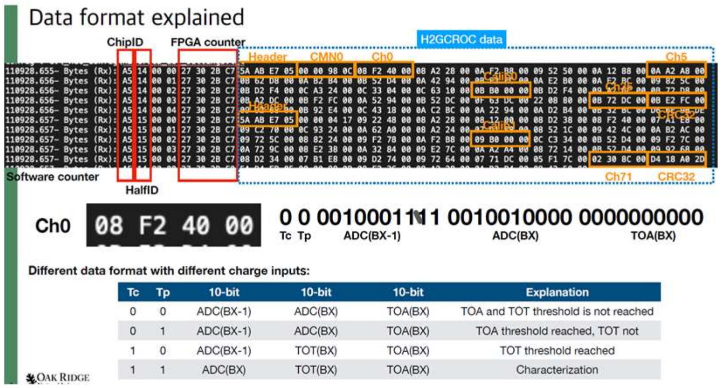
- Routing data
- Formatting data
- Processing data
  - Cluster finding
  - Software triggering
  - Sanity Checkers
  - QA Monitoring
  - Metadata
  - Slow controls information
  - Building time frames

- Scalers / continuously running DAQ components



DAQ Computing: Data Format Case Study (H2GCROC3A)

This is the data format of a prototype ASIC – not the final ASIC, my points are conceptual! (N. Nowitski, Sat 4:30pm)



Tc	Tp	10-bit	10-bit	10-bit	Explanation
0	0	ADC(BX-1)	ADC(BX)	TOA(BX)	TOA and TOT threshold is not reached
0	1	ADC(BX-1)	ADC(BX)	TOA(BX)	TOA threshold reached, TOT not
1	0	ADC(BX-1)	TOT(BX)	TOA(BX)	TOT threshold reached
1	1	ADC(BX)	TOT(BX)	TOA(BX)	Characterization

- Granularity of readout is half chip (36 channels).
- Data is 3x10 bit words per sample, but the meaning of the three words depends on the Tc, Tp flags
- Data is NOT pedestal subtracted, rather the pedestals are shifted via calibration to the same value
- Can configure fixed # time samples read out
- Can configure to disable certain Tc/Tp Combinations

Potential interesting data:

information ↑ / Data volume ↓

- Full ASIC data (despite containing up to 35 channels of data without hits) {RAW}
- Channels without hits removed, but containing n samples for channels hit {CHANNEL}
- Channels with samples suppressed and only TOA/TOT information {TOATOT}
- Processed TOA/TOT information improved using sample information, but without saving samples {ENHANCED}

➡ Suppress Banks in some timeframes:

/TF#/NHCAL/DAM01/RAW	(1 in 1000)
/TF#/NHCAL/DAM01/CHANNEL	(1 in 10)
/TF#/NHCAL/DAM01/TOATOT	(1 in 10)
/TF#/NHCAL/DAM01/ENHANCED	(always)

➡ Implications for data formats & readers

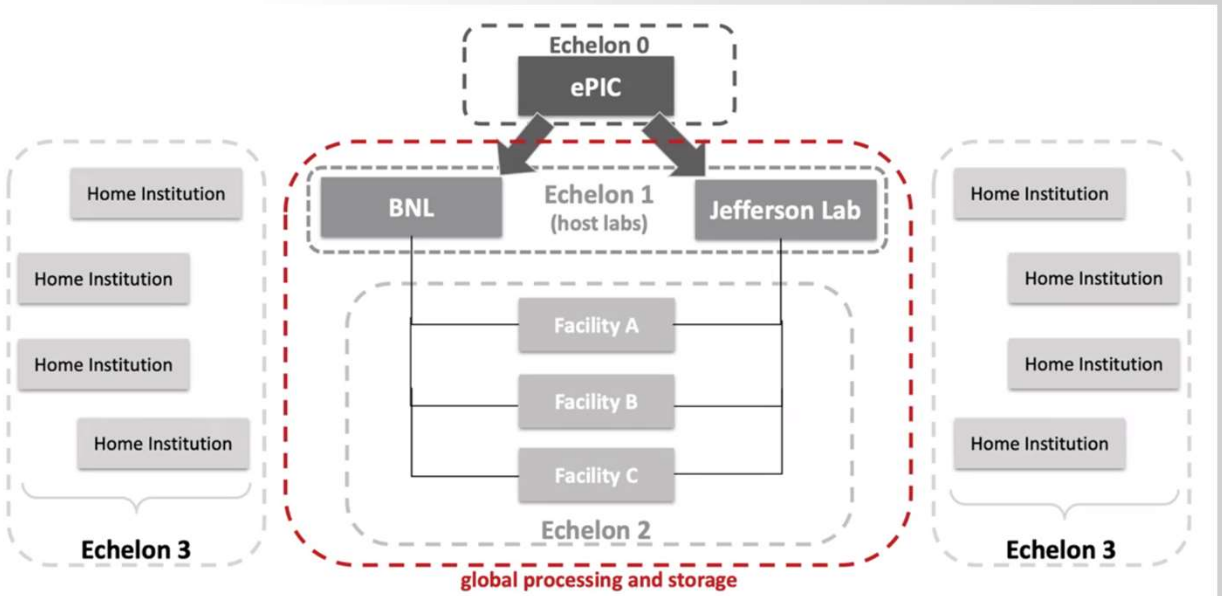
- ➢ Generic named data navigation
- ➢ Detailed bank reader plugins

Boundary between Online and Offline

- Streaming RO computing model talk(M. Diefenthaler, Saturday 11:30)
- The streaming architecture allows for some blurring of the offline and online processing (QA/ Monitoring/Calibrations)
- Require an interface to allow sharing of code between offline / online processing
- Local buffering (Echelon 0) provides elasticity in the transfer of data to computing facilities
- Offline buffering (Echelon 1) will allow several weeks for calibration and reconstruction.

ePIC Scientific Computing

- The Resource Review Board (RRB) looks at the international detector and computing resources.
- The ePIC computing groups are actively implementing scientific software as well as working with the DAQ to define the interface with the streaming DAQ system.



Summary

- We have given an overview of the planned ePIC DAQ system
- Described how the streaming DAQ model fits into the design of the ePIC DAQ
- Overview of the ePIC DAQ system components and the status of their development
- We are still in the early stages of constructing the DAQ system, if you are interested in contributing:

Mailing list: https://lists.bnl.gov/mailman/admin/eic-projdet-daq-l

Indico page: https://indico.bnl.gov/category/409/

Questions?

Back-up slides

Detector System	Channels	RDO	Gb/s (RDO)	Gb/s (Tape)	DAM Boards	Readout Technology	Notes
Si Tracking: 3 vertex layers, 2 sagitta layers, 5 backward disks, 5 forward disks	7 m^2 16B pixels 5,200 MAPS sensors	400	26	26	17	MAPS: Several flavors: curved its-3 sensors for vertex Its-2 staves / w improvements	Fiber count limited by Artix Transceivers
MPGD tracking: Electron Endcap Hadron Endcap Inner Barrel Outer Barrel	16k 16k 30k 140k	8 8 30 72	1	.2	5	uRWELL / SALSA uRWELL / SALSA MicroMegas / SALSA uRWELL / SALSA	64 Channels/Salsa, up to 8 Salsa / FEB&RDO 256 ch/FEB for MM 512 ch/FEB for uRWELL
Forward Calorimeters: LFHCAL HCAL insert* ECAL W/SciFi Barrel Calorimeters: HCAL ECAL SciFi/PB ECAL ASTROPIX Backward Calorimeters: NHCAL ECAL (PWO)	63,280 8k 16,000 7680 5,760 500M pixels 3,256 2852	74 9 64 9 32 230 18 12	502	28	19	SiPM / HG2CROC SiPM / HG2CROC SiPM / Discrete SiPM / HG2CROC SiPM / HG2CROC Astropix SiPM / HG2CROC SiPM / Discrete	Assume HGCROC 56 ch * 16 ASIC/RDO = 896 ch/RDO 32 ch/FEB, 16 FEB/RDO estimate, 8 FEB/RDO conserve. HCAL 1536x5 *HCAL insert not in baseline Assume similar structure to its-2 but with sensors with 250k pixels for RDO calculation. 24 ch/feb, 8 RDO estimate, 23 RDO conservative
Far Forward: B0: 3 MAPS layers 1 or 2 AC-LGAD layer 2 Roman Pots 2 Off Momentum ZDC: Crystal Calorimeter 32 Silicon pad layer 4 silicon pixel layers 2 boxes scintillator	300M pixel 1M 1M (4 x 135k layers x 2 dets) 640k (4 x 80k layers x 2 dets) 400 11,520 160k 72	10 30 64 42 10 10 10 2	15	8	8	MAPS AC-LGAG / EICROC AC-LGAD / EICROC AC-LGAD / EICROC APD HGCROC as per ALICE FoCal-E	3x20cmx20cm 600^cm layers (1 or 2 layers) 13 x 26cm layers 9.6 x 22.4cm layers There are alternatives for AC-LGAD using MAPS and low channel count DC-LGAD timing layers
Far Backward: Low Q Tagger 1 Low Q Tagger 2 Low Q Tagger 1+2 Cal 2 x Lumi PS Calorimeter Lumi PS tracker	1.3M pixels 480k pixels 700 1425/75 80M pixels	12 12 1 1 24	150	1	4	Timepix4 Timepix4 (SiPM/HG2CROC) / (PMT/FLASH) Timepix4	
PID-TOF: Barrel Endcap	2.2M 5.6 M	288 212	31	1	17	AC-LGAD / EICROC (strip) AC-LGAD / EICROC (pixel)	bTOF 128 ch/ASIC, 64 ASIC/RDO eTOF 1024 pixel/ASIC, 24-48 ASIC/RDO (41 ave)
PID-Cherenkov: dRICH pfRICH DIRC	317,952 69,632 69,632	1242 17 24	1240 24 11	13.5 12.5 6	28 1 1	SiPM / ALCOR HRPPD / EICROC (strip or pixel) HRPPD / EICROC (strip or pixel)	Worse case after radiation. Includes 30% timing window. Requires further data volume reduction software trigger

Design Criteria for High Level Data Format

- Low overhead
- Generic
  - The file should have the structure of a read-only file system with banks taking part of files
  - Identification of banks by name (or by ID)
  - Bank Names or IDs should be maintained in an external database with plugin information for data bank decoding
- No backtracking while writing
  - No sizes at start or end of file
  - It is useful to support sizing directories within the header structure for efficiency
- Fractal (sub-files are files!)
  - Two valid files can be appended and remain valid files
  - Two valid files can be separated (at bank boundaries) and remain valid files
  - Physically embedding a bank into the file has simple controllable behavior
    - Embeds as a relative directory retaining substructure
    - Embeds as an absolute directory retaining substructure
- Low level library tools supporting header building and scatter gather to files, ethernet and IPC

These features mean that in subsystem testing / stand alone systems are simple

- They use identical readers
- They can use identical analysis codes
- They can be written at any stage of DAQ computing for integration and testing