ePIC Stream Computing Use Case: Autonomous Alignment and Calibration

Jin Huang (BNL)

Some initial considerations and hope to inspire more discussions

Calibration in ePIC computing

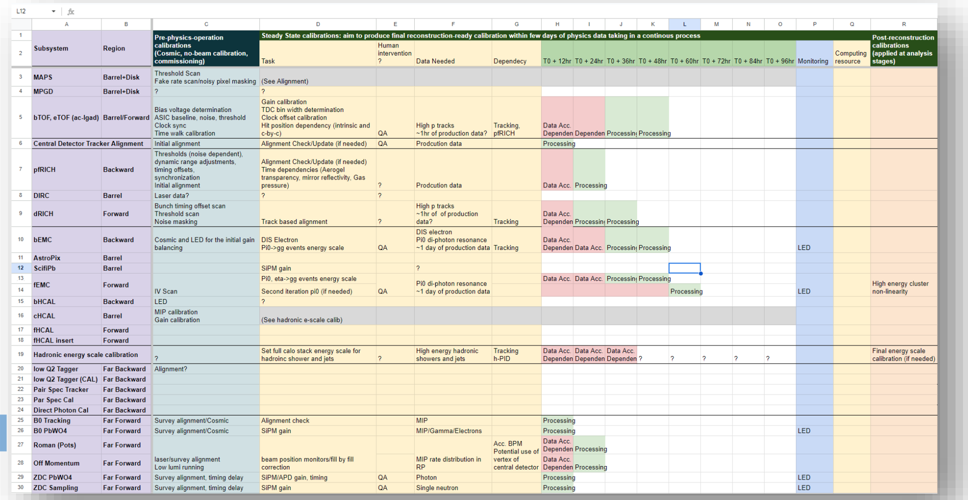
- ePIC aim to prompt reconstruction of experiment data at Echelon-1 facilities: aim to have latency of days, and < 3 weeks in steady state running
- The latency is driven by calibration
  - Collision/Calib data statistics required
  - Interdependency of detector calib.
- Concerted effort in SRO WG mapping out the calibration workflow
  - Link to each discusses in backup

Working document for calibration workflow

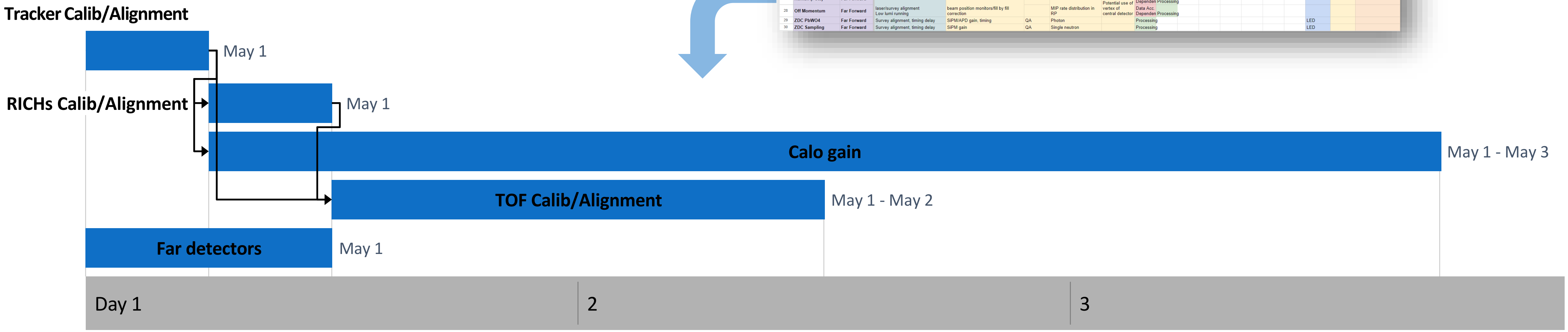
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1			Pre-physics-operation calibrations (Cosmic, no-beam calibration, commissioning)	Steady State calibrations: aim to produce final reconstruction-ready calibration within few days of physics data taking in a continous process														Post-reconstruction calibrations (applied at analysis stages)
2	Subsystem	Region		Task	Human intervention ?	Data Needed	Dependecy	T0 + 12hr	T0 + 24hr	T0 + 36hr	T0 + 48hr	T0 + 60hr	T0 + 72hr	T0 + 84hr	T0 + 96hr	Monitoring	Computing resource	
3	MAPS	Barrel+Disk	Threshold Scan Fake rate scan/noisy pixel masking	(See Alignment)														
4	MPGD	Barrel+Disk	?	?														
5	bTOF, eTOF (ac-lgad)	Barrel/Forward	Bias voltage determination ASIC baseline, noise, threshold Clock sync Time walk calibration	Gain calibration TDC bin width determination Clock offset calibration Hit position dependency (intrinsic and c-by-c)	QA	High p tracks ~1hr of production data?	Tracking, pfRICH	Data Acc. Dependen	Dependen	Processin	Processing							
6	Central Detector Tracker Alignment		Initial alignment	Alignment Check/Update (if needed)	QA	Prodcution data		Processing										
7	pfRICH	Backward	Thresholds (noise dependent), dynamic range adjustments, timing offsets, synchronization Initial alignment	Alignment Check/Update (if needed) Time dependencies (Aerogel transparency, mirror reflectivity, Gas pressure)	?	Prodcution data		Data Acc.	Processing									
8	DIRC	Barrel	Laser data?	?	?													
9	dRICH	Forward	Bunch timing offset scan Threshold scan Noise masking	Track based alignment	?	High p tracks ~1hr of production data?	Tracking	Data Acc. Dependen	Processin	Processing								
10	bEMC	Backward	Cosmic and LED for the initial gain balancing	DIS Electron Pi0->gg events energy scale	QA	DIS electron Pi0 di-photon resonance ~1 day of production data	Tracking	Data Acc. Dependen	Data Acc.	Processin	Processing					LED		
11	AstroPix	Barrel																
12	SciFiPb	Barrel		SiPM gain		?												
13	fEMC	Forward	IV Scan	Pi0, eta->gg events energy scale	QA	Pi0 di-photon resonance ~1 day of production data		Data Acc.	Data Acc.	Processin	Processing					LED		High energy cluster non-linearity
14				Second iteration pi0 (if needed)								Processing						
15	bHCAL	Backward	LED	?														
16	cHCAL	Barrel	MIP calibration Gain calibration	(See hadronic e-scale calib)														
17	fHCAL	Forward																
18	fHCAL insert	Forward																
19	Hadronic energy scale calibration		?	Set full calo stack energy scale for hadroinc shower and jets	?	High energy hadronic showers and jets	Tracking h-PID	Data Acc. Dependen	Data Acc. Dependen	Data Acc. Dependen	?	?	?	?	?			Final energy scale calibration (if needed)
20	low Q2 Tagger	Far Backward	Alignment?															
21	low Q2 Tagger (CAL)	Far Backward																
22	Pair Spec Tracker	Far Backward																
23	Par Spec Cal	Far Backward																
24	Direct Photon Cal	Far Backward																
25	B0 Tracking	Far Forward	Survey alignment/Cosmic	Alignment check		MIP		Processing										
26	B0 PbWO4	Far Forward	Survey alignment/Cosmic	SiPM gain		MIP/Gamma/Electrons		Processing								LED		
27	Roman (Pots)	Far Forward	laser/survey alignment Low lumi running	beam position monitors/fill by fill correction		MIP rate distribution in RP	Acc. BPM Potential use of vertex of central detector	Data Acc. Dependen	Processing									
28	Off Momentum	Far Forward						Data Acc. Dependen	Processing									
29	ZDC PbWO4	Far Forward	Survey alignment, timing delay	SiPM/APD gain, timing	QA	Photon		Processing								LED		
30	ZDC Sampling	Far Forward	Survey alignment, timing delay	SiPM gain	QA	Single neutron		Processing								LED		

Calibration workflow

Working document for calibration workflow



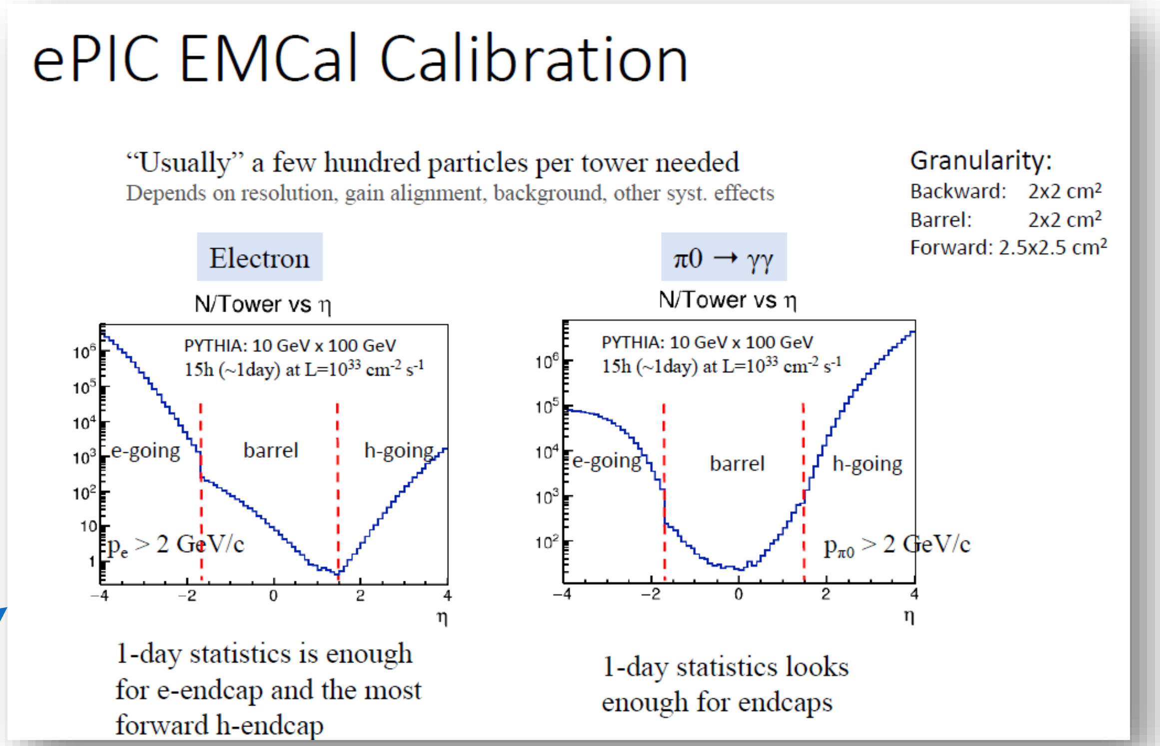
- Calibration workflow seems fits into the prompt reconstruction computing model. Inputs welcomed.
- High level summary plot:

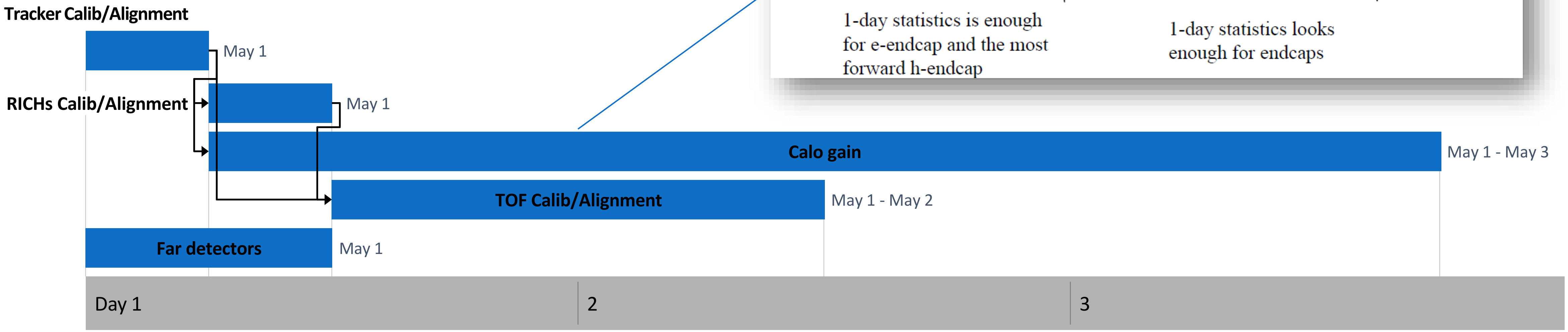


Calibration workflow

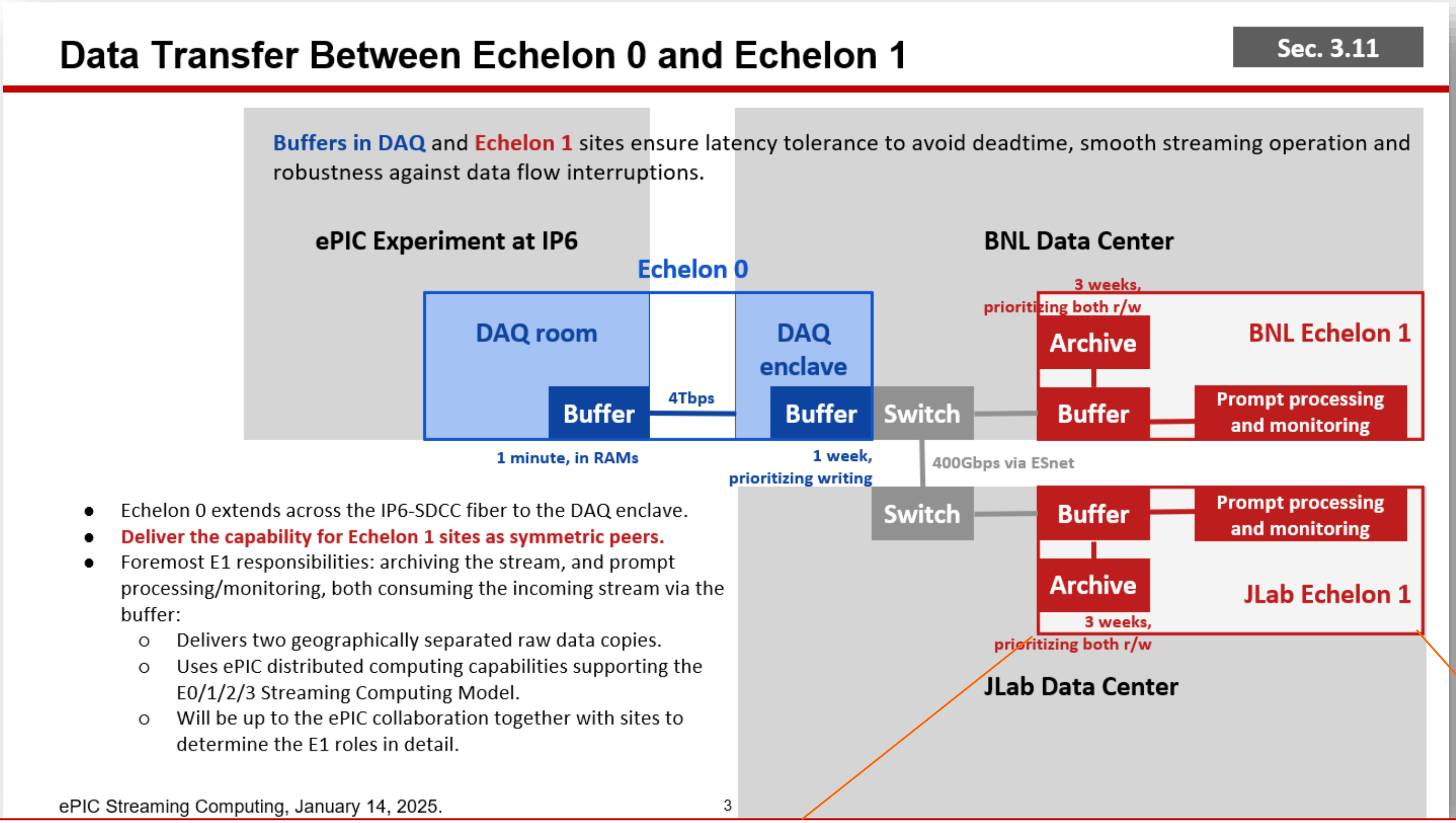
Critical path goes through calo calibration which require 1-day of data + two iteration of calibration processing

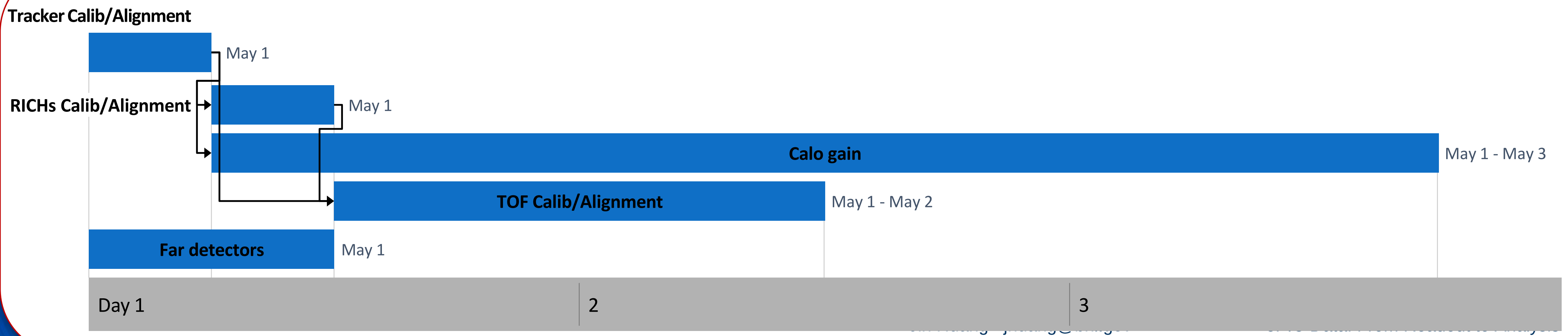
Work from A.Bazilevsky
Cited in calo calib meeting [ref]





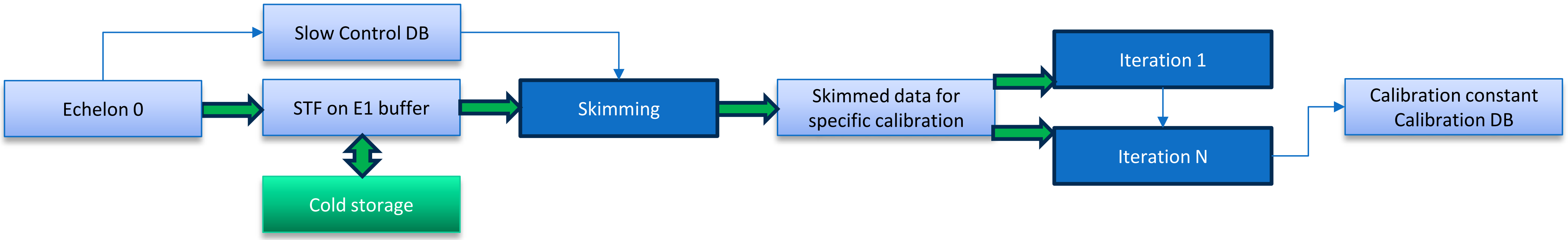
Streaming computing integration





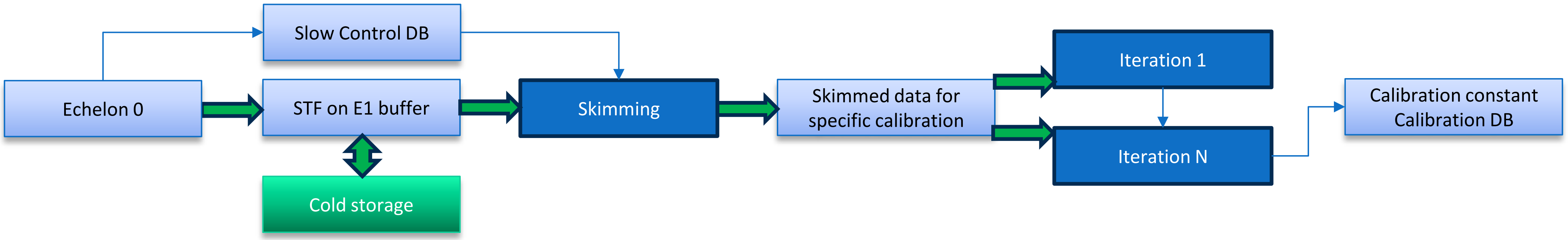
Discussion on streaming computing integration: Calibration workflow

- Calibration workflow has strong integration with prompt reconstruction and Run QA
- A set of calibration would only initiate if QA identify a calibration drift
  - Example tracking residual analysis found a shift in the residual distribution after a run period → tracker + RICH alignment update
  - Calibration workflow starts with automated/shifter-driven QA check flagging issue
- Another set of calibration would process ALL data
  - Example is calorimeter EM energy scale calibration which will use all pi0 and eta0 resonance data
  - Calibration workflow starts with arrival of super time frame (STF)
- Robustness against new problems
  - Much of the calibration is a detector debugging process, in particular at beginning of ePIC



Discussion on streaming computing integration: Data staging and computing

- Calibration jobs should be guaranteed to run with resources of two Echeon-1 computing centers
  - Ensuring completing the calibration jobs in days and start of physics-quality reconstruction
- Skimming pass for calibration input data
  - Many calibration tasks require rarer events: e.g. pi0 in barrel EMCal; high momentum pion for RICH
  - Start with skimming pass of “raw data” in STF that is equivalent to high level triggering
    - (1) allow faster processing of multi-iterations (e.g. two iteration for pi0-driven E-scale)
    - (2) reduces Echelon-1 buffer disk IO read load, so the STF raw data are only read in entirety twice (2x data taking rate): calibration skimming and full production.
- Slow control/facility data: (1) primary input via database (2) backup in raw data stream



Discussion on streaming computing integration: Calibration constant management

- Calibration constant be reproducible for given data and software tags
  - Folded into the reproducibility of ePIC results from DAQ output (“raw data”)
- Avoid frequent calibration constant update beyond the varying detector and facility conditions
  - Stable calibration constant help accumulate statistics for data driven performance/uncertainty determination
  - Caution against continuously updating calibration processes: a brief excursion of calibration constant could lead to a large systematic uncertainty (e.g. high energy tail for calorimeter)
- Also for reproducibility: Calibration results should be tagged in database → tagged onto meta data of production output files for analyzers

Extra Information

Feel free to share your views

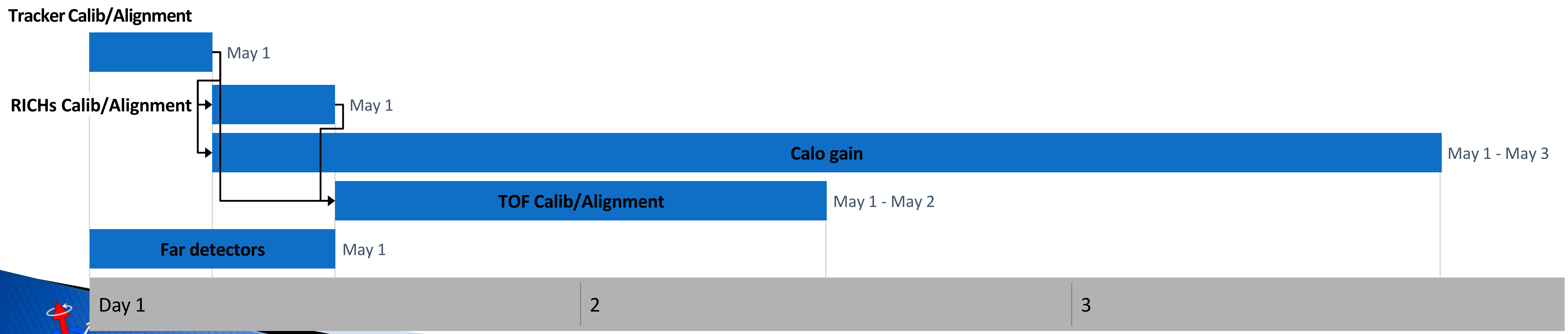
Live note on indico [link]

Relevant WG meetings

- Alignment, TOF: Dec 19 https://indico.bnl.gov/event/21619/
- SVT sensor, Barrel Hcal: Jan 23 https://indico.bnl.gov/event/21785/
- dRICH: Jan 30 https://indico.bnl.gov/event/22114/
- Backward, Forward EMCal: Feb 27 https://indico.bnl.gov/event/22412/
- Far forward: Mar 12 https://indico.bnl.gov/event/22676/
- AI driven calibration: Apr 16 https://indico.bnl.gov/event/23034/

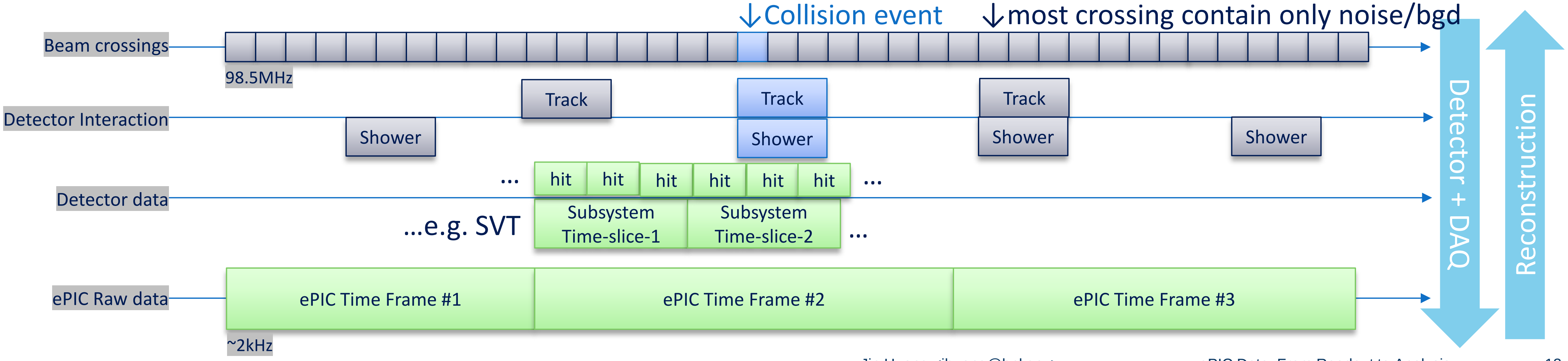
Open discussions / path forward

- So far the calibration workflow seems fit well into reconstruction latency goal
- Suggestions always welcomed. And further subsystem inputs needed:
  - Hadronic energy scale, Barrel EMCal, DIRC, Far backward detectors
- Computing resource estimation (so far seems << reconstruction)
- Summarize into next update of computing model and computing review



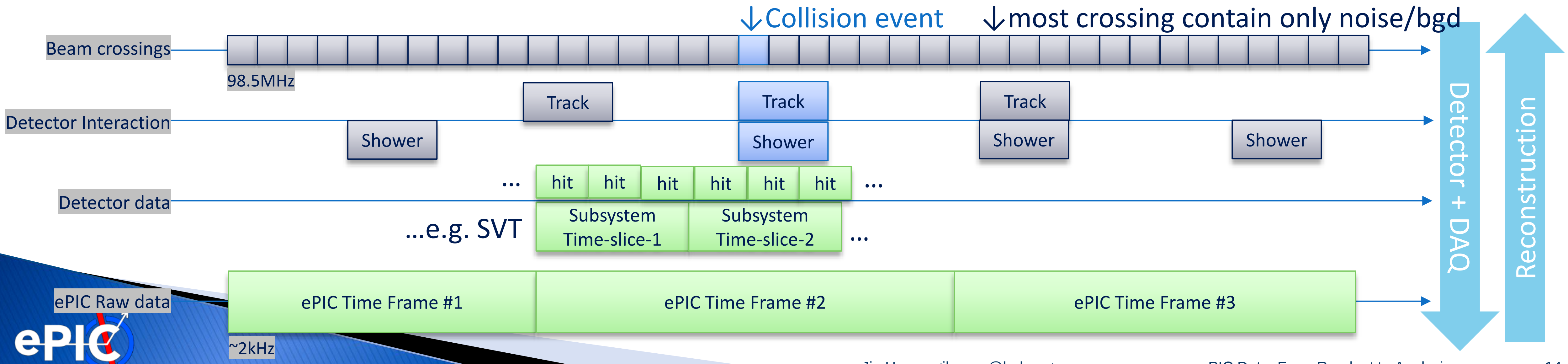
Summary of consensus in SRO WG Apr-9 meeting

- Preference not to align time frame length with respect to the EIC beam rotation.
- Event keying: primary key is 64-bit beam clock (BCO) counter; secondary convenient key is tuple run-timeframe-BCOInTimeFrame; reconstruction will generate event counter tagging
- Run structure will be used, driven by configuration changes; plus continuous readout information on beam/detector monitoring
- Redundant information in storing slow control data: database and raw data file embedding. Need to follow up on the implementation of SC data flow from online to offline.



Time-Frames Introduction

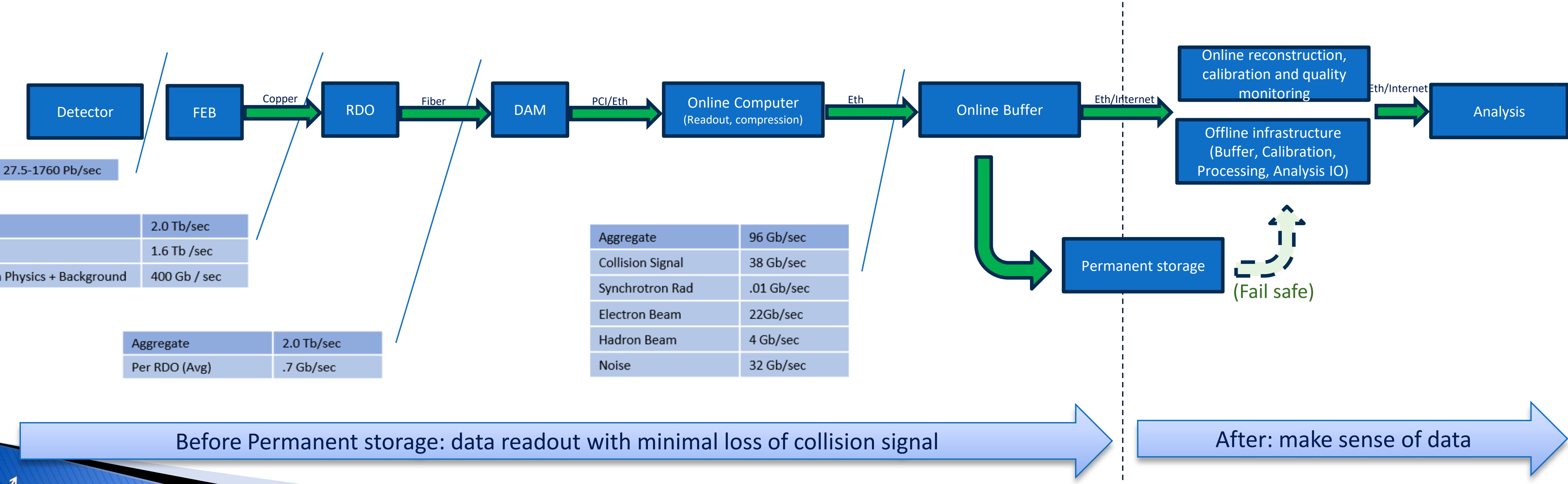
- We plan to use this meeting to follow up on Nathan's talk on time-frame-based reconstruction, solidify a few open concept in our WG and make progress on their implementation in EICRecon
- ePIC Time Frame concept is developing towards a spec doc in DAQ and SAR WGs;
  - Update discussion on Apr 11 DAQ meeting , please join: https://indico.bnl.gov/event/22945/
  - <=2^16 crossing: 16-bit integer sufficient to locate hit's BX in Time Frame; <=665us/300 events/10MB
  - Exact length defined by GTU sync signal: most flexible
  - We could choose to align with EIC beam evolution (1260BX,): simpler to locate abort gap and spin states
- Time Frames will be order in data files, internally carry header-payload (a.k.a data bank/packets) data chunks from each detector component.



Streaming DAQ – Computing : consideration 2

For kickstart the discussion, please interrupt to discuss at any moment

- Sooner or later, a copy of data is stored and saved for permanent storage
- This stage of first permanent storage could be viewed as a DAQ – computing boundary



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	22Gb/sec
Hadron Beam	4 Gb/sec
Noise	32 Gb/sec