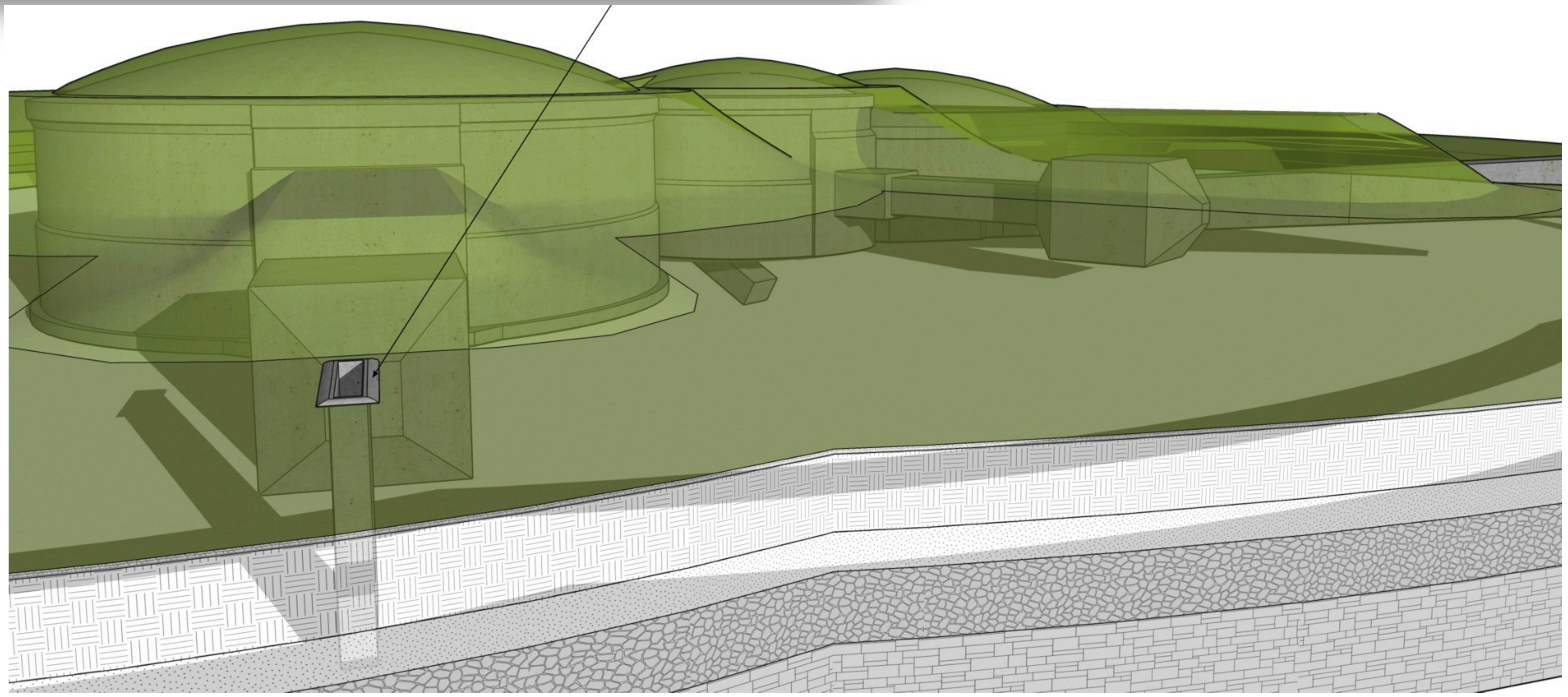
Bi-weekly WP2 meeting

Tuesday 22 Oct 2024, 14:00 → 16:10 Europe/Rome

Alberto Annovi (Istituto Nazionale di Fisica Nucleare) , Piergiulio Lenzi (Istituto Nazionale di Fisica Nucleare)

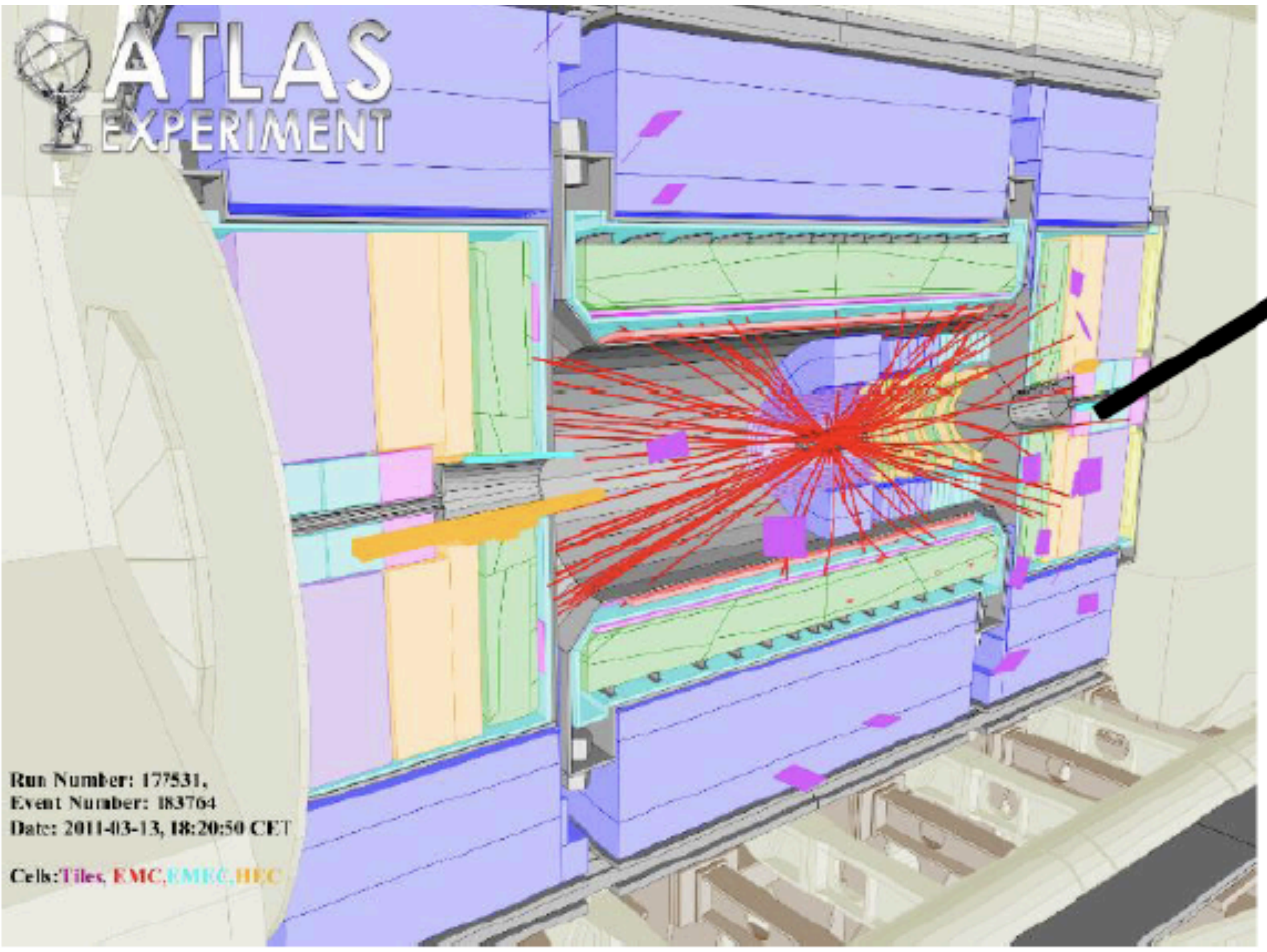


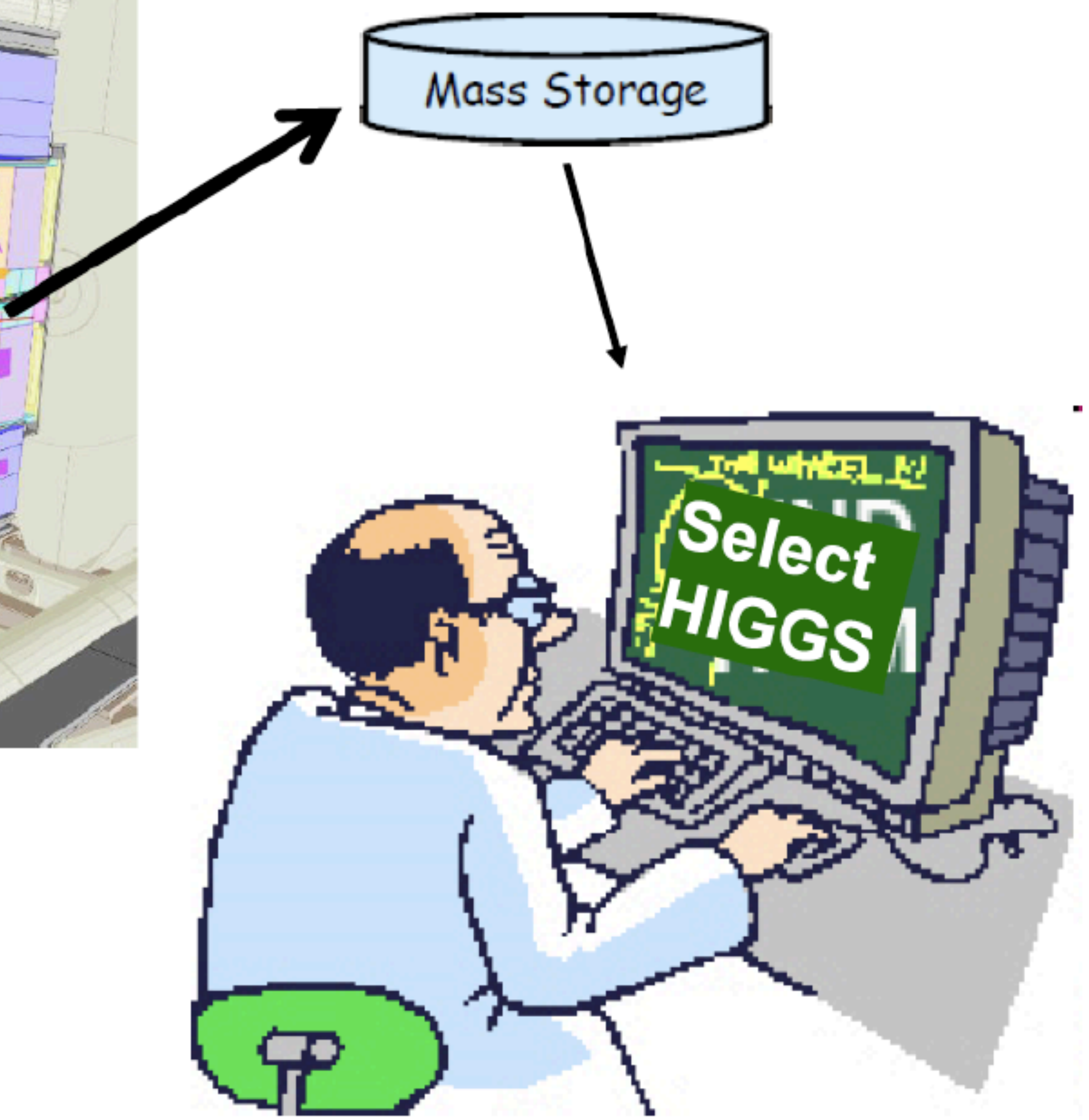
Online data reduction for the BDX experiment at Jefferson Lab

M.Battaglieri (INFN)

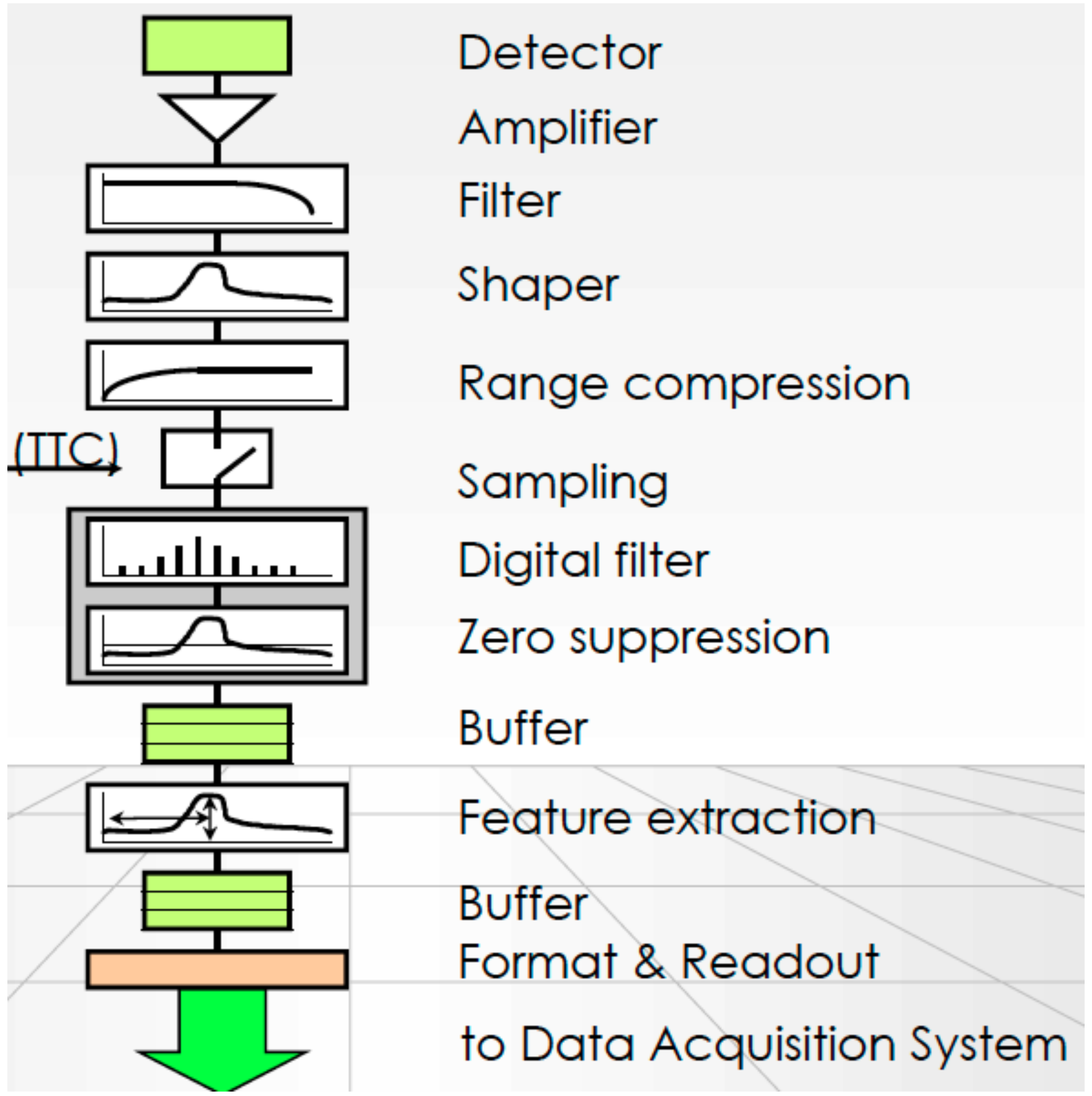
Jefferson Lab · clas · e- JLab12 · Brookhaven National Laboratory · ePIC · INFN

From signals to physics



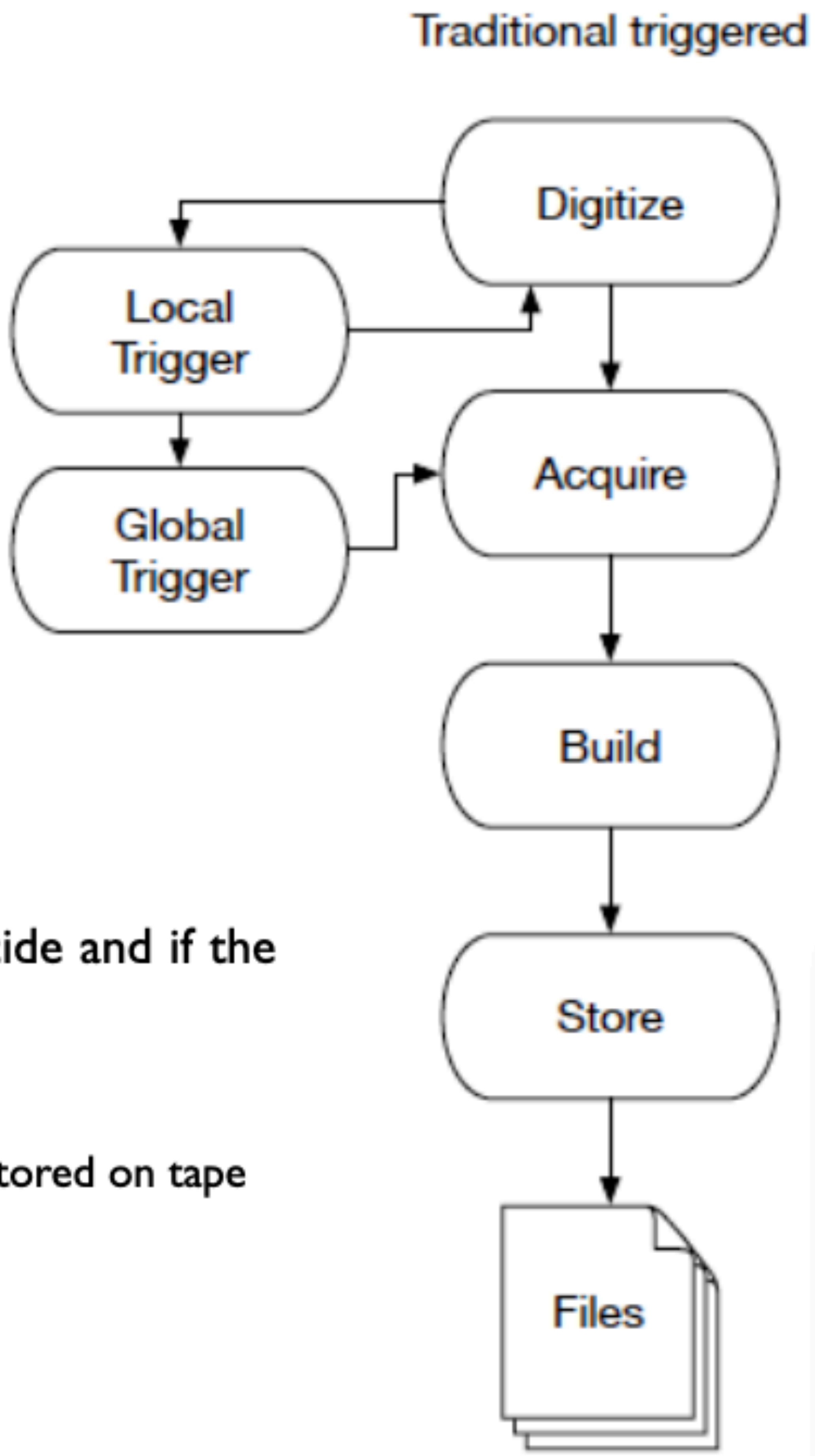


DAQ chain



Triggered DAQ

Traditional (triggered) DAQ



✳ (few) trigger Channels participating send (partial) information to trigger logic

✳ All channels continuously measured, hits stored in short term memory

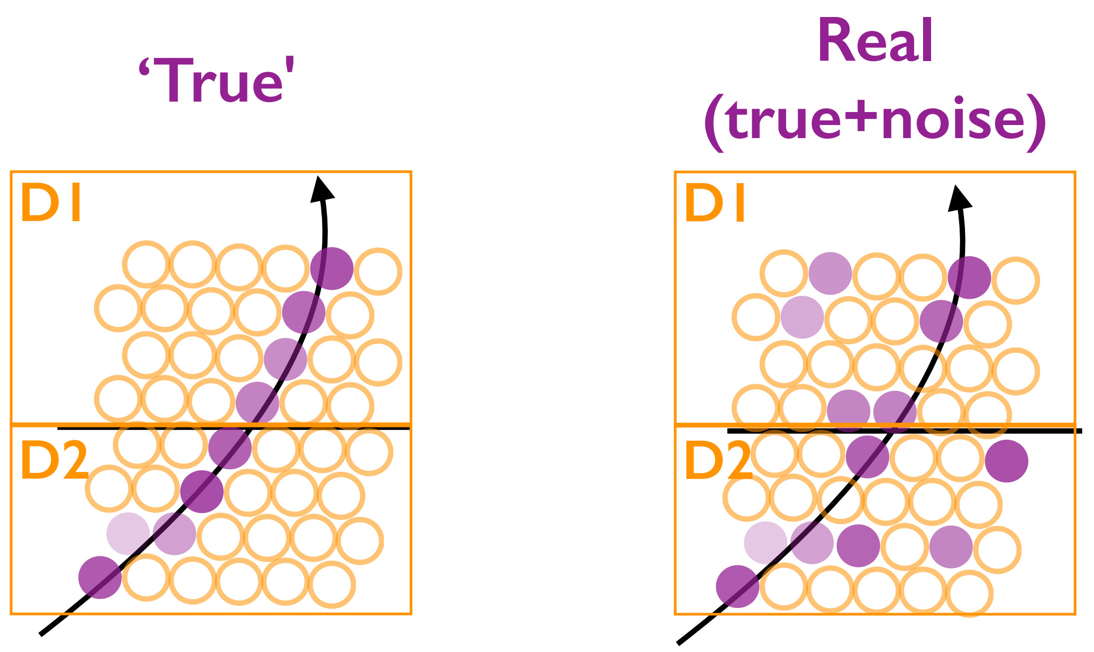
✳ Trigger logic takes time to decide and if the trigger condition is satisfied:
- a new 'event' is defined
- trigger signal back to the FEE
- data read from memory and stored on tape

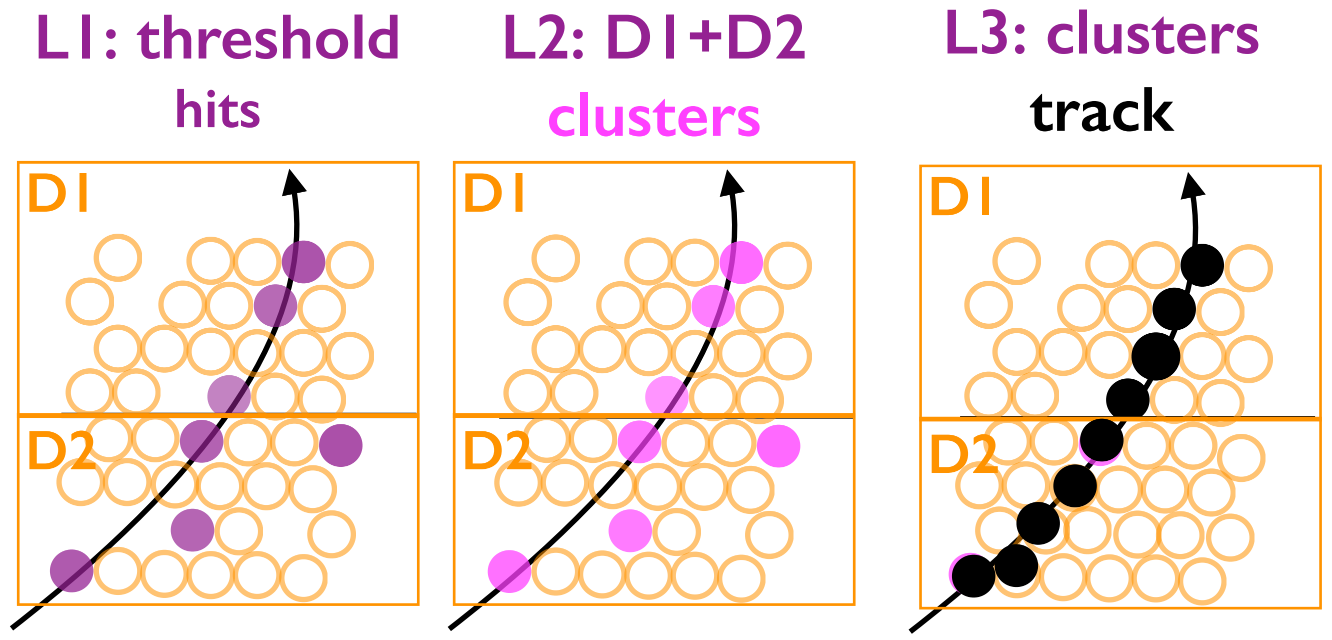
Traditional triggered DAQ

- **Pros**
  - we know it works reliably!
- **Drawbacks:**
  - only few information forms the trigger
  - Trigger logic (FPGA) difficult to implement and debug
  - not easy to change and adapt to different conditions

Trigger logic

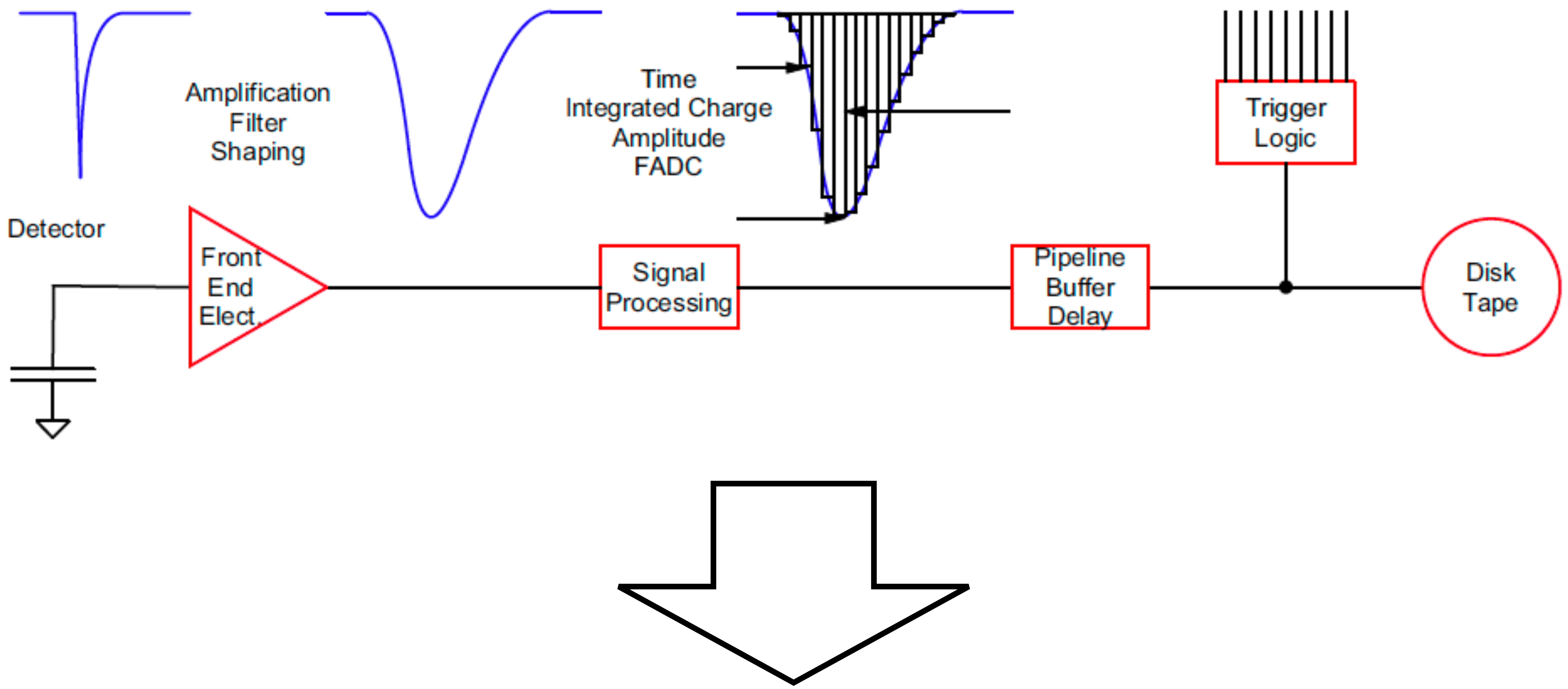
★ decides if/when to collect detector information
★ Select 'events' over 'background'
★ Save data on disk for further processing
★ Different levels
- L1: threshold on FEE
- L2: combine information from different sub-detector components
- L3: requires info processing



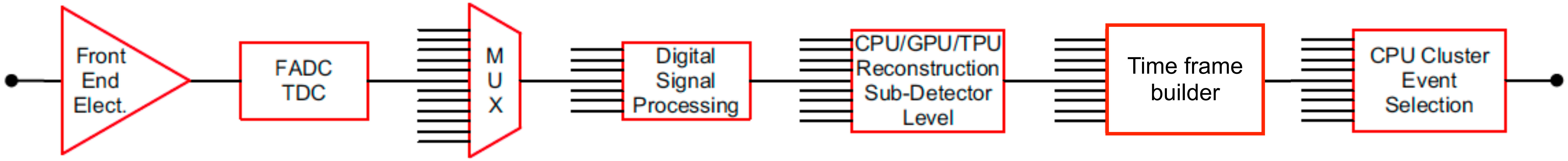


Streaming readout

Triggered DAQ



Streaming readout DAQ



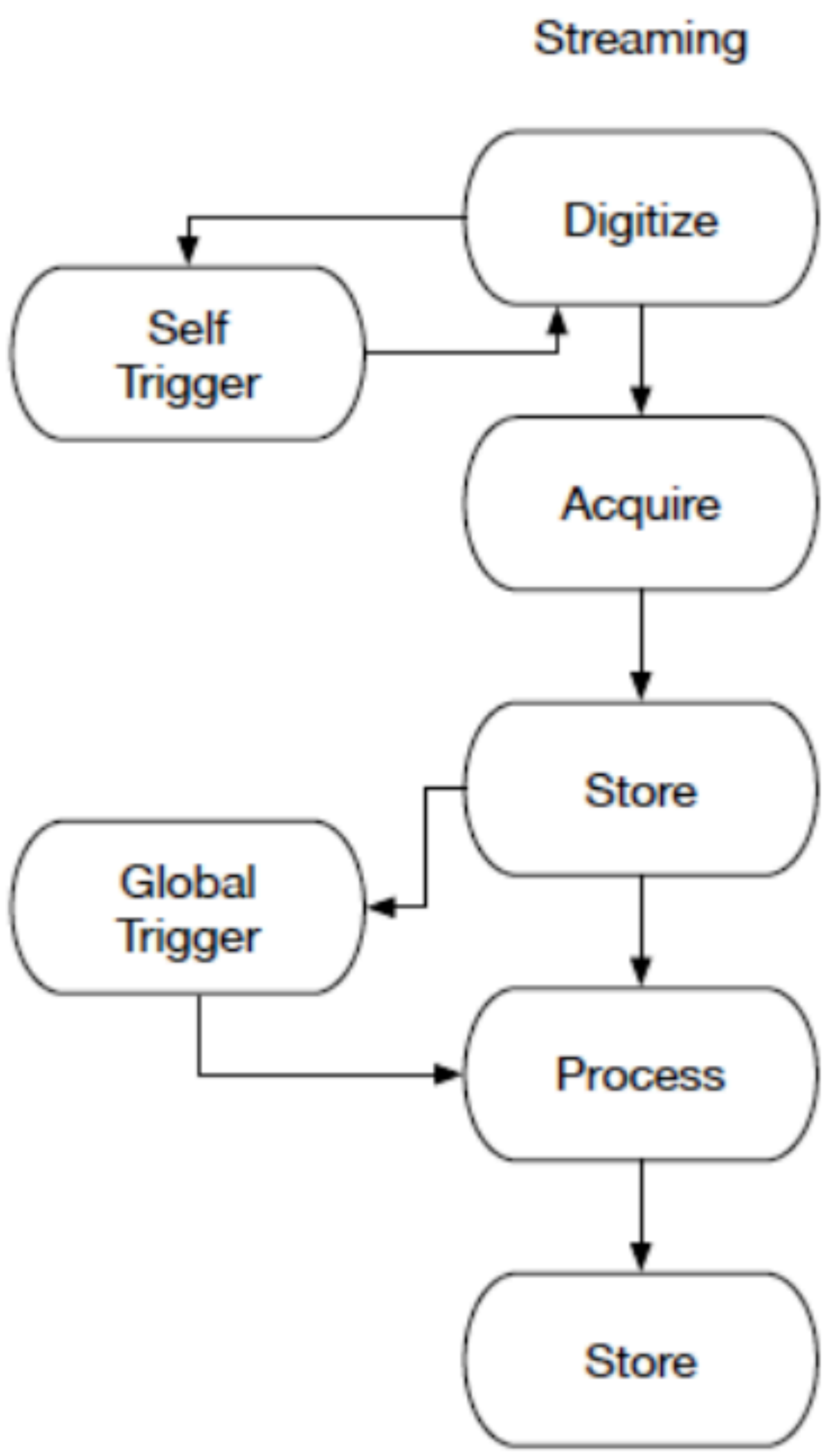
Streaming RO

Streaming read out (SRO)

✶ A HIT MANAGER receives hits from FEE, order them and ship to the software defined trigger

✶ Software defined trigger re-aligns in time the whole detector hits applying a selection algorithm to the time-slice

- the concept of 'event' is lost
- time-stamp is provided by a synchronous common clock distributed to each FEE



✶ All channels continuously measured and hits streamed to a HIT manager (minimal local processing) with a time-stamp

SRO DAQ

- **Pros**
  - All channels can be part of the trigger
  - Sophisticated tagging/filtering algorithms
  - high-level programming languages
  - scalability
- **Drawbacks:**
  - we do not have the same experience as for TRIGGERED DAQ

Why SRO is so important?

✶ **High luminosity experiments**
- Write out the full DAQ bandwidth
- Reduce stored data size in a smart way (reducing time for off-line processing)

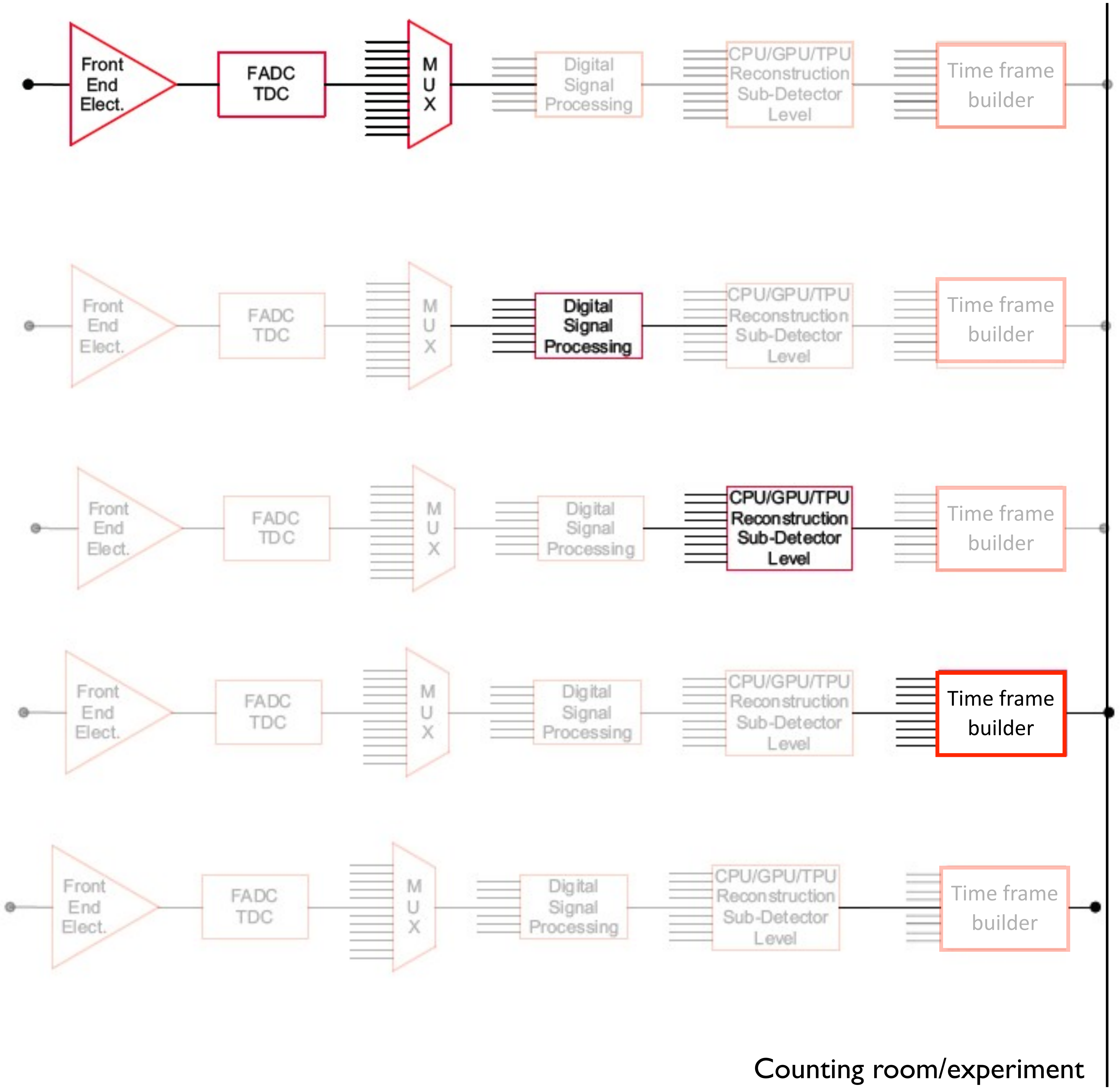
✶ **Shifting data tagging/filtering from the front-end (hw) to the back-end (sw)**
- Optimize real-time rare/exclusive channel selection
- Use of high-level programming languages
- Use of existing/ad-hoc CPU/GPU farms
- Use of available AI/ML tools
- (future) use of quantum-computing

✶ **Scaling**
- Easier to add new detectors in the DAQ pipeline
- Easier to scale
- Easier to upgrade

Many NP and HEP experiments adopt a SRO DAQ

- CERN: LHCb, ALICE, AMBER
- FAIR: CBM
- DESY: TPEX
- FRIBS: GRETA
- BNL: sPHENIX.ePIC
- JLAB: SOLID, BDX, CLAS12, …

Streaming RO

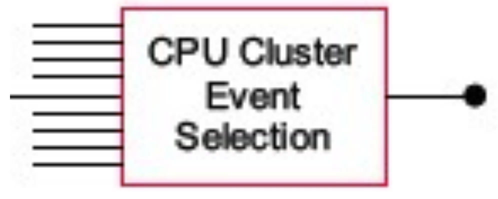


- FEE optimised for SRO
  - ASICS (cheap) or fADC (multiplexing) at (O($10/ch)
  - TDC if necessary to replace fADC
  - Zero-suppression mode
  - Fast readout (optical link)

- Signal pre-processing with fast hw (dedicated FPGA)
  - de-multiplexing fADC info
  - Charge, time, amplitude
  - Data compression
  - Data monitoring
  - Add other information (e.g. ch_ID eTimeStamp)

- CPU/GPU/TPU sub-detector analysis (single stream)
  - Local clusters, track segments, PID, …
  - Time-frame building
  - If necessary only store high-level data dumping raw

- TF-Router Time frame construction
  - Use time stamps to reorganise data from all streams in time frames



Data center

- Full reconstruction CPU analysis (for each time frame)

AI/ML shall play a significant role in each of these steps

AI-supported algorithms for SRO

Real Time data analysis

- In the SRO scheme, data analysis is performed online [this does not prevent to save unbiased frames for further analysis!]
- A św trigger is released based on real-time data analysis
- SRO and real-time data processing shall use AI:
  - to adapt data analysis to the changed conditions of the run (e.g. thresholds)
  - to identify data features in real-time (e.g.clusters)
  - to extract calibration constants from a data sub-set
  - to define algorithms to run (fast!) in real time on heterogeneous systems (e.g. CPU+GPU+FPGA)

Partial Real-Time data reconstruction: clustering

- Look at all detector information (hit: x, y, t, E) to learn correlations: clusters of objects share common features
- Define a metric in a space and identify cluster features
- Tests on minimum bias trigger data before real-time
- Hyperparameters optimization based on data

Fast inference

- Fast algorithms to extract data features to be used in data selections (and reduction)
- Mimicking a smart 'trigger'
- provide partial reconstructed quantity quickly

Calibration

- Use smart algorithms to extract data features and correct detector parameters varying over time
- toward a self-calibrating detector

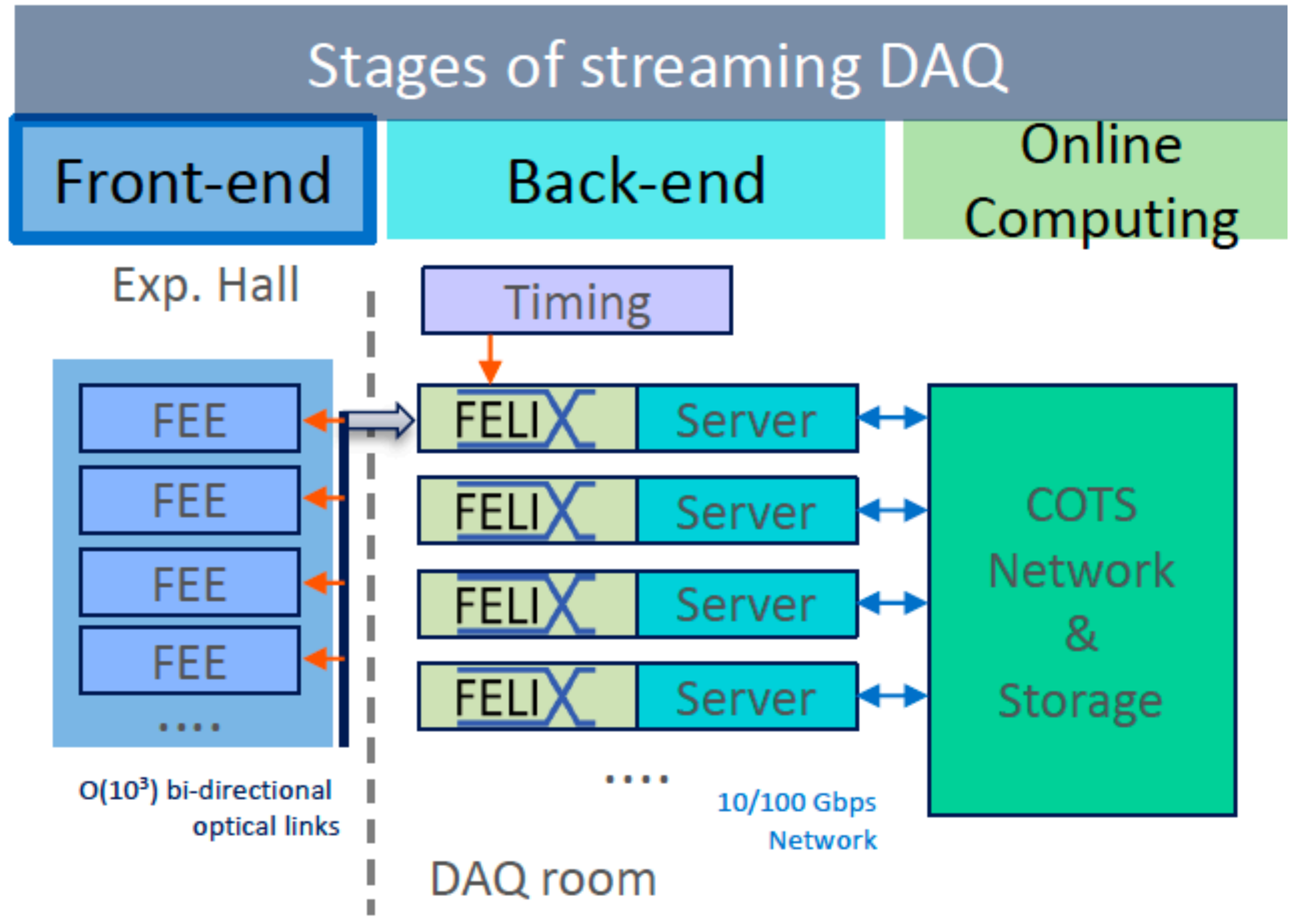
Data reduction

- reduce data volume to a manageable level with minimum bias

Realtime data reduction

Data reduction represents a main challenge in SRO

- Traditional DAQ: triggering (+ high level triggering/ reconstruction and compression) reduces data volume
- Streaming DAQ needs to reduce data real-time: zero-suppression, feature building, lossy compression



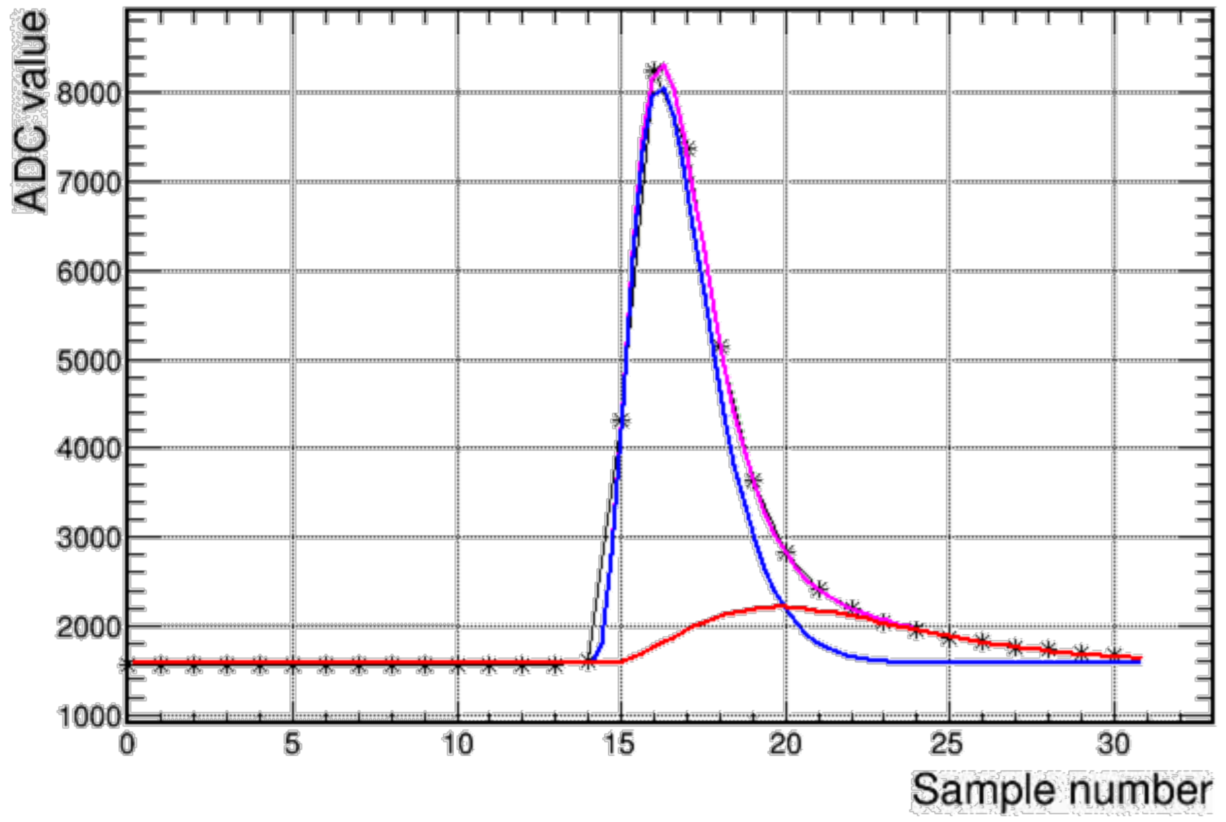
Opportunities for real-time AI but also a challenge:

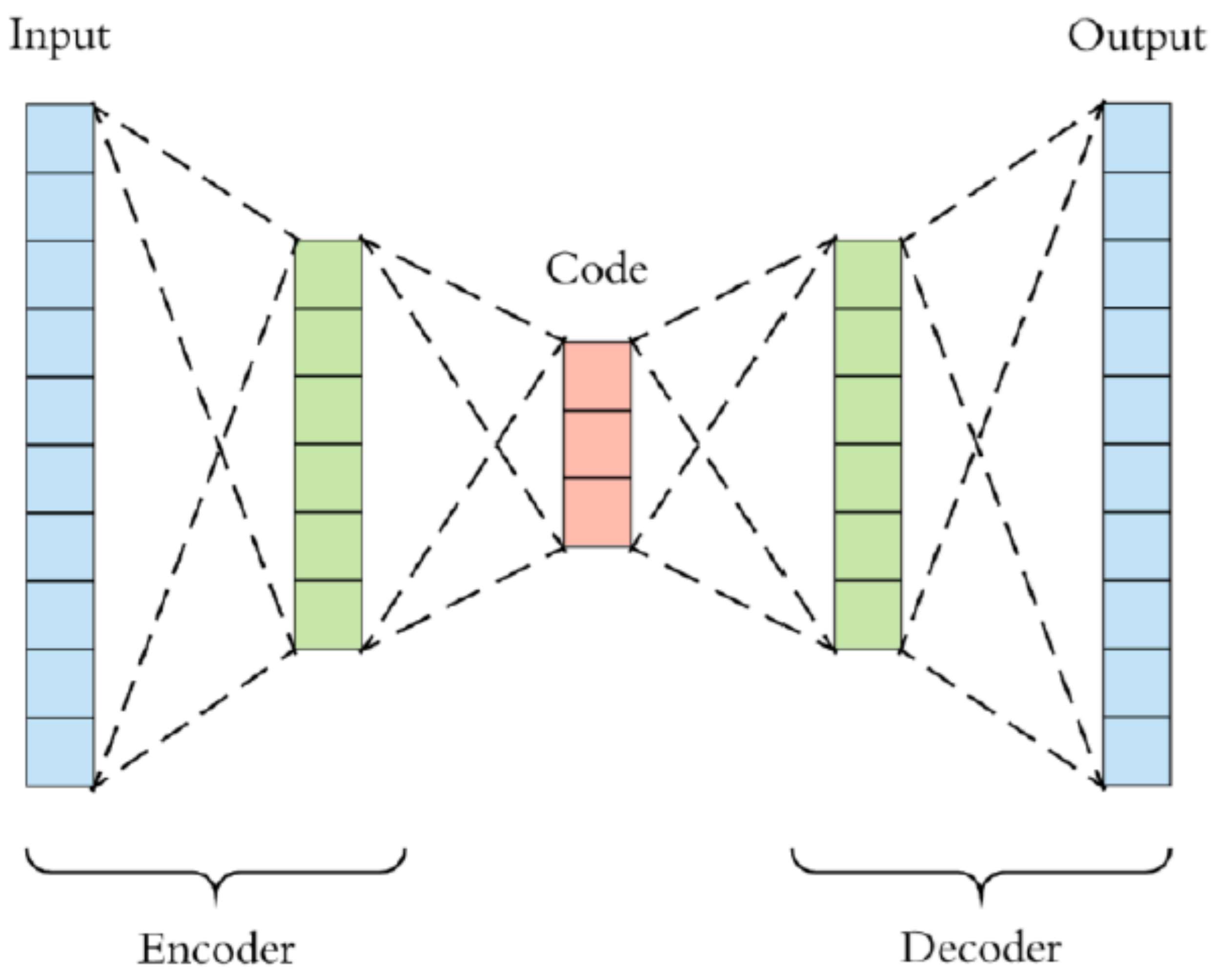
- reliable data reduction
- Applicable at each stages of streaming DAQ (front-end electronics, readout back-end, online computing)
- Data quality monitoring, fast calibration/reconstruction

Front end electronics

- Digitization (ADC, TDC, pixel readout)
- Data reduction strategy to immediately apply zero-suppression
- Real-time AI data reductions:
  - Improved zero-suppression (e.g.small signal recovery)
  - Feature building
  - Compression
- Target hardware: ASIC, (smaller) FPGAs Common requirement of low-power consumption, radiation tolerant

- Waveform digitizer: output data in ADC time series
- NN can be used in the FE to extract features (e.g. amplitude and time)
- Fit limited resources in FEE FPGA or ASIC
- quantized-aware training and pruning

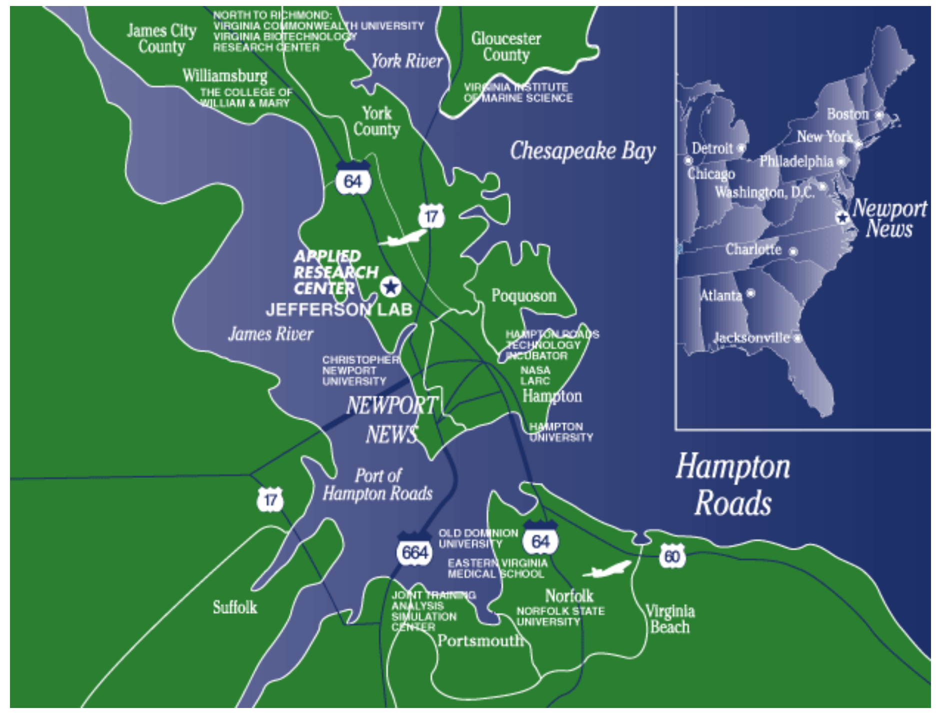




Autoencoder

- Charge (Energy) and time are compact to stream but partial
- fast and efficient way to preserve the full (anagogic) wave-form information
- Reduce the traffic on the first stages of the SRO DAQ pipeline

Jefferson Lab



Jefferson Lab's accelerator site

✳Primary Beam: Electrons

✳ Beam Energy: 12 GeV

- 10 > λ > 0.1 fm
- nucleon → quark transition
- baryon and meson excited states

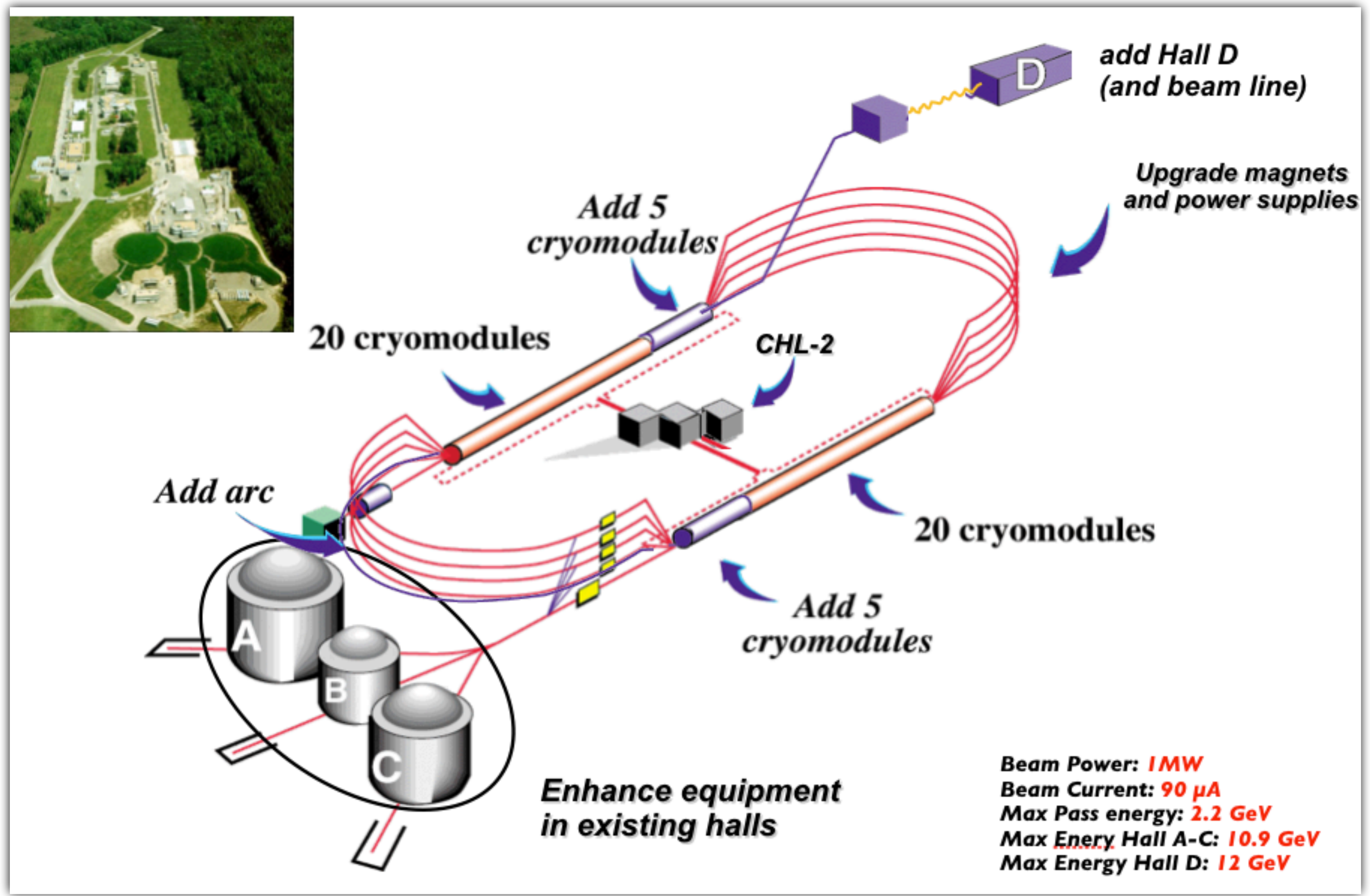
✳100% Duty Factor (cw) Beam

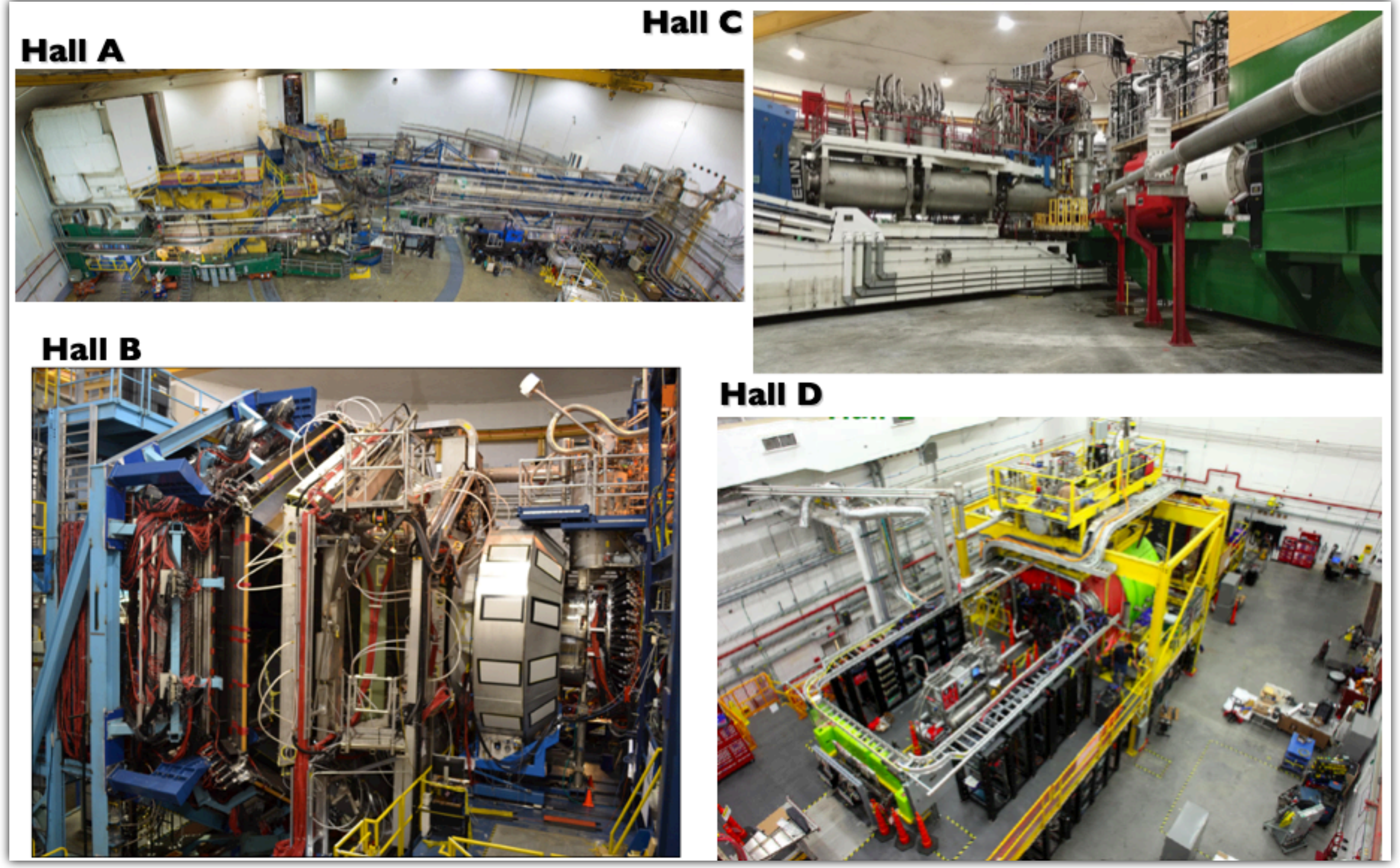
- coincidence experiments
- Four simultaneous beams
- Independent E and I

✳ Polarization

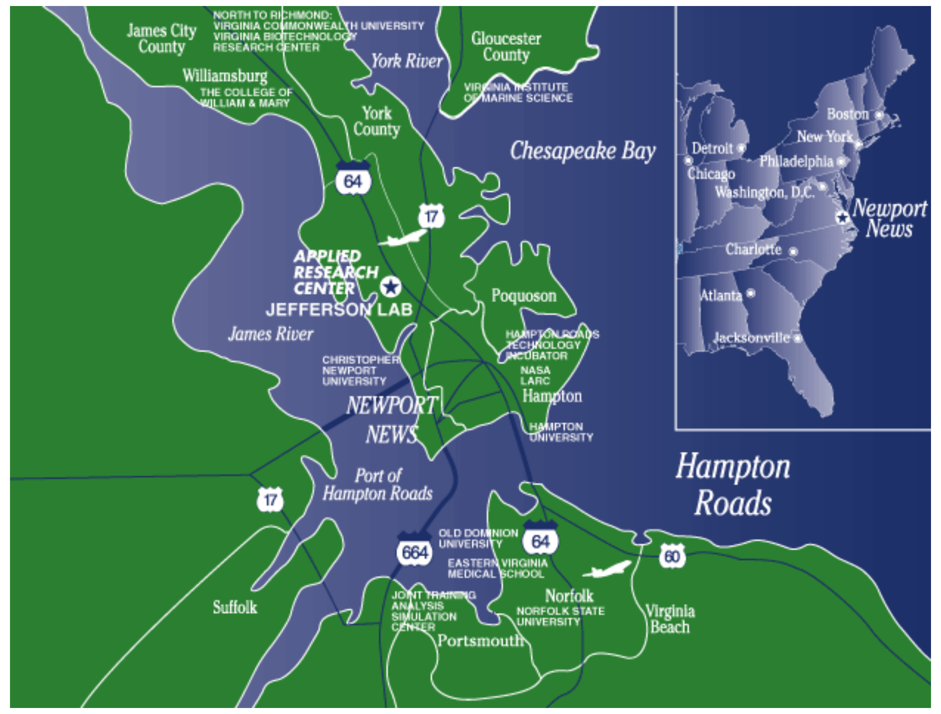
- spin degrees of freedom
- weak neutral currents

Luminosity > 10^7 -10^8 x SLAC at the time of the original DIS experiments!





Jefferson Lab




Jefferson Lab's accelerator site

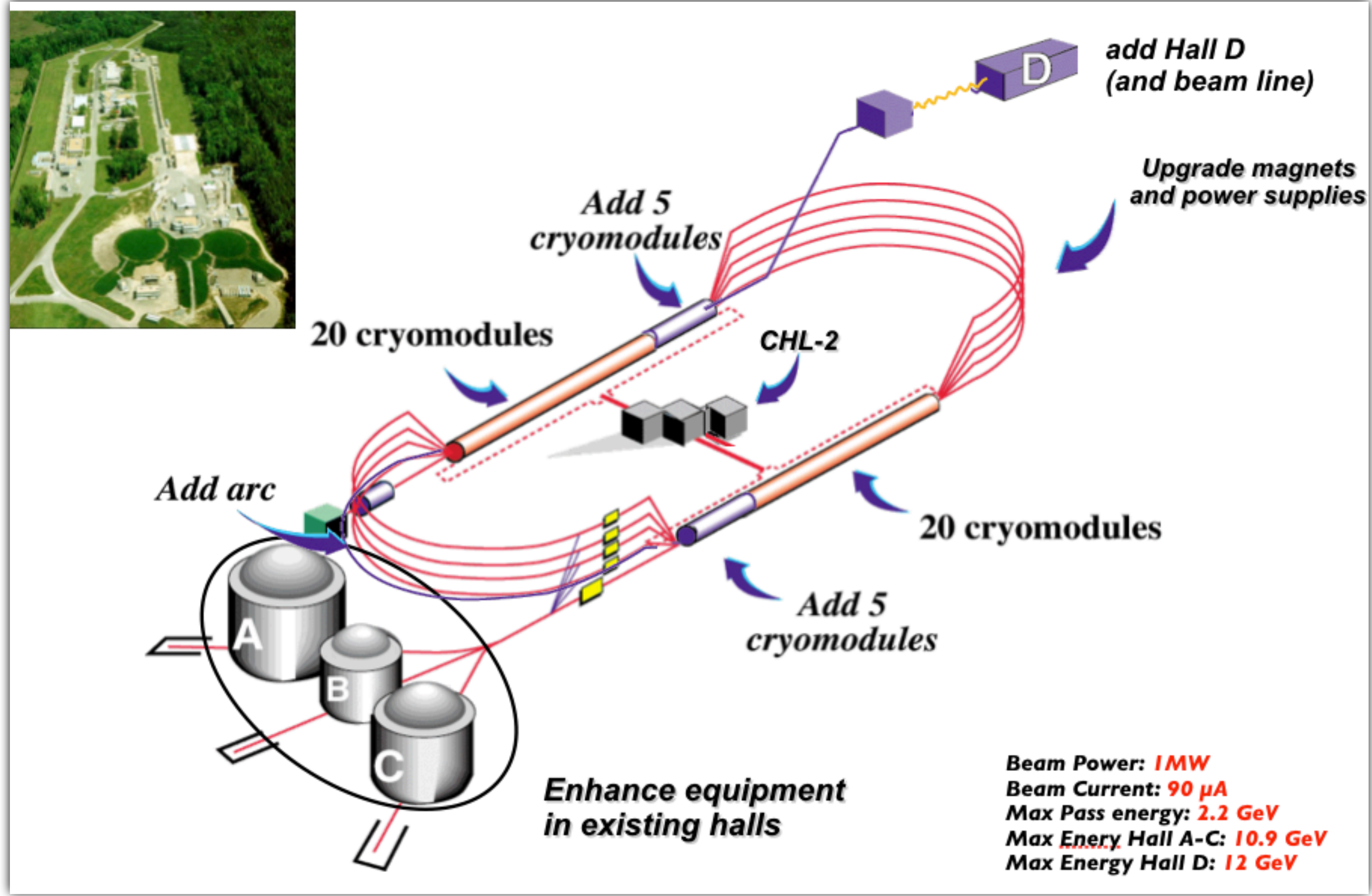
✳Primary Beam: Electrons

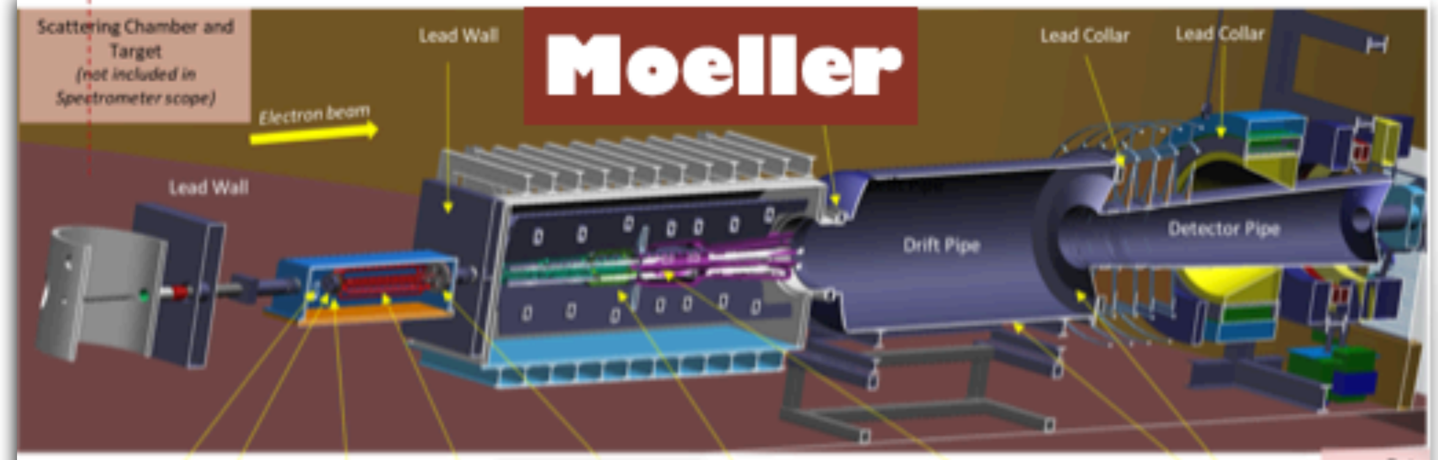
✳ Beam Energy: 12 GeV
- 10 > λ > 0.1 fm
- nucleon → quark transition
- baryon and meson excited states

✳100% Duty Factor (cw) Beam
- coincidence experiments
- Four simultaneous beams
- Independent E and I

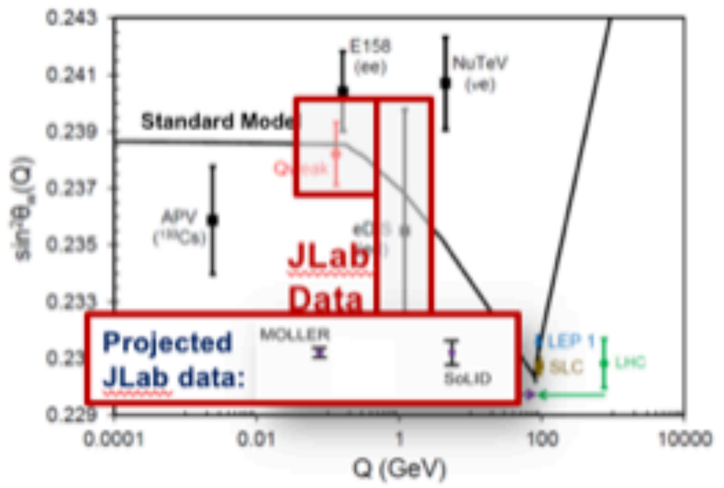
✳ Polarization
- spin degrees of freedom
- weak neutral currents

Luminosity > 10^7 - 10^8 x SLAC
at the time of the original DIS experiments!

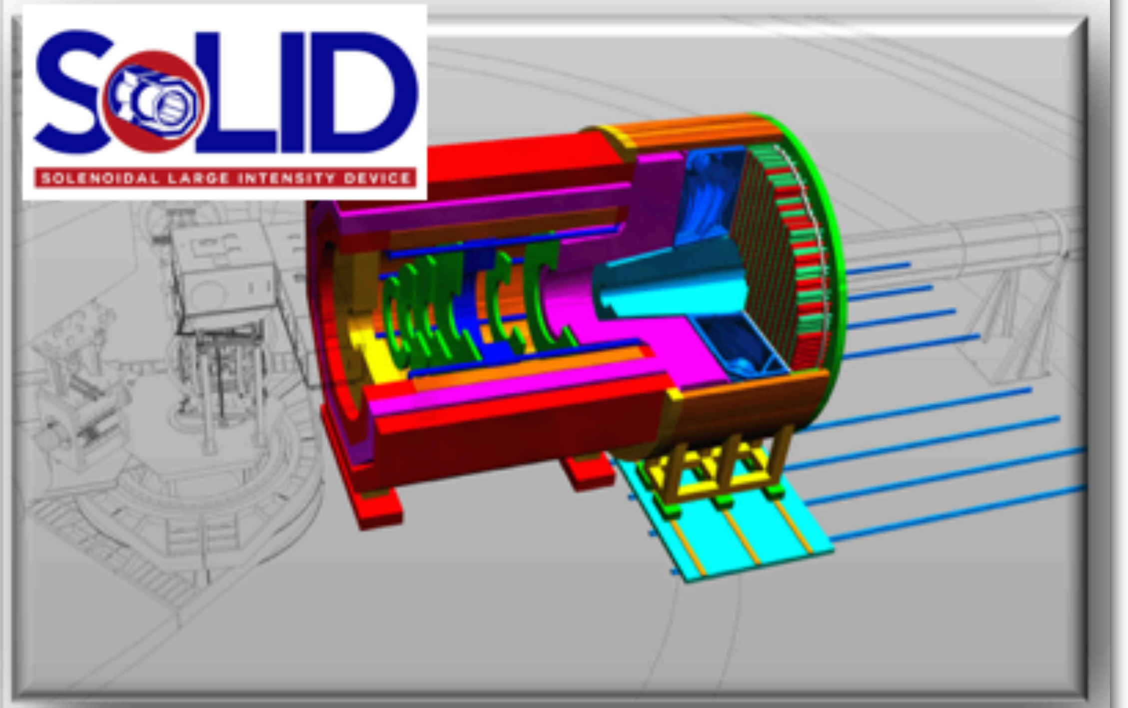


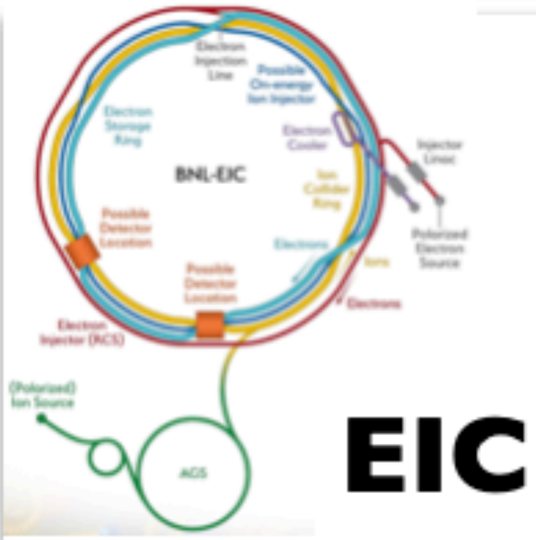


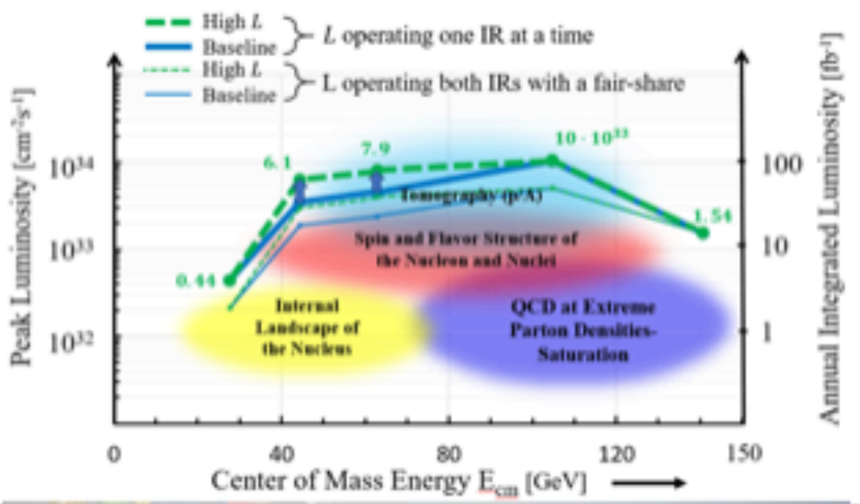
- Unique discovery space for new physics up to 38 TeV mass scale, with a purely leptonic probe
- CD-1 approved Dec 2020
- Expected to operate in FY26



- **SO**lenoidal **L**arge **I**ntensity **D**evice – new multipurpose detector facility optimized for high luminosity ($10^{37\text{-}39}$ cm^{-2} s^{-1}) and large acceptance







- Luminosity 100-1000 times that of HERA
- Polarized protons and light nuclear beams
- Nuclear beams of all A (p→U)
- Center mass variability with minimal loss of luminosity

- Large acceptance
- Frwrd/Bckw angles

- Precise vertexing
- HRes Tracking
- Excellent PID

The Beam Dump eXperiment - BDX

Spokespersons:
M.Battaglieri (INFN), M.Bondi (INFN), A.Celentano (INFN), M.DeNapoli (INFN), R.DeVita (JLab), G.Krnjaic (FNAL)

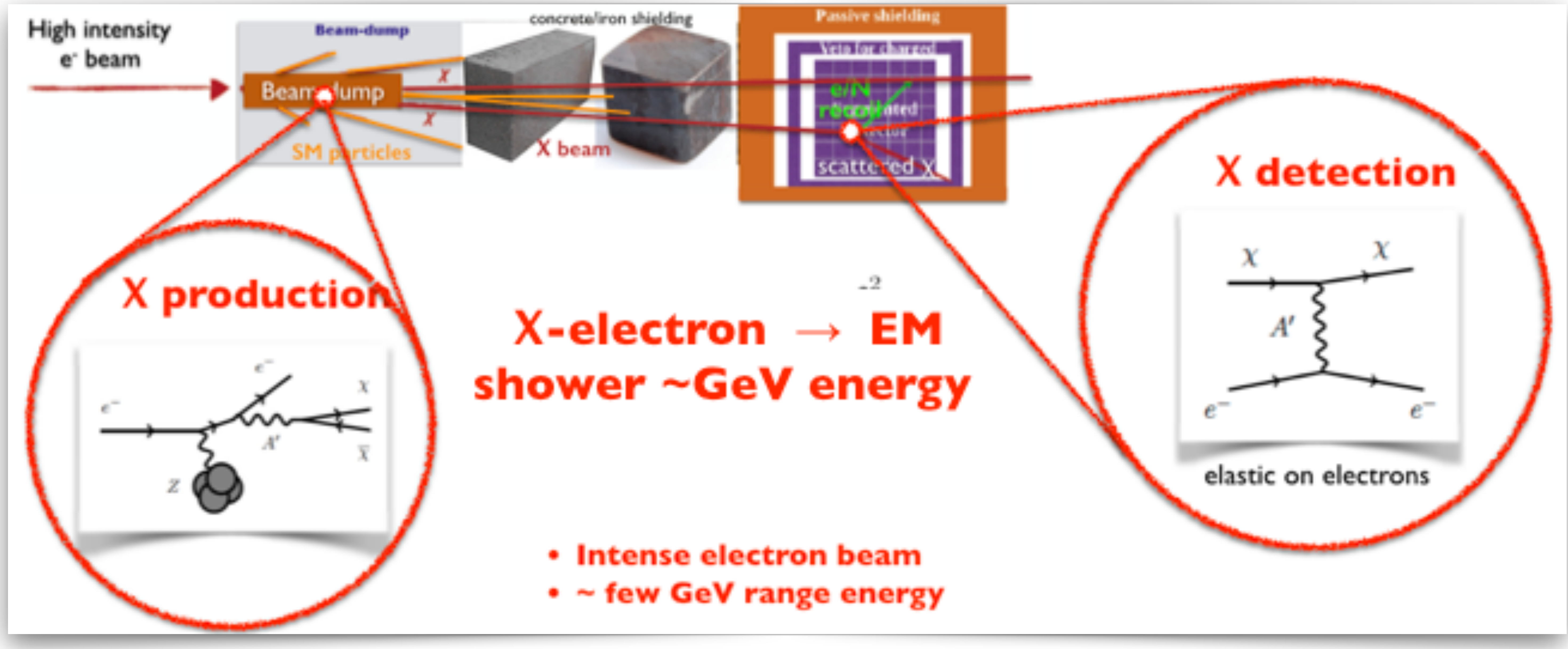
★ Unique experiment able to **PRODUCE** and **DETECT** Light Dark Matter

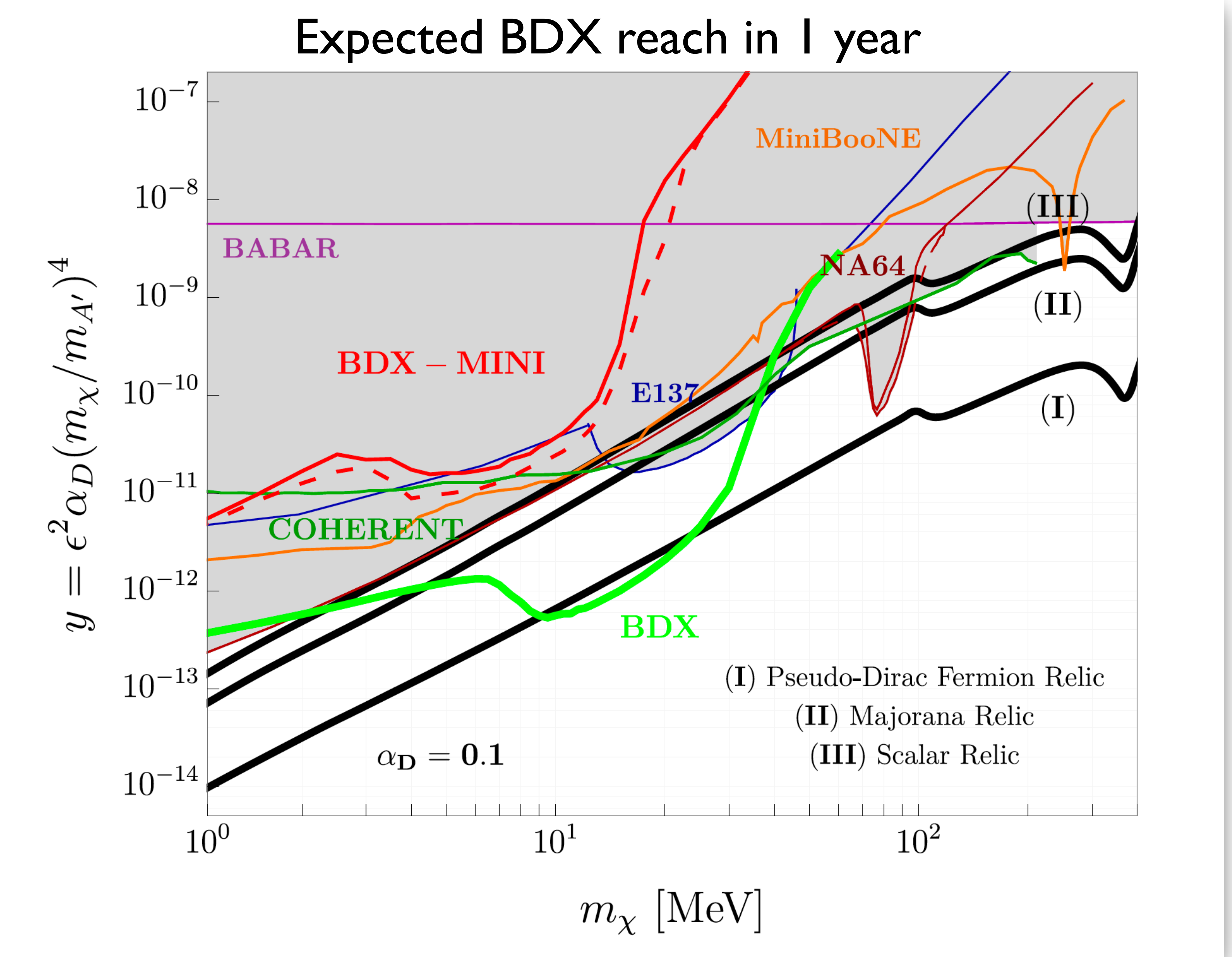
Two-step process:

I) An electron radiates an A' and the A' promptly decays to a χ (DM) pair

II) The χ (in-)elastically scatters on an e⁻/nucleon in the detector producing a visible recoil (GeV)

★ BDX will improve by 2 orders of magnitude current exclusion limits in LDM parameter space with sensitivity to the most viable scenarios (eg. relic LDM)





★ JLab offers the best condition for BDX:

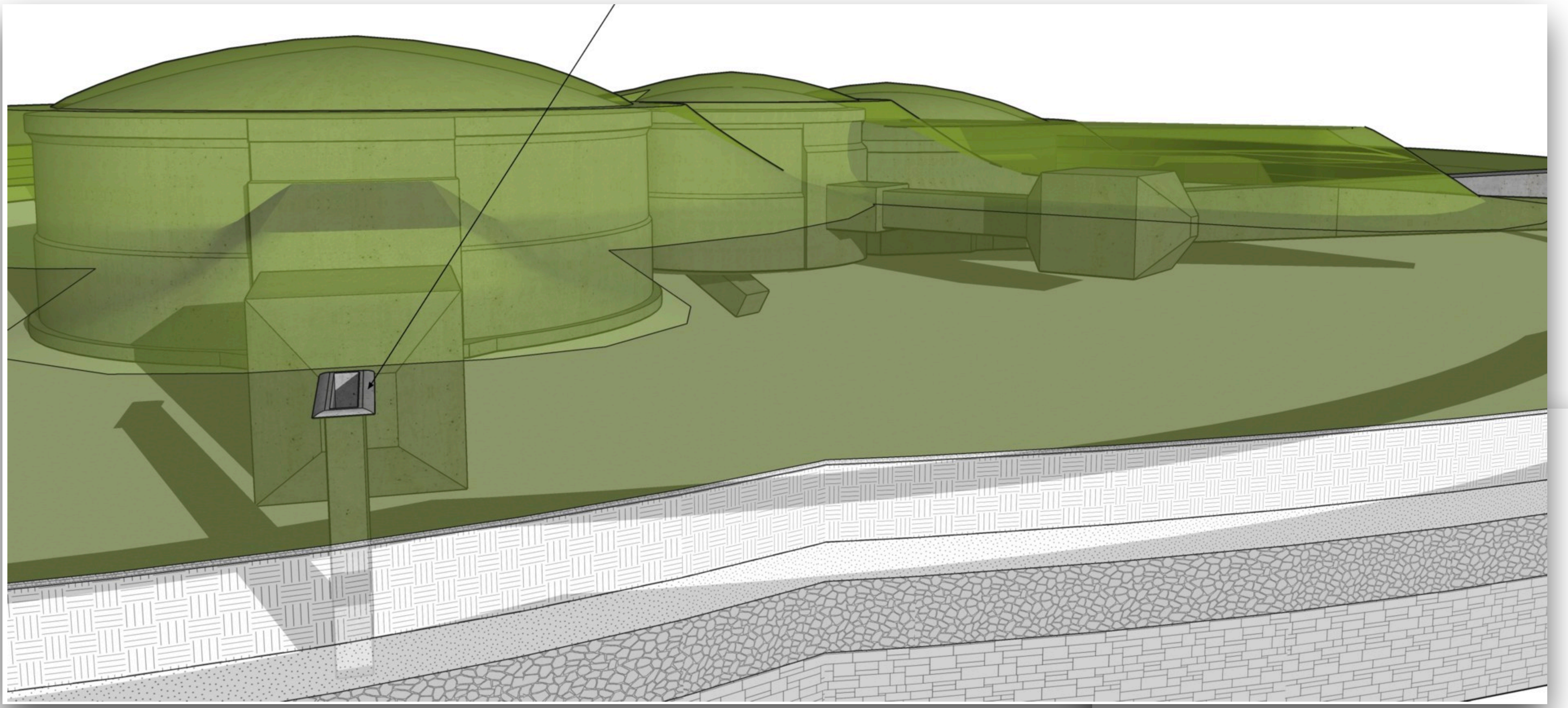
- A high energy beam: 11 GeV
- The highest available electron beam current: ~65 uA
- The highest integrated charge: 10^{22} EOT (41 weeks)
- Fully parasitic wrt Hall-A physics program (Moeller experiment)

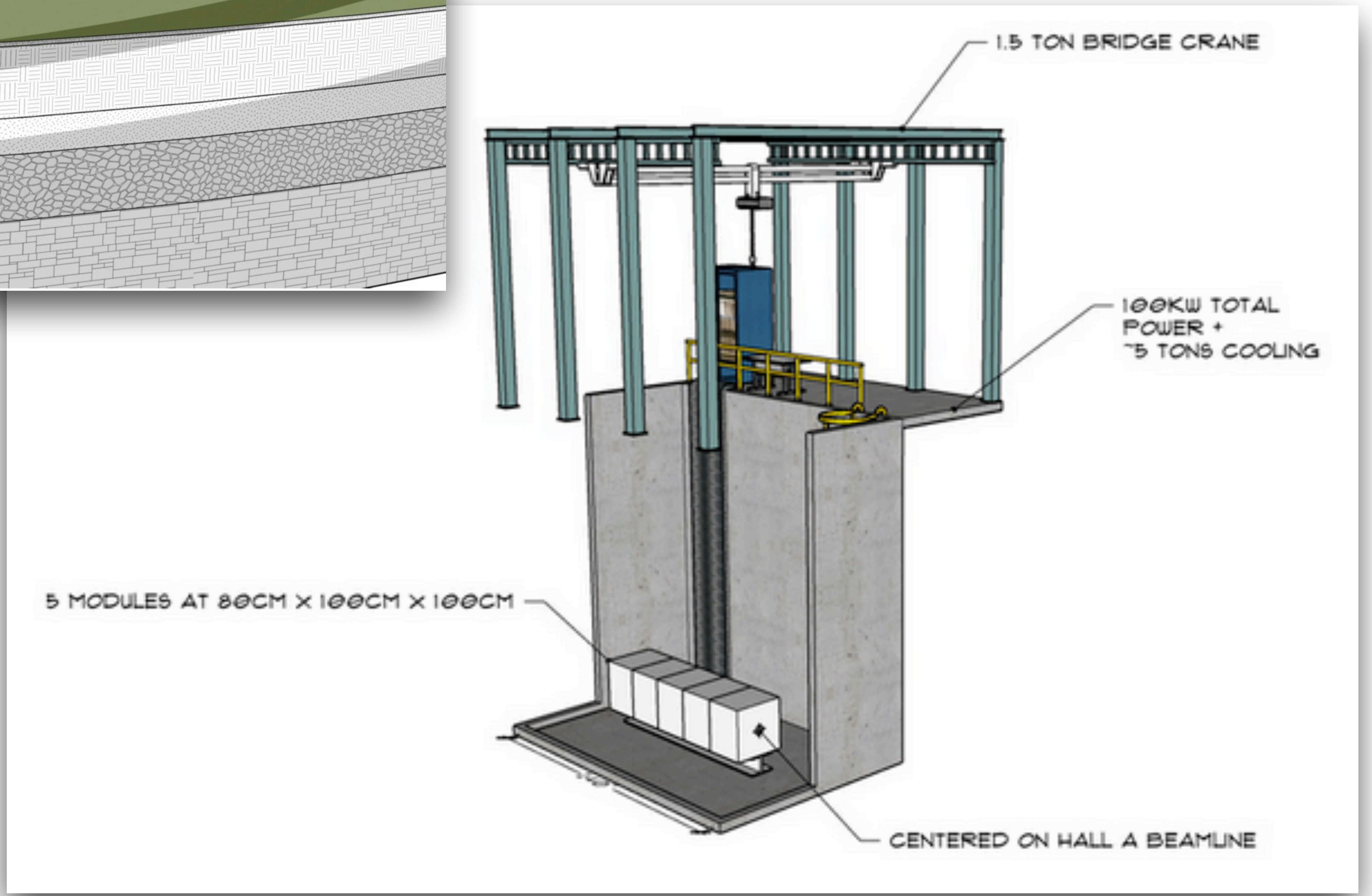
★ Approved by JLab PAC-46 in July 2018 (reconfirmed in 2023 by PAC-51) with maximum scientific rating (A) and waiting for scheduling

The Beam Dump eXperiment - BDX

BDX infrastructures:

★ New underground hall downstream of the Hall-A beam-dump and shielding wall (~2m of lead between the dump and the concrete vault or ~7m of iron downstream of the concrete vault)

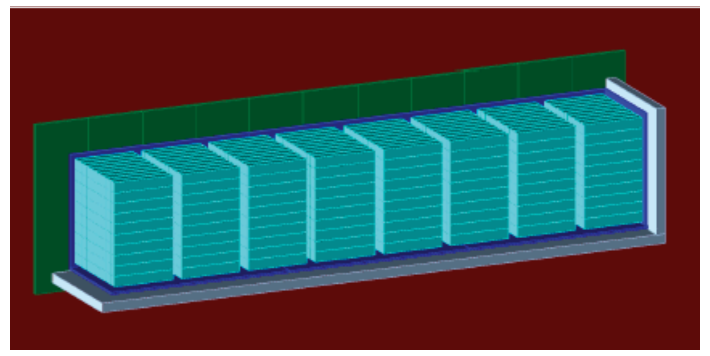
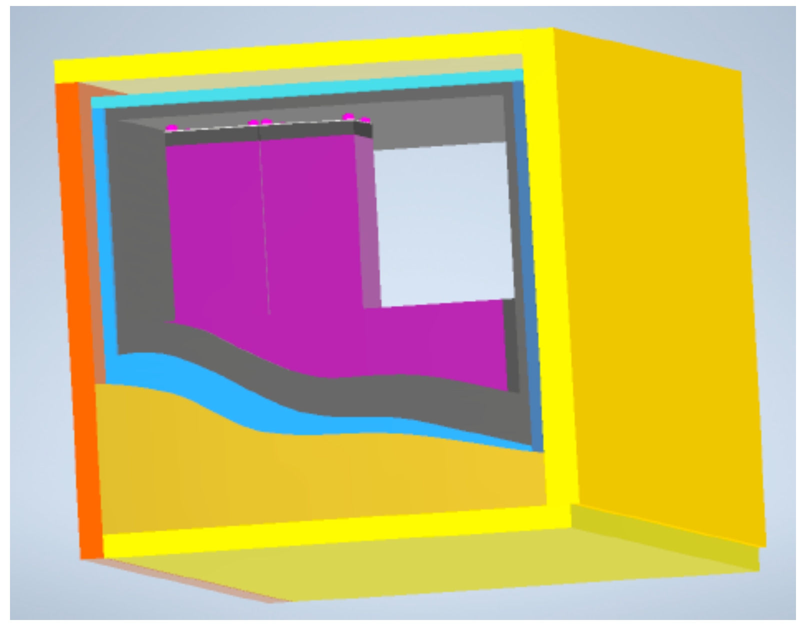
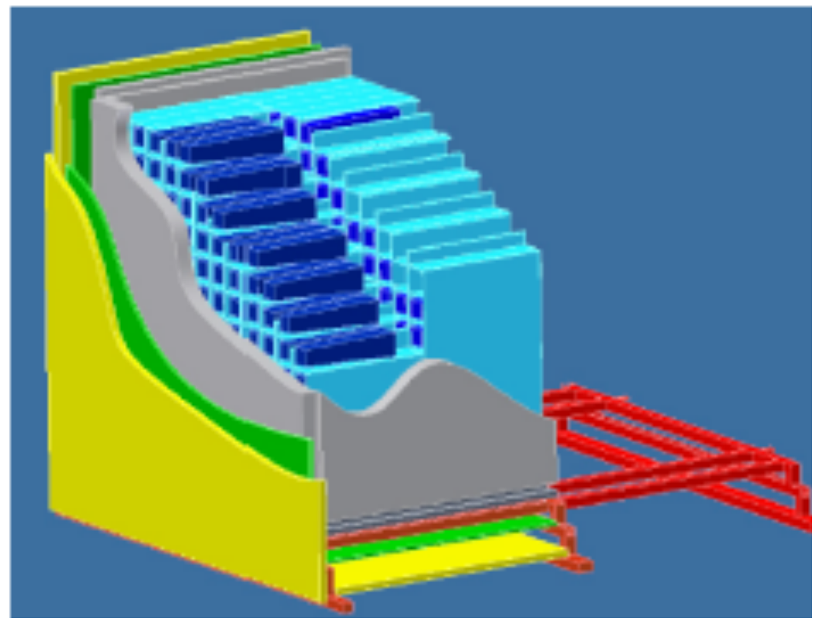




BDX Detector

BDX detector:

- ★ BDX detector: 8 modules made by an EMCal and a surrounding veto
- ★ EMCal (each module): 2x100 CsI(Tl) crystals from BaBar ECal fully refurbished with SiPMs and fADCs readout
- ★ The calorimeter is hermetically closed in a two-layer active veto (plastic scintillator) and passive (lead)

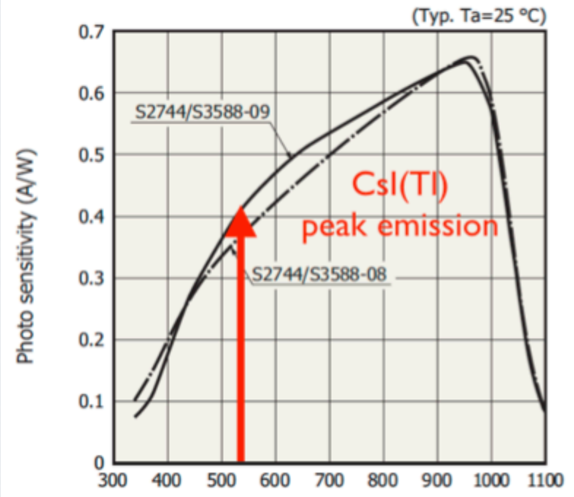
★ Calorimeter options B, C, D, …
- PANDA PbWO crystals
- BGO ex-GRAAL crystals (+BaF2 nex-TAPS?)
- CMS PbWO?

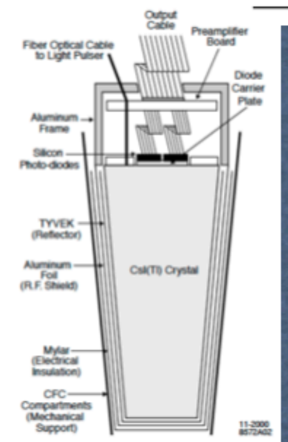
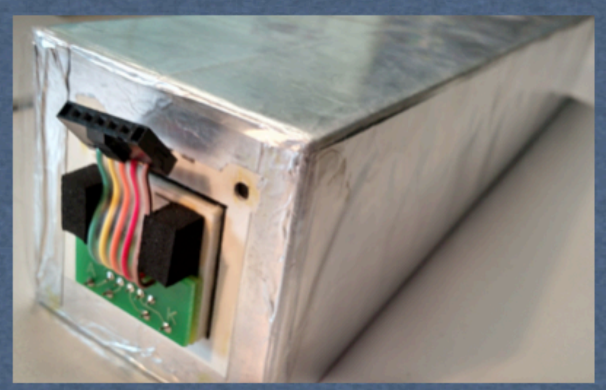
★ DAQ
- ~1000 fADC cos provided by Hall-B (PRAD-II exp)
- INFN FastElectronics designed a16 chs bias/preamp board
- SRO DAQ crate with 4 fADC boards available at JLab for test
- BDX DAQ server procured and under tests at INFN-GE

Reuse of BaBar crystals

BABAR em calorimeter
- ★ 6580 CsI(Tl) ~5(6)x5(6)x30cm³ tapered geometry
- ★ 820 end cups + 5760 barrel
- ★ 2x Hamamatsu S2744-08 silicon diodes readout, thermalized
- ★ ~3.9 10³ pe/MeV (with reduced shaping time)

Parameter	Values
Radiation length	1.85 cm
Molière radius	3.8 cm
Density	4.53 g/cm³
Light yield	50,000 γ/MeV
Light yield temp. coeff.	0.28%/°C
Peak emission λ_{max}	565 nm
Refractive index (λ_{max})	1.80
Signal decay time	680 ns (64%)
	3.34 μs (36%)



Radiation damage
- ★ Barrel exposed to 2.2krad
- ★ -14% LY after full operation
- ★ Thermal annealing not efficient
- ★ ~0.1%/day recovery they should be fine by now

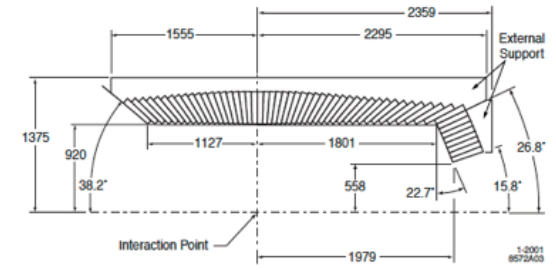
Trapezoidal cross-section

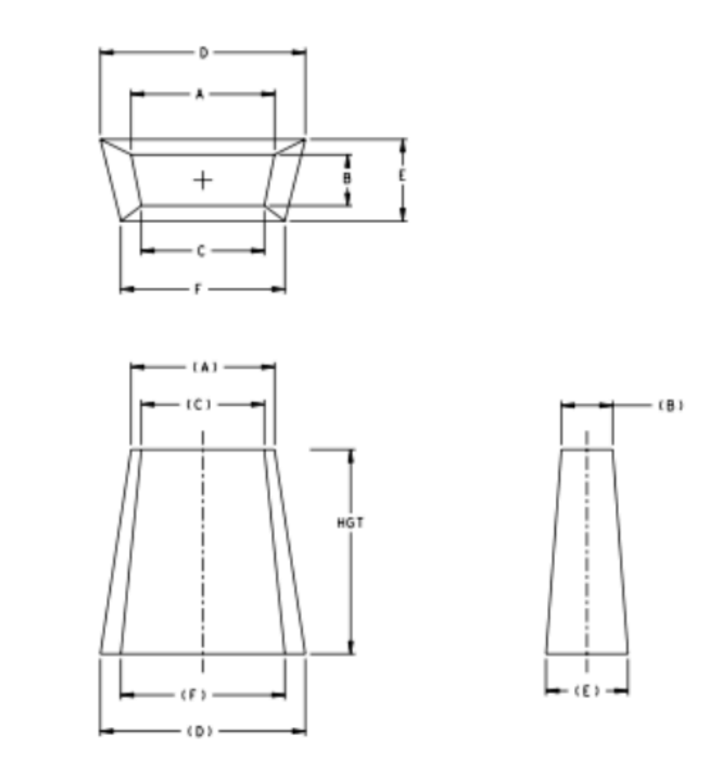
Front face : 4.7 x 4.7 cm²
Back face : 6 x 6 cm²

Crystal length (from Endcap): 30.5 cm/16.5 X_0 (80 crystals) and 32.4 cm/17.5 X_0

- ★ SLAC trashed the BaBar EM endcap
- ★ 1680 similar size crystals in the barrel
- ★ They still have the barrel but storage conditions are not know
- ★ Planned an on-site visit to check and plan the loan to JLab

Babar Calorimeter





Barrel

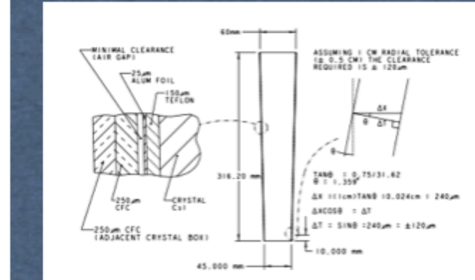
θ Row	Number Needed	Volume (cc)	A (cm)	B (cm)	C (cm)	D (cm)	E (cm)	F (cm)	Height (cm)
1	120	882.5	4.956	4.773	4.736	5.779	5.523	5.524	32.55
2	120	888.4	4.953	4.773	4.736	5.812	5.559	5.559	32.55
3	120	894.6	4.950	4.773	4.736	5.846	5.595	5.594	32.55
4	120	901.0	4.947	4.773	4.736	5.880	5.633	5.631	32.55
5	120	907.6	4.943	4.773	4.736	5.916	5.672	5.669	32.55
6	120	914.4	4.939	4.773	4.736	5.952	5.712	5.708	32.55
7	120	921.3	4.935	4.773	4.736	5.988	5.753	5.748	32.55
8	120	896.7	4.930	4.773	4.736	5.994	5.766	5.759	31.62
9	120	903.5	4.925	4.773	4.736	6.030	5.808	5.800	31.62
10	120	910.4	4.919	4.773	4.736	6.065	5.850	5.841	31.62
11	120	917.3	4.913	4.773	4.736	6.101	5.893	5.882	31.62
12	120	924.4	4.907	4.773	4.736	6.136	5.936	5.924	31.62
13	120	931.4	4.900	4.773	4.736	6.171	5.980	5.965	31.62
14	120	938.3	4.892	4.773	4.736	6.204	6.023	6.007	31.62
15	120	910.6	4.884	4.773	4.736	6.197	6.028	6.009	30.69
16	120	916.9	4.876	4.773	4.736	6.226	6.069	6.048	30.69
17	120	922.9	4.867	4.773	4.736	6.254	6.108	6.086	30.69
18	120	928.7	4.857	4.773	4.736	6.279	6.147	6.123	30.69
19	120	934.2	4.847	4.773	4.736	6.302	6.184	6.158	30.69
20	120	939.3	4.837	4.773	4.736	6.322	6.220	6.191	30.69
21	120	943.9	4.826	4.773	4.736	6.339	6.253	6.222	30.69
22	120	911.3	4.814	4.773	4.736	6.306	6.237	6.204	29.76
23	120	914.6	4.802	4.773	4.736	6.315	6.264	6.228	29.76
24	120	917.2	4.790	4.773	4.736	6.320	6.287	6.248	29.76
25	120	919.2	4.778	4.773	4.736	6.321	6.307	6.266	29.76
26	120	920.6	4.765	4.773	4.736	6.318	6.322	6.279	29.76
27	120	921.2	4.752	4.773	4.736	6.310	6.334	6.289	29.76
28	120	882.4	4.740	4.773	4.736	6.299	5.896	6.294	29.76
29	120	882.4	4.740	4.773	4.736	6.299	5.896	6.294	29.76
30	120	921.2	4.752	4.773	4.736	6.310	6.334	6.289	29.76
31	120	920.6	4.765	4.773	4.736	6.318	6.322	6.279	29.76
32	120	919.2	4.778	4.773	4.736	6.321	6.307	6.266	29.76
33	120	917.2	4.790	4.773	4.736	6.320	6.287	6.248	29.76
34	120	914.6	4.802	4.773	4.736	6.315	6.264	6.228	29.76
35	120	911.3	4.814	4.773	4.736	6.306	6.237	6.204	29.76
36	120	907.5	4.826	4.773	4.736	6.293	6.208	6.177	29.76
37	120	903.2	4.837	4.773	4.736	6.277	6.176	6.147	29.76
38	120	898.4	4.847	4.773	4.736	6.258	6.142	6.115	29.76
39	120	893.4	4.857	4.773	4.736	6.236	6.106	6.081	29.76
40	120	888.0	4.867	4.773	4.736	6.212	6.068	6.045	29.76
41	120	882.3	4.876	4.773	4.736	6.185	6.029	6.008	29.76
42	120	876.5	4.884	4.773	4.736	6.157	5.990	5.971	29.76
43	120	870.5	4.892	4.773	4.736	6.127	5.949	5.932	29.76
44	120	864.4	4.900	4.773	4.736	6.096	5.909	5.893	29.76
45	120	858.3	4.907	4.773	4.736	6.064	5.868	5.854	29.76
46	120	852.2	4.913	4.773	4.736	6.031	5.827	5.815	29.76
47	120	846.1	4.919	4.773	4.736	5.998	5.787	5.776	29.76
48	120	840.0	4.925	4.773	4.736	5.965	5.747	5.737	29.76
49	120	834.1	4.930	4.773	4.736	5.931	5.708	5.699	29.76

Endcup

θ Row	Number Needed	Volume (cc)	A (cm)	B (cm)	C (cm)	D (cm)	E (cm)	F (cm)	Height (cm)
1	80	741.0	4.298	4.670	3.944	4.995	5.426	4.583	32.55
2	80	802.4	4.641	4.670	4.289	5.394	5.428	4.985	32.55
3	80	863.6	4.982	4.670	4.632	5.792	5.430	5.386	32.55
4	100	739.6	4.247	4.670	3.970	4.941	5.432	4.618	32.55
5	100	788.3	4.519	4.670	4.244	5.257	5.433	4.937	32.55
6	100	836.9	4.790	4.670	4.517	5.573	5.433	5.256	32.55
7	120	737.7	4.209	4.670	3.984	4.898	5.434	4.636	32.55
8	120	778.4	4.437	4.670	4.214	5.162	5.433	4.903	32.55
9	120	819.1	4.665	4.670	4.445	5.427	5.432	5.171	32.55

θ Interval (radians)	Length (X_0)	# Rings	Crystals /Ring
Barrel			
2.456 – 1.214	16.0	27	120
1.213 – 0.902	16.5	7	120
0.901 – 0.655	17.0	7	120
0.654 – 0.473	17.5	7	120
Endcap			
0.469 – 0.398	17.5	3	120
0.397 – 0.327	17.5	3	100
0.326 – 0.301	17.5	1	80
0.300 – 0.277	16.5	1	80

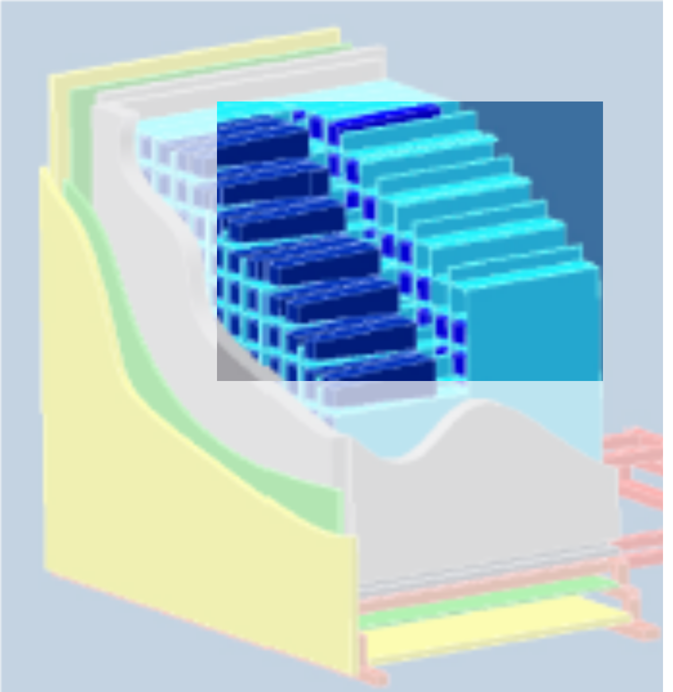
Tolerances



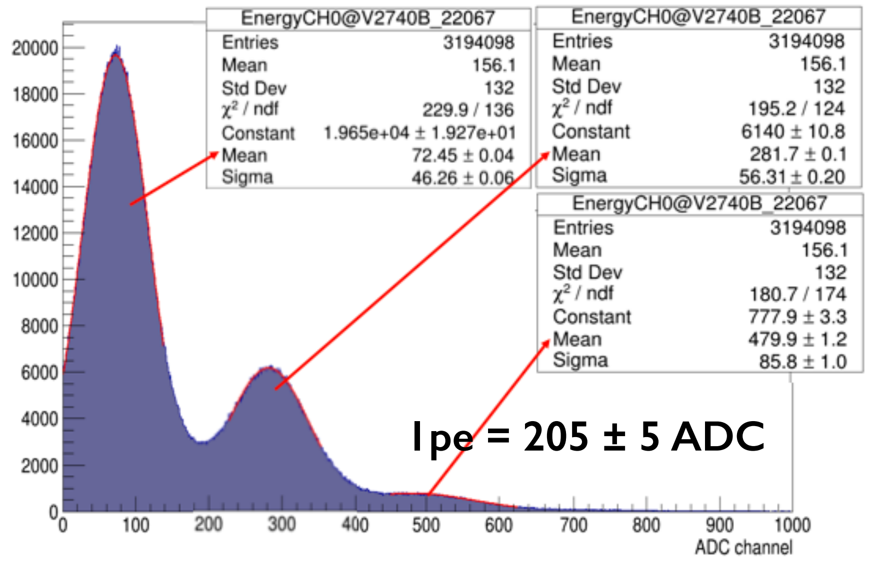
BDX DAQ

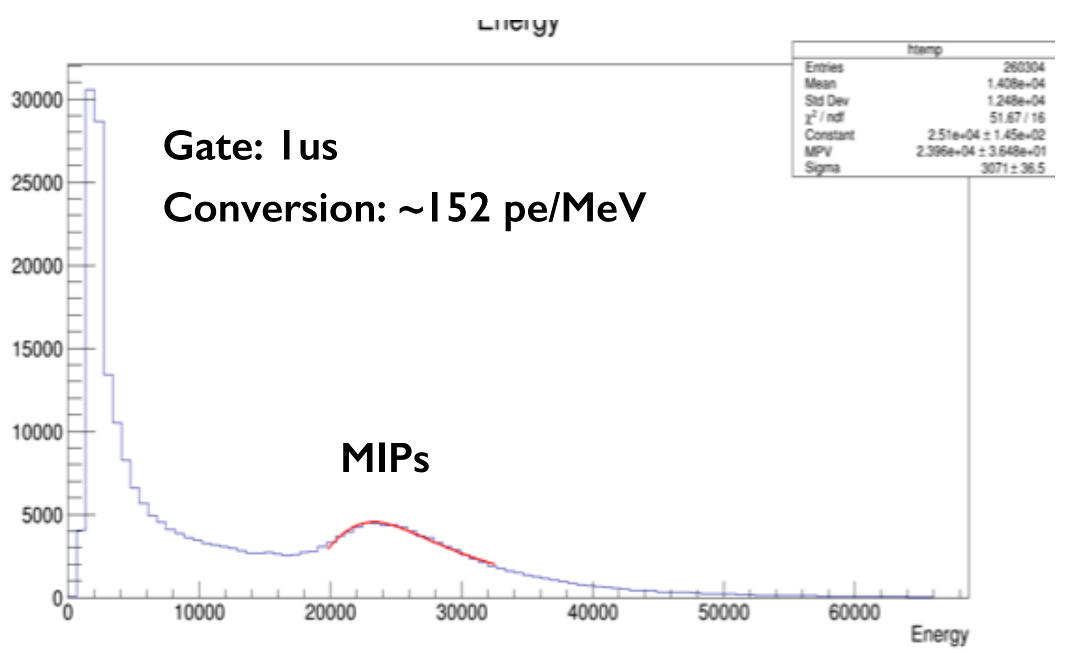
- No amplification, MCX-MCX to MCX-3M patch panel

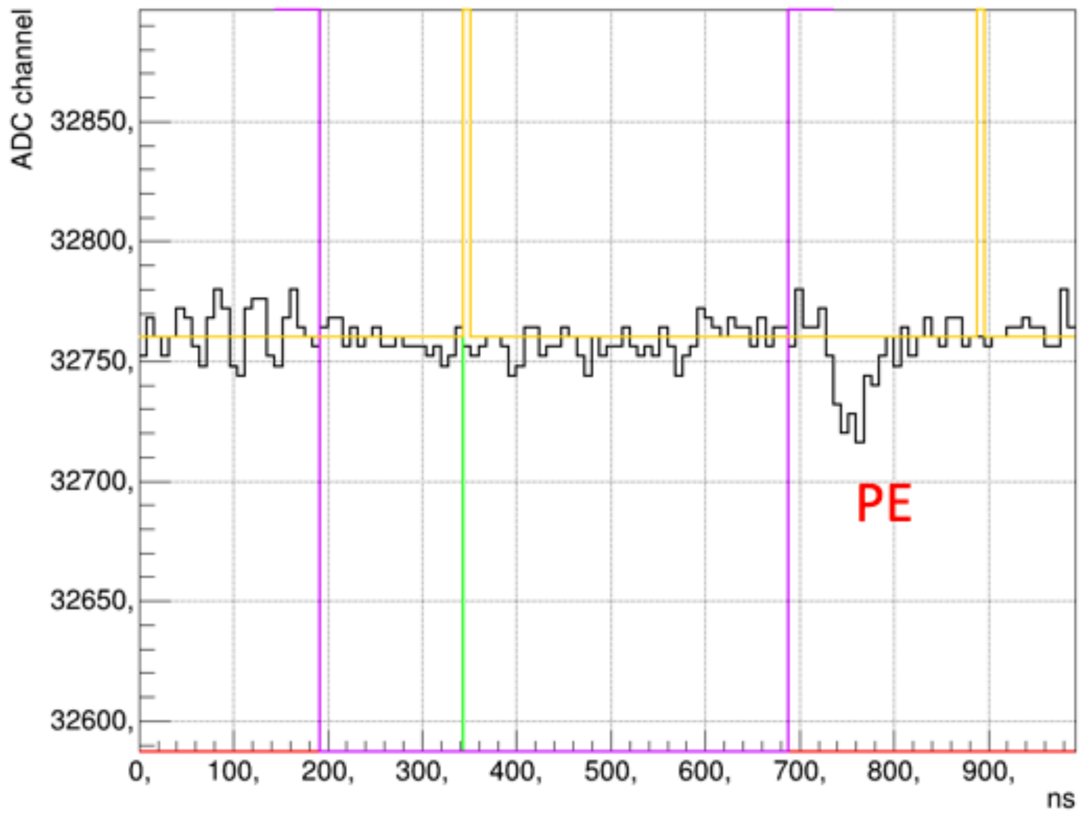
BDX DAQ

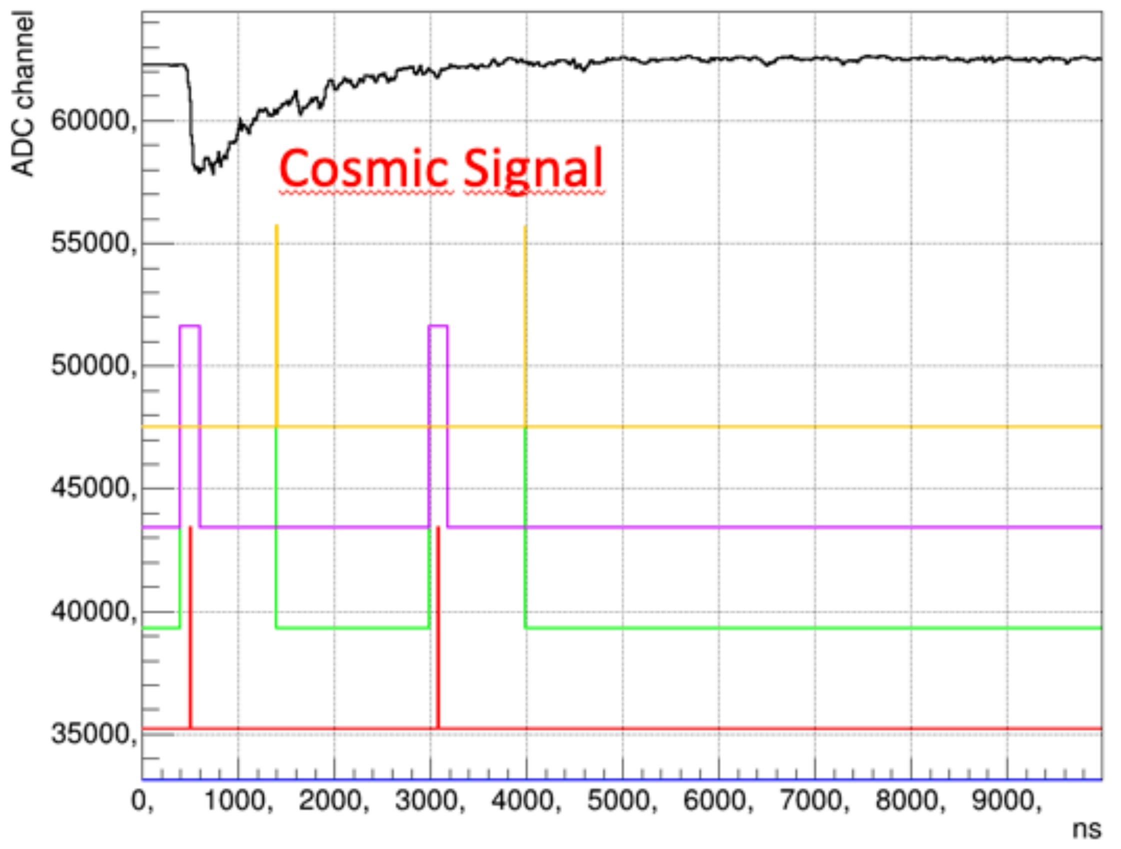


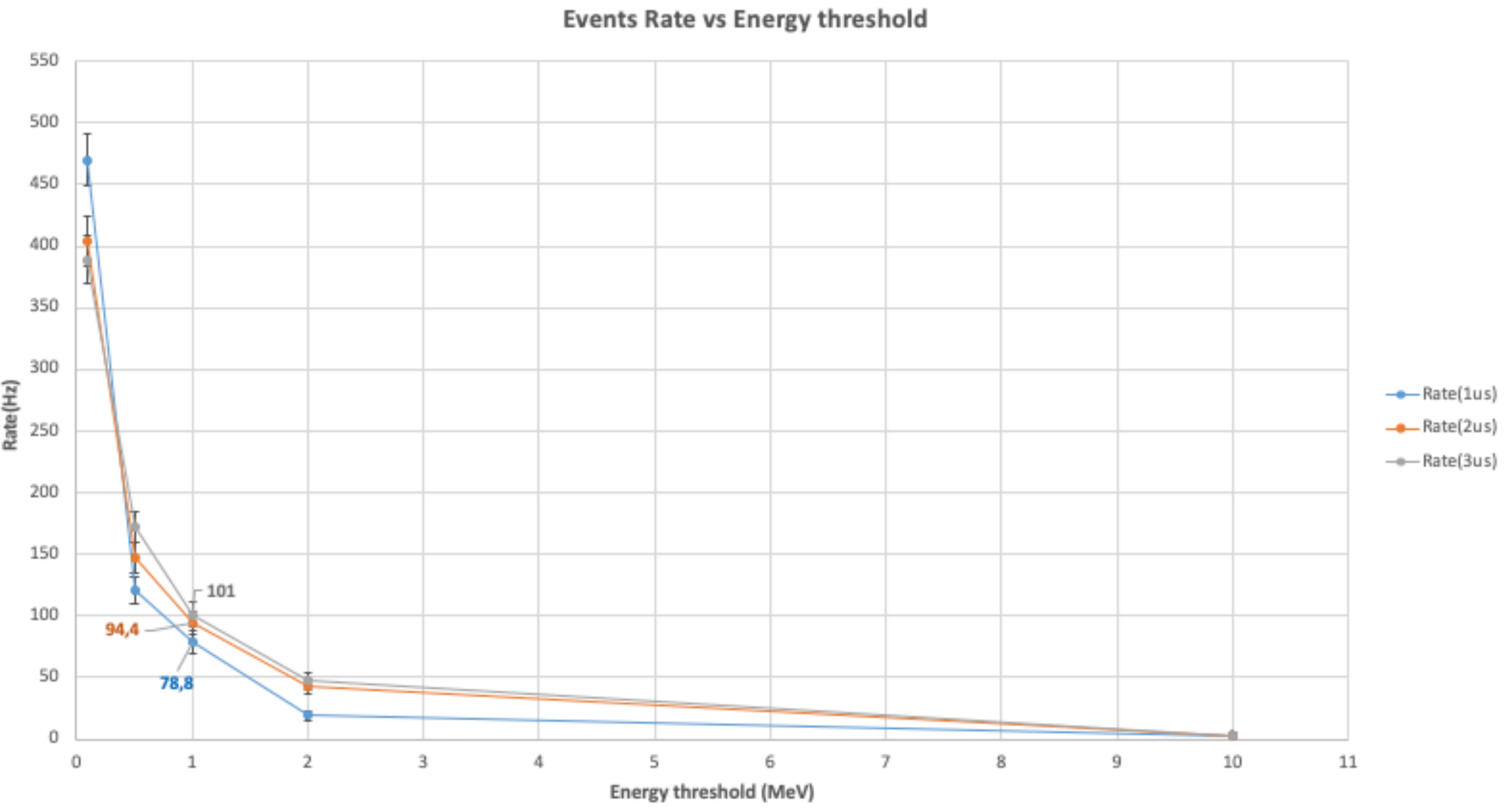
Calorimeter: x200 6x6 S13360-6075 SiPM, CsI(Tl) crystals

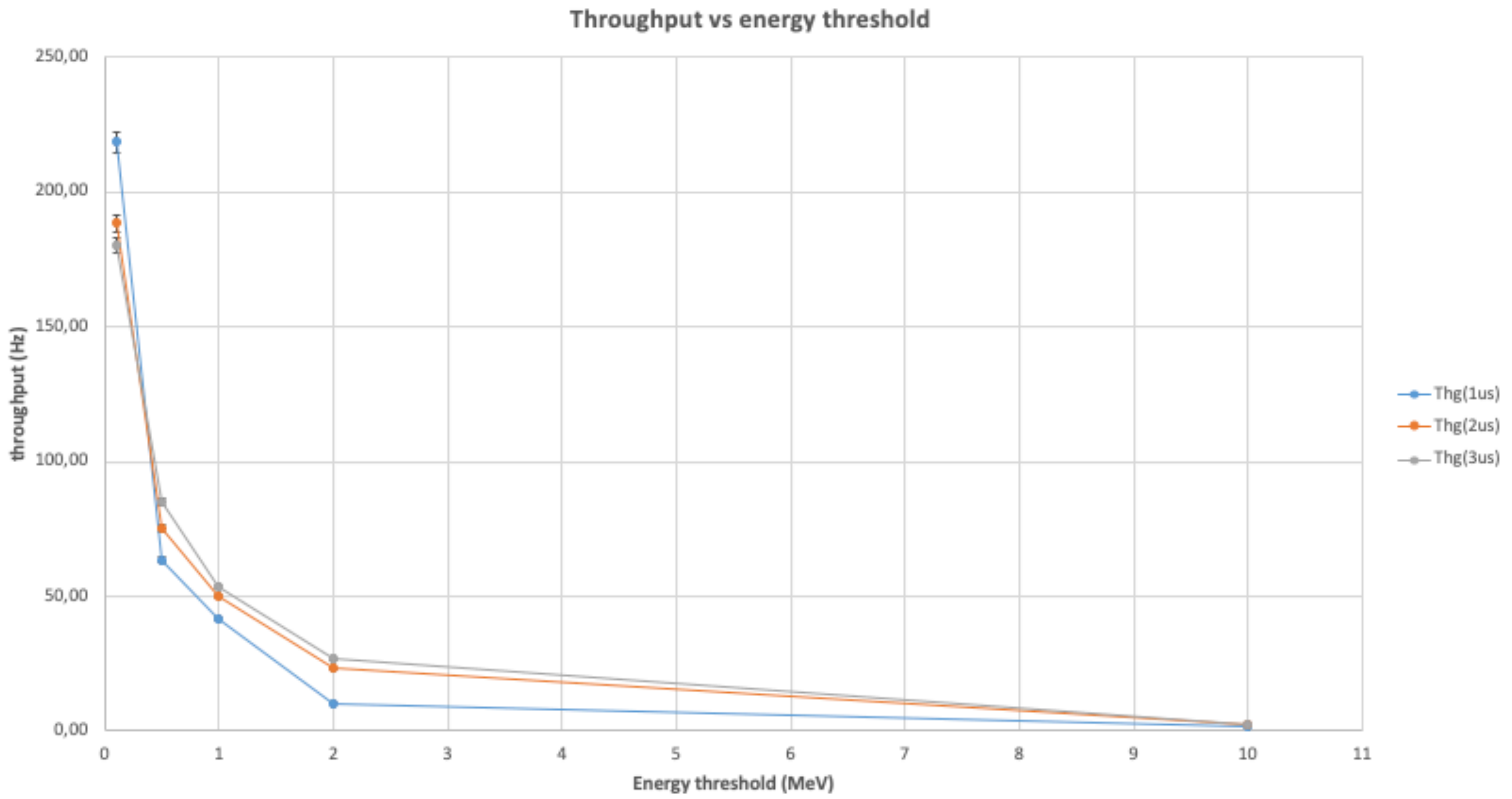


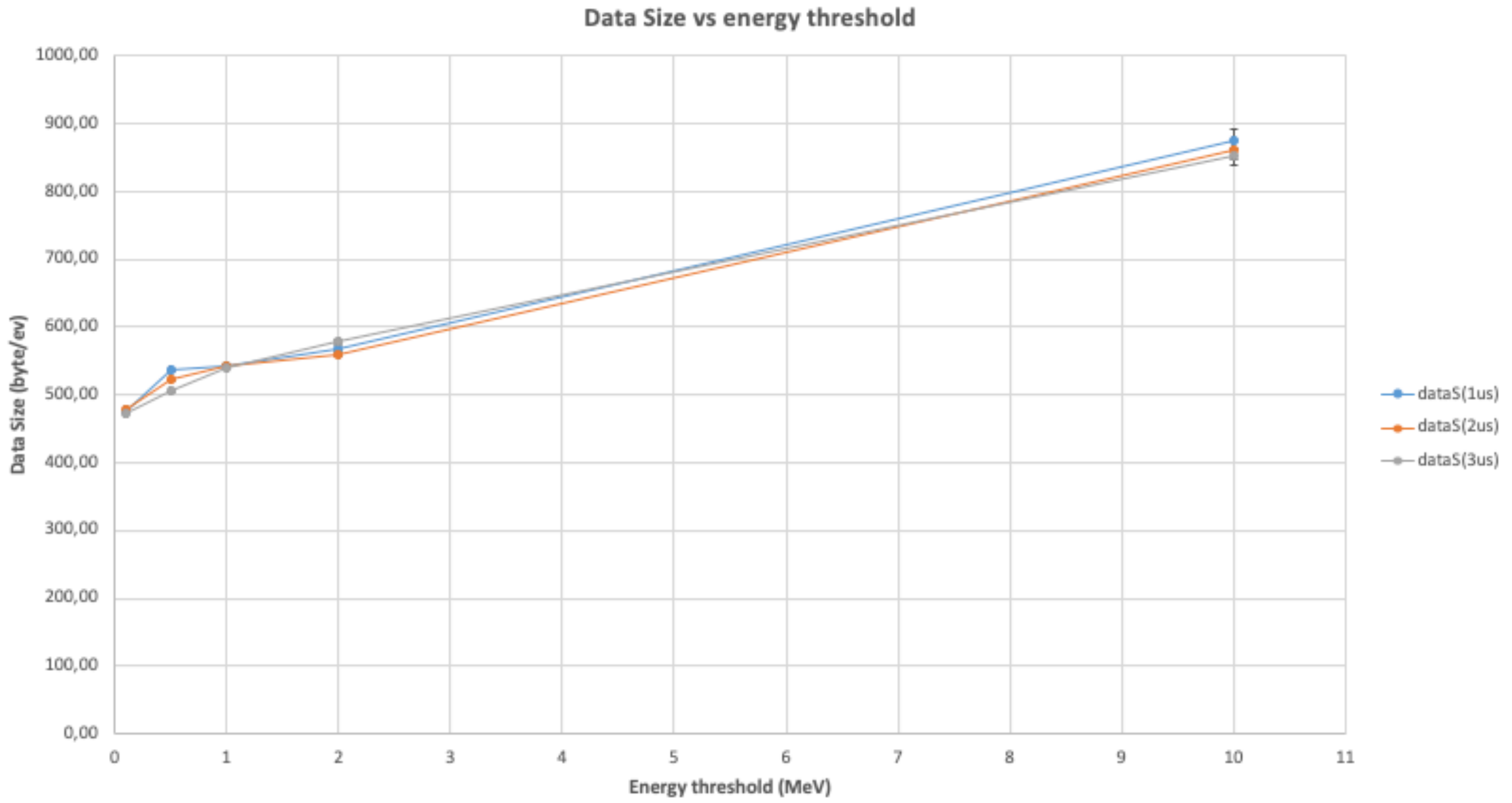












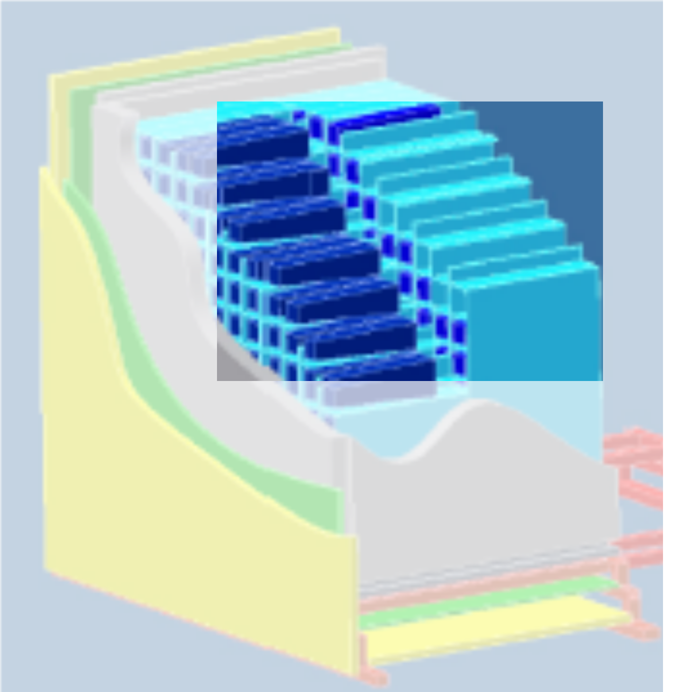
2 MeV threshold (~300 pe)

1 crystal
R_{Events} ~ 20 Hz
$R_{Data\text{-}NO\text{-}WF}$ (2us time window) ~ 0.5 kB/s
$R_{Data\text{-}WF}$ (2us time window) ~ 30 kB/s

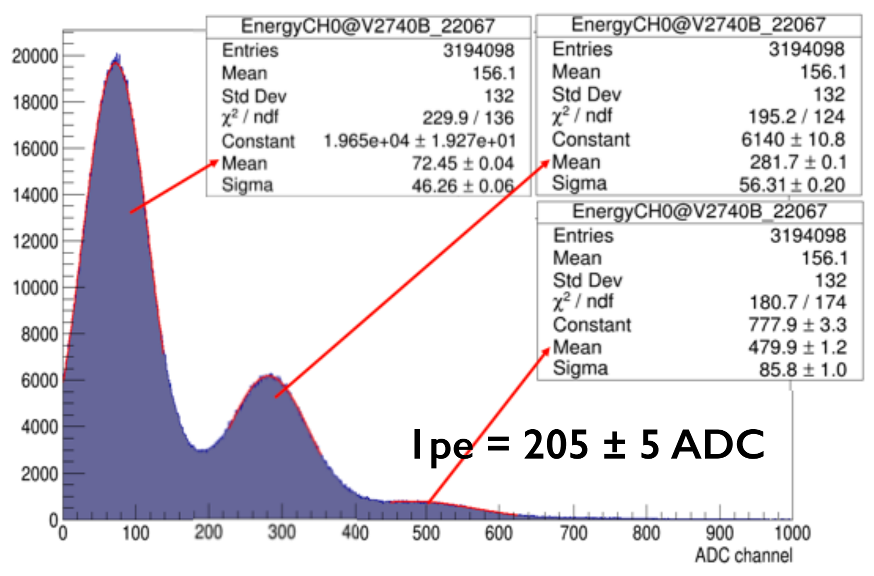
1 Module (200 crystals)
R_{Events} ~ 4 kHz
$R_{Data\text{-}NO\text{-}WF}$ (2us time window) ~ 0.1 MB/s
$R_{Data\text{-}WF}$ (2us time window) ~ 6 MB/s

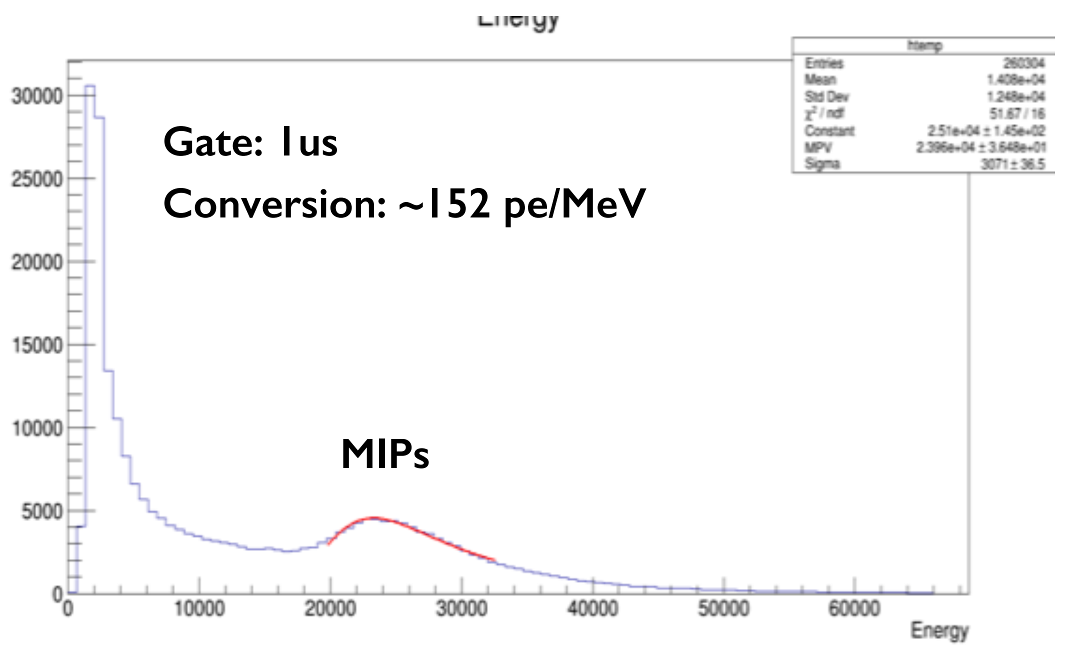
Whole calorimeter (~1000 crystals)
R_{Events} ~ 20 kHz
$R_{Data\text{-}NO\text{-}WF}$ (2us time window) ~ 0.5 MB/s
$R_{Data\text{-}WF}$ (2us time window) ~ 30 MB/s

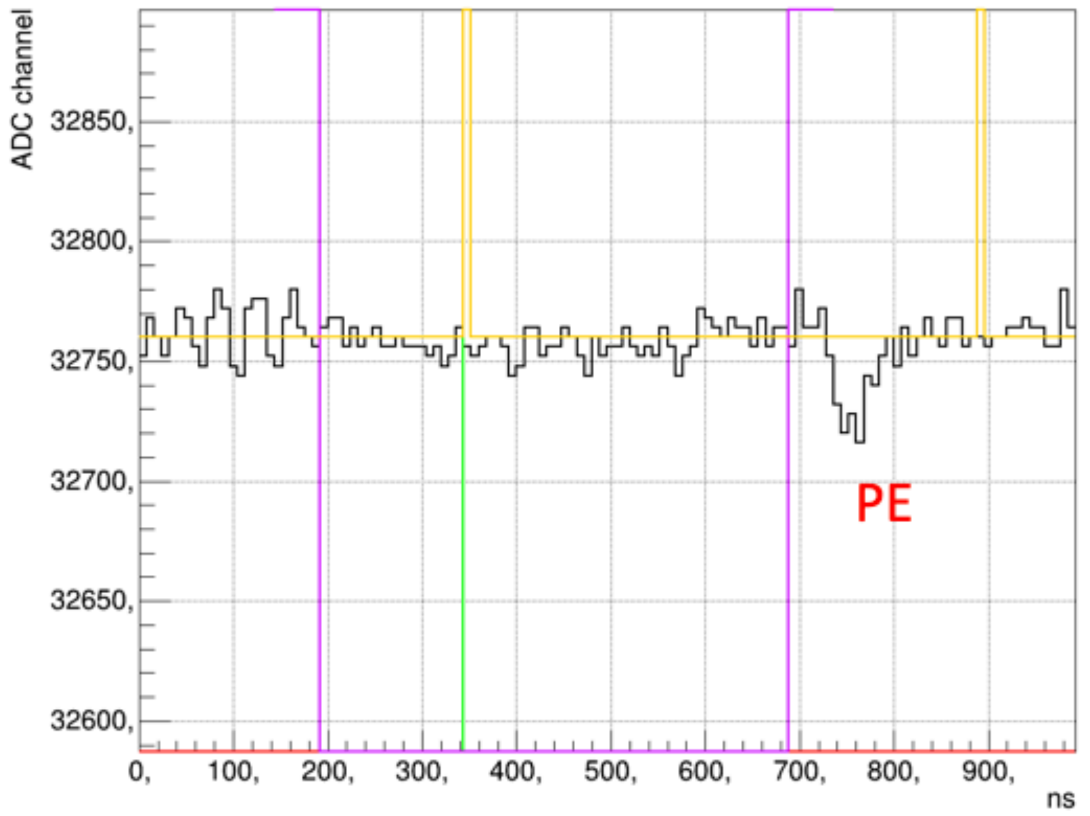
BDX DAQ

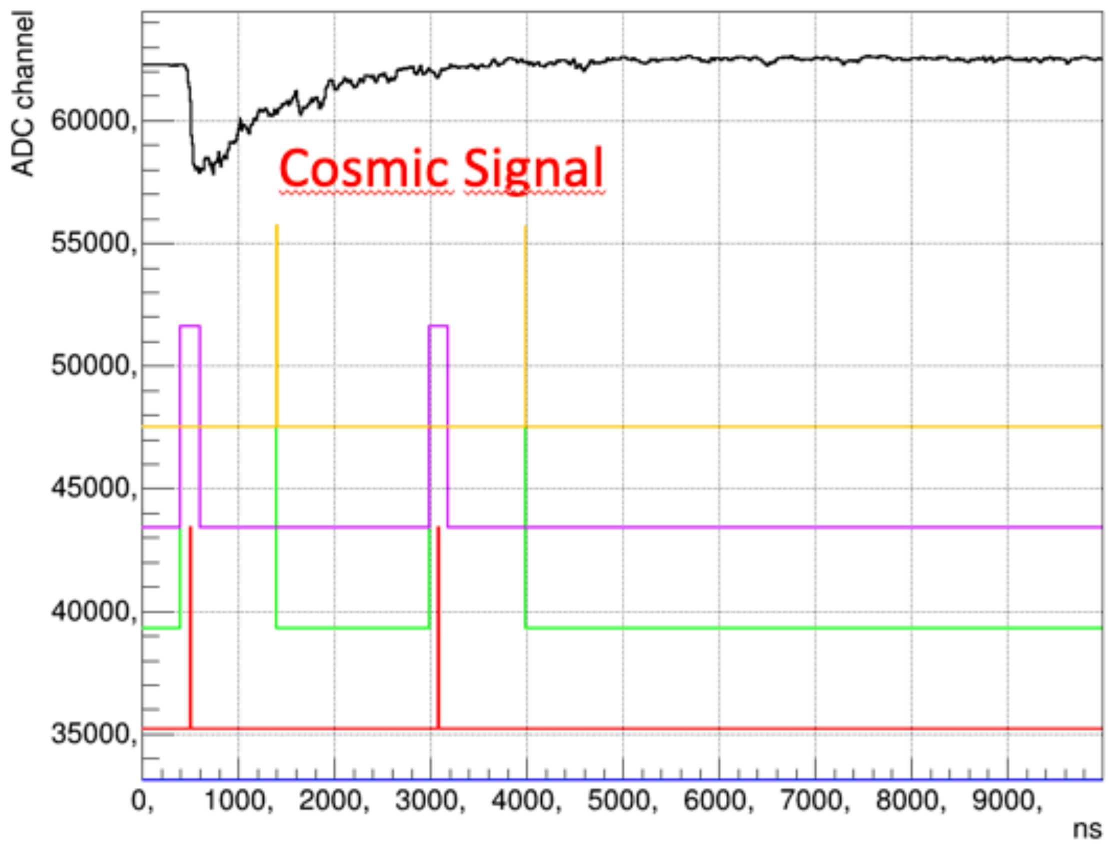


Calorimeter: x200 6x6 S13360-6075 SiPM, CsI(Tl) crystals









Event and data rates (no waveform saved)

Caen Digitalizer v2470b 64 ch, 16 bit, 125 Gs/s

Thr(V) mV	Thr(E) MeV	Rate(1us) Hz	Rate(2us) Hz	Rate(3us) Hz	dataS(1us) byte/ev	dataS(2us) byte/ev	dataS(3us) byte/ev	Thg(1us) kB/s	Thg(2us) kB/s	Thg(3us) kB/s
5	0,1	556,4	499,1	460	8,30	8,40	8,52	4,52	4,10	3,83
	0,2	403,2	391,4	391,7	8,49	8,61	8,65	3,35	3,30	3,32
	0,5	151,5	176,5	219,3	9,41	9,34	9,15	1,40	1,62	1,97
	1	85,29	104,6	124,7	9,03	9,00	8,97	0,76	0,93	1,10
	2	17,48	44,47	50,06	9,52	9,15	9,12	0,17	0,41	0,45
	5	2,95	3,52	4,23	8,16	8,19	8,76	0,03	0,04	0,04
	10	2,485	2,72	2,68	8,24	8,78	8,92	0,03	0,03	0,03

Event and data rates (with waveform)

Thr(V) mV	Thr(E) MeV	Rate(1us) Hz	Rate(2us) Hz	Rate(3us) Hz	dataS(1us) byte/ev	dataS(2us) byte/ev	dataS(3us) byte/ev	Thg(1us) kB/s	Thg(2us) kB/s	Thg(3us) kB/s
5	0,1	470	404	389	475,67	477,34	473,81	218,33	188,33	180,00
	0,2									
	0,5	121	147	172	535,91	522,40	506,00	63,33	75,00	85,00
	1	78,8	94,4	101	541,36	542,29	540,65	41,67	50,00	53,33
	2	18,73	42,63	47,23	569,09	560,30	578,00	10,42	23,33	26,67
	5									
	10	2,367	2,45	3,117	876,41	860,65	851,69	2,03	2,07	2,60

2 MeV threshold (~300 pe)

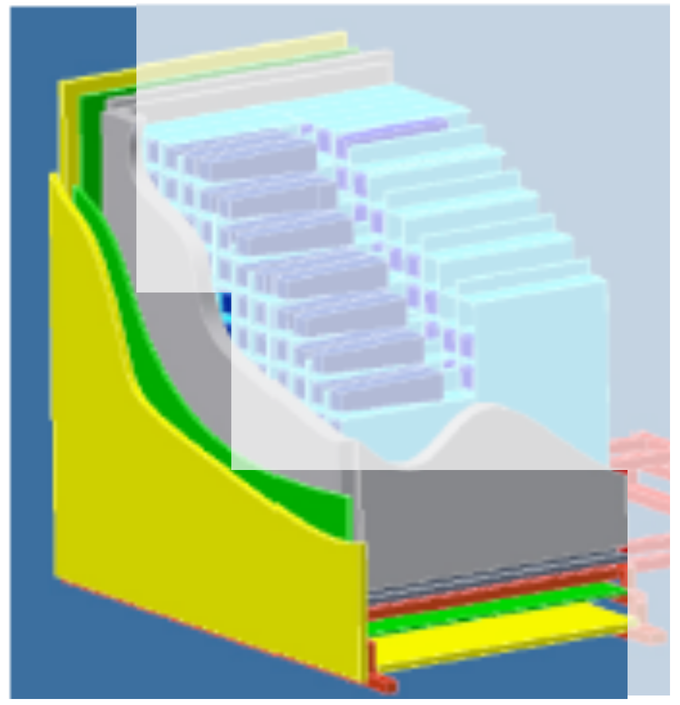
1 crystal
R_{Events} ~ 20 Hz
$R_{Data\text{-}NO\text{-}WF}$ (2us time window) ~ 0.5 kB/s
$R_{Data\text{-}WF}$ (2us time window) ~ 30 kB/s

1 Module (200 crystals)
R_{Events} ~ 4 kHz
$R_{Data\text{-}NO\text{-}WF}$ (2us time window) ~ 0.1 MB/s
$R_{Data\text{-}WF}$ (2us time window) ~ 6 MB/s

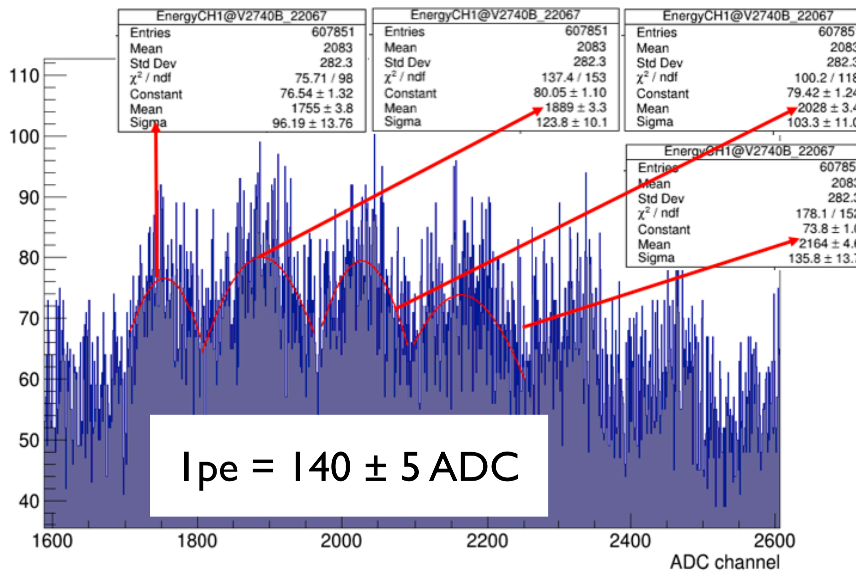
Whole calorimeter (~1000 crystals)
R_{Events} ~ 20 kHz
$R_{Data\text{-}NO\text{-}WF}$ (2us time window) ~ 0.5 MB/s
$R_{Data\text{-}WF}$ (2us time window) ~ 30 MB/s

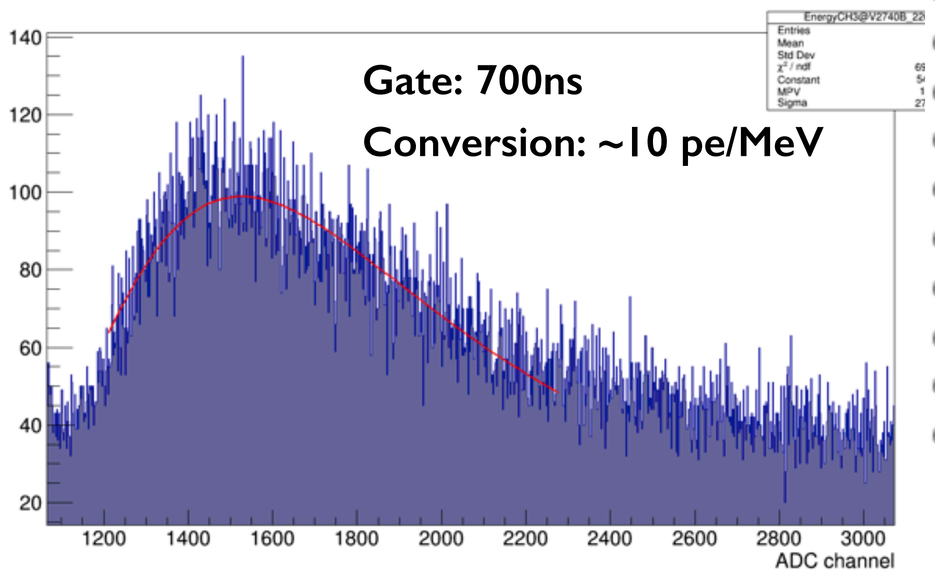
BDX DAQ

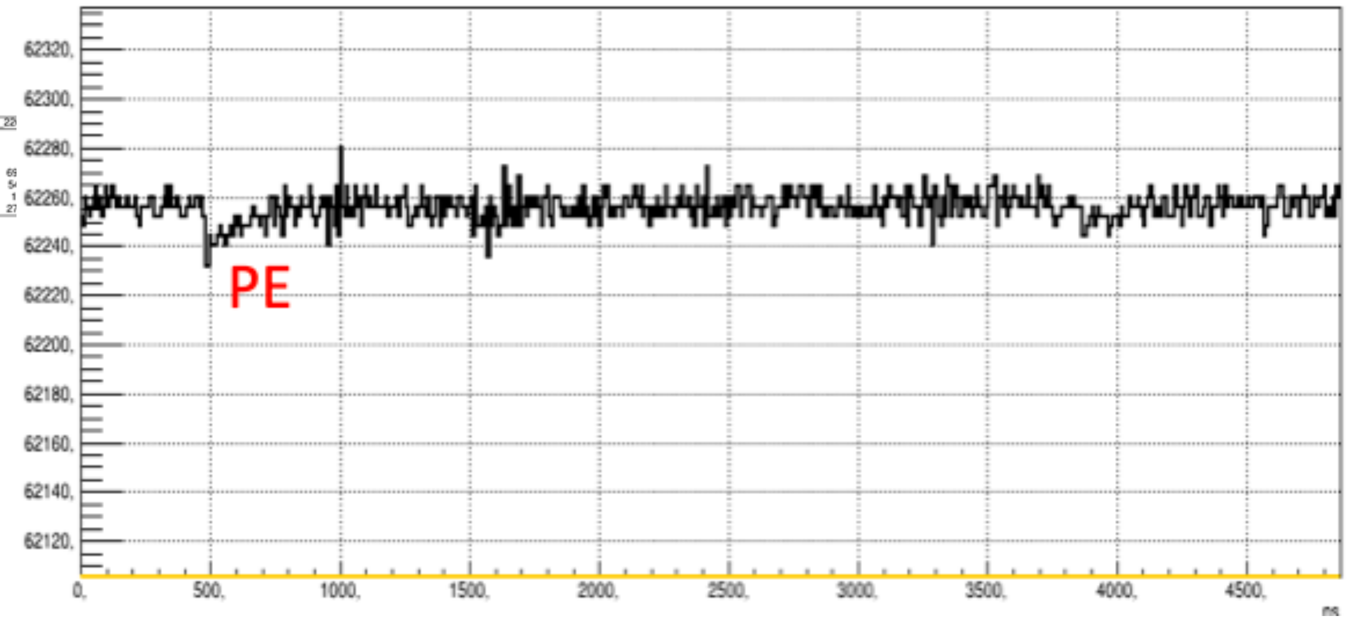
MIPs

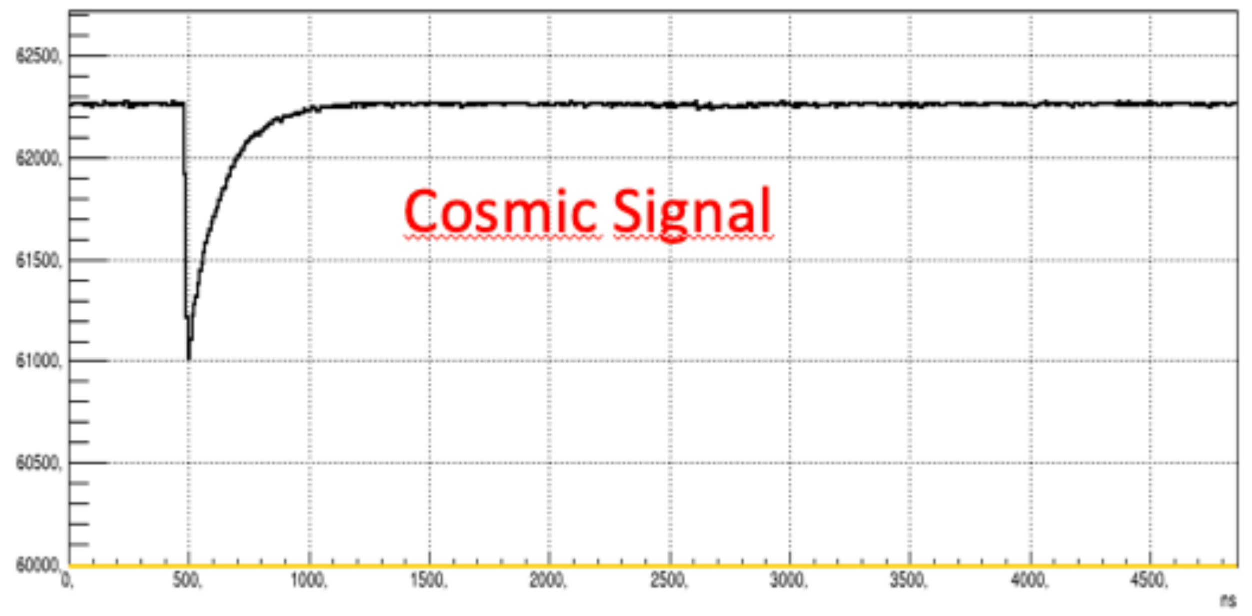


Veto: x 64 3x3 S13360-3075, SiPM, Plastic Scintillator









Event and data rates (with waveform)

Caen Digitalizer v2470b 64 ch, 16 bit, 125 Gs/s

Thr(V) mV	Thr(E) MeV	Rate(700ns) Hz	Rate(2us) Hz	Rate(3us) Hz	dataS(1us) byte/ev	dataS(2us) byte/ev	dataS(3us) byte/ev	Thg(1us) kb/m	Thg(2us) kb/m	Thg(3us) kb/m
5 mv	0,1	3698	4564	5681	423,66	427,41	430,20	91800	114300	143200
	0,2	946	1760	2575	416,74	448,00	448,04	23100	46200	67600
	0,5	120	153	202	422,34	441,68	427,47	2970	3960	5060
	1	99	102	97,8	408,83	425,42	385,75	2372	2543	2211
	2	39	44,6	44,7	259,74	255,06	264,80	594	667	694
	5	8,71	9,22	9,3	197,02	217,59	221,22	101	118	121
	10									

PE	Thr(E) keV	Rate(500ns) kHz	Rate(700ns) Hz	dataS(500ns) byte/ev	dataS(700ns) byte/ev	Thg(500ns) MB/s	Thg(700ns) MB/s
1	50	10,6		441		4,68	
3	150	4,15		455		1,89	
5	250	2,06		465		0,96	
7	350	1,44		478		0,69	
9	450	0,89		487		0,43	
11	550	0,387		490		0,19	
13	650	0,257		492		0,13	
15	750	0,165		495		0,081	

0.15 keV threshold (~3 pe)

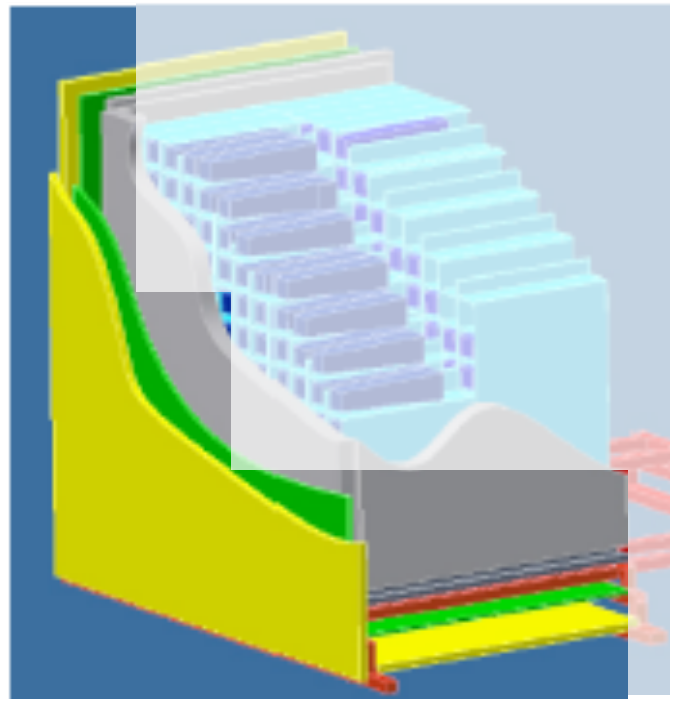
1 SiPM
R_{Events} ~ 5 kHz
$R_{Data\text{-}WF}$ (500ns time window) ~ 2 MB/s

1 Module (64 chs)
R_{Events} ~ 320 kHz
$R_{Data\text{-}WF}$ (500ns time window) ~ 130 MB/s

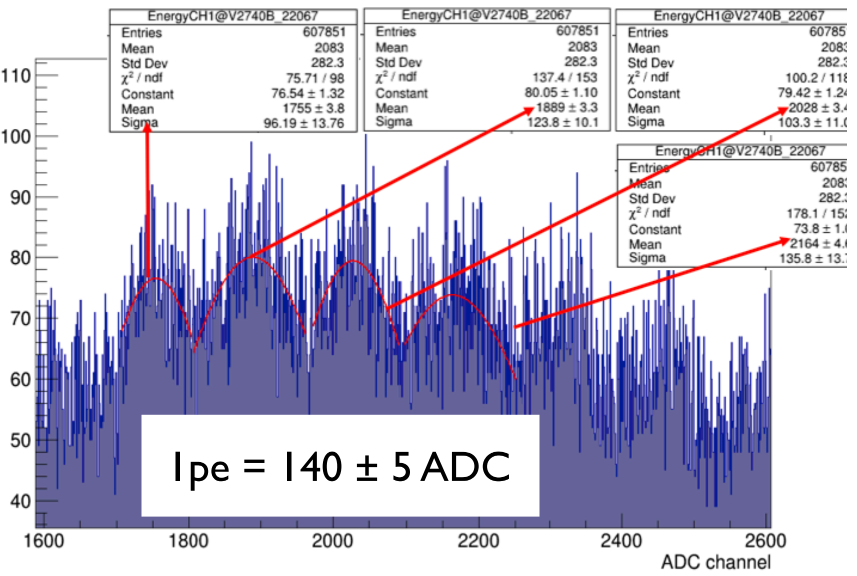
Whole Veto (~256 cos)
R_{Events} ~ 1.3 MHz
$R_{Data\text{-}WF}$ (500ns time window) ~ 0.5 GB/s

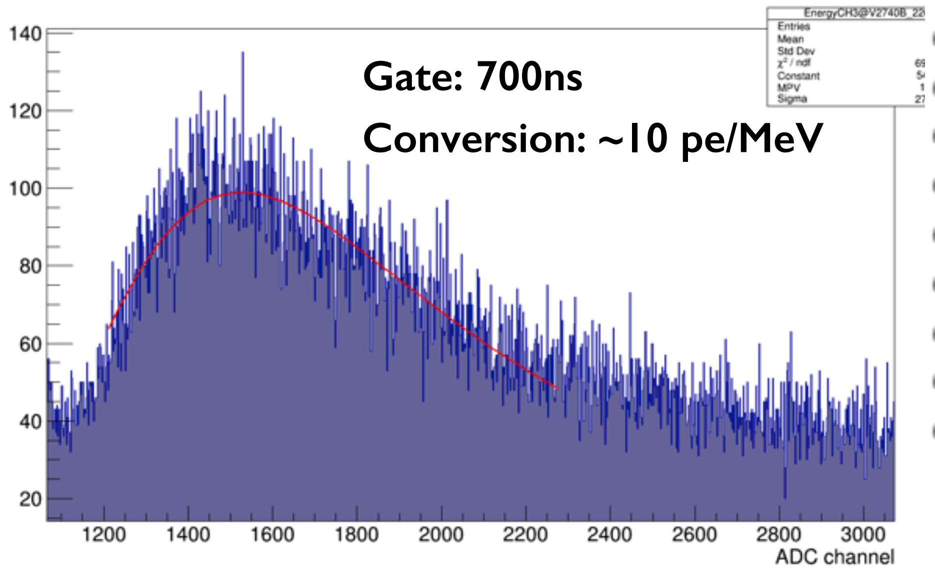
BDX DAQ

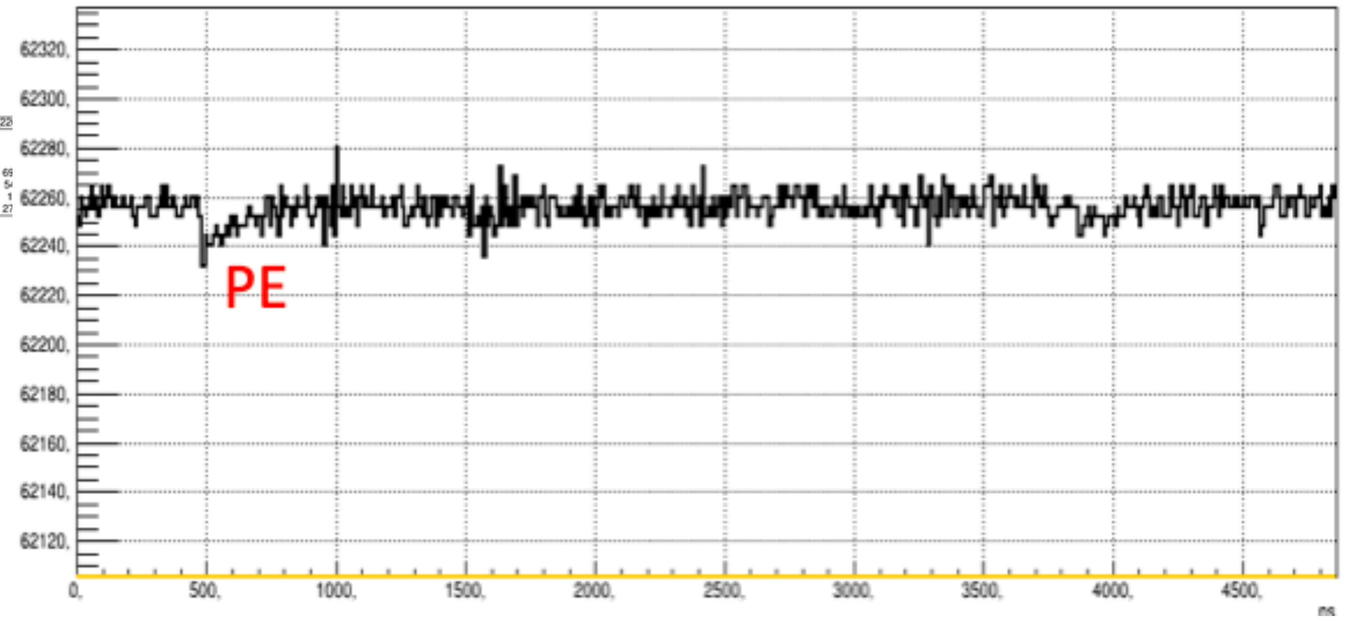
MIPs

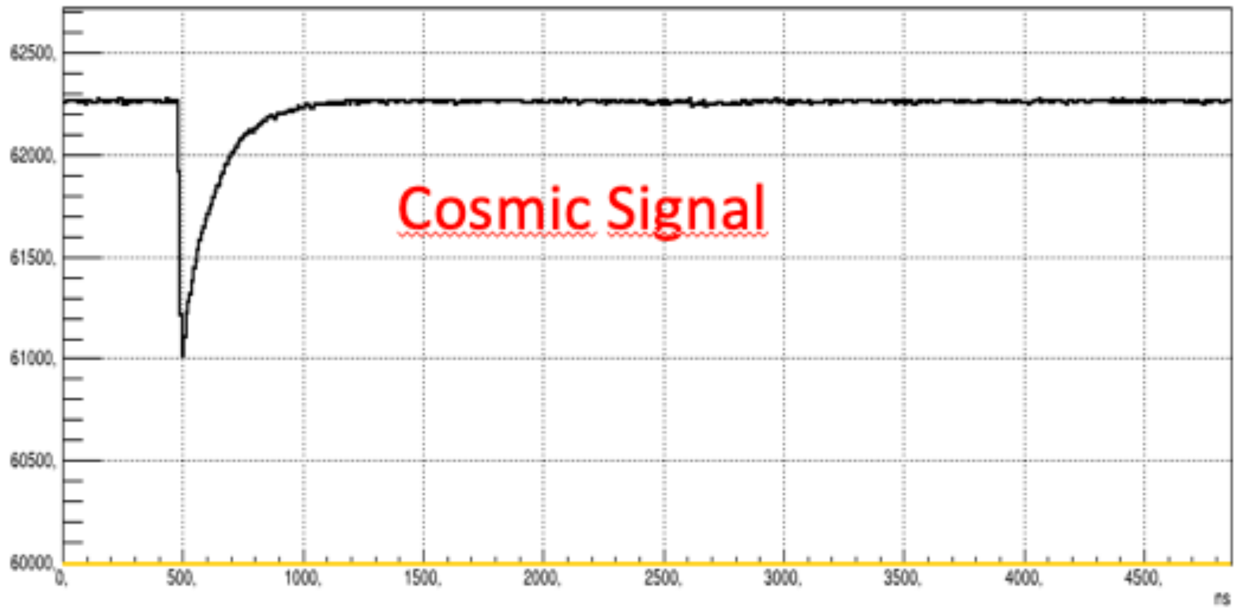


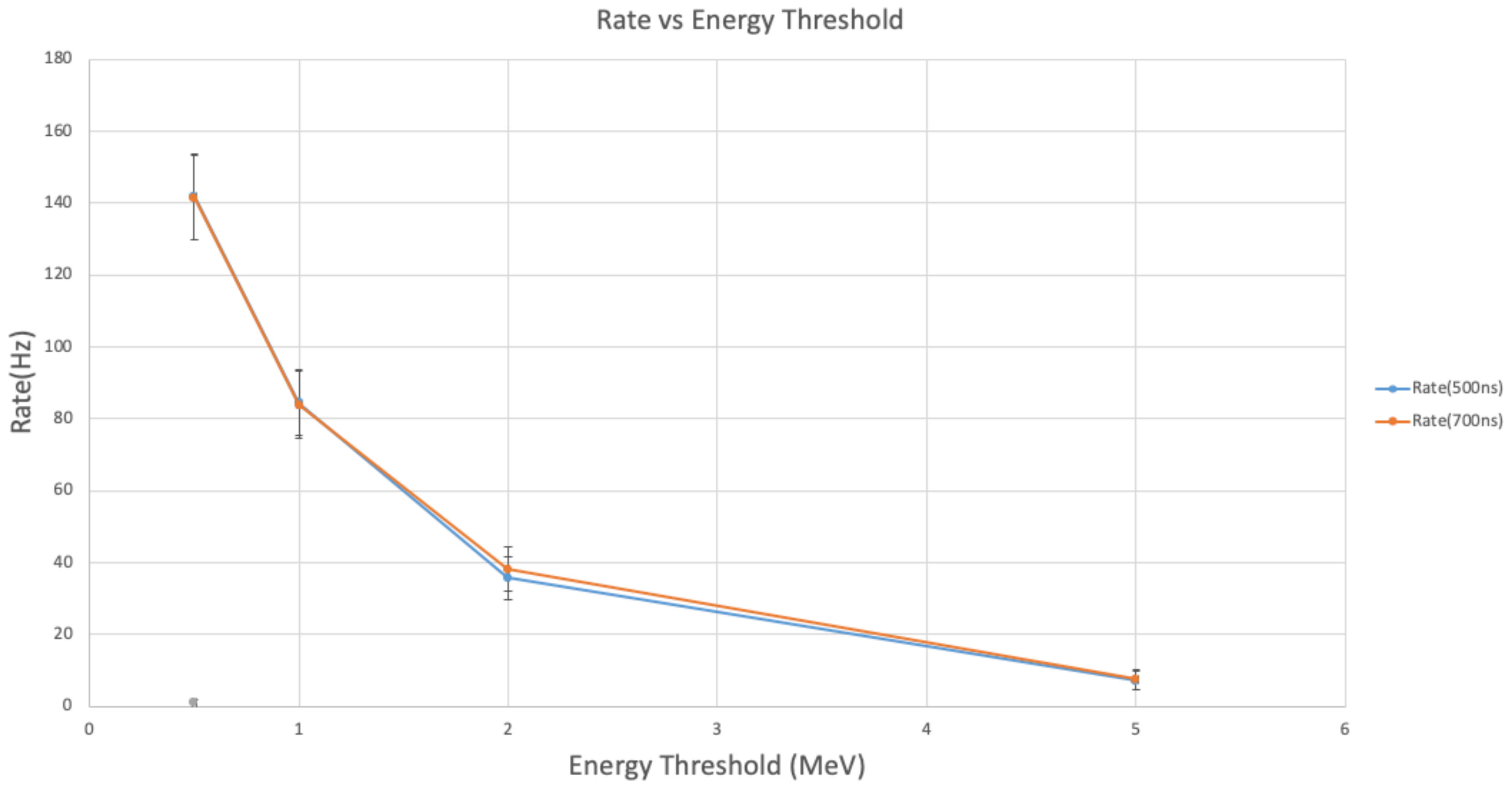
Veto: x 64 3x3 S13360-3075, SiPM, Plastic Scintillator

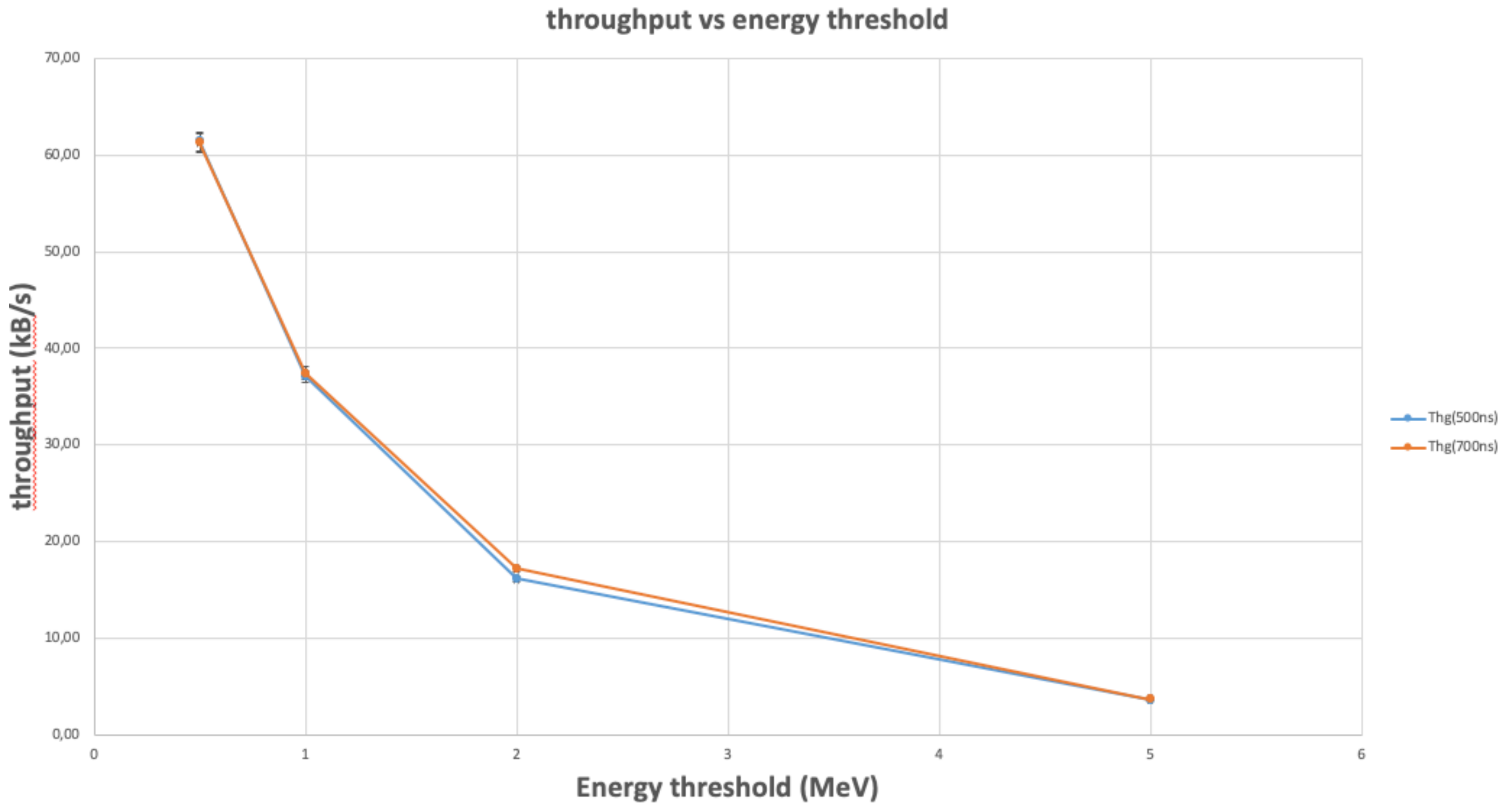


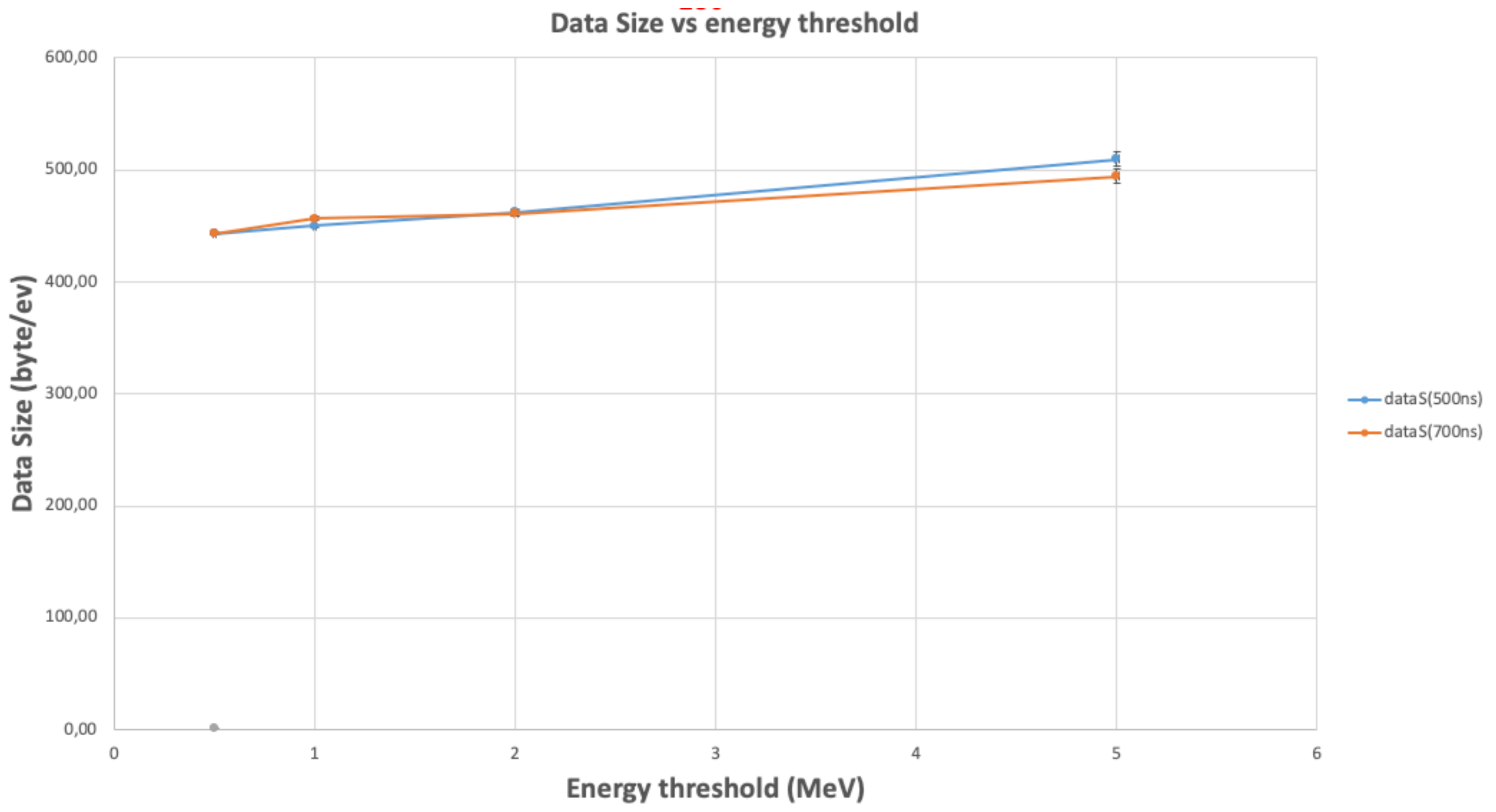












Caen Digitalizer v2470b 64 ch, 16 bit, 125 Gs/s

0.15 keV threshold (~3 pe)

1 SiPM
R_{Events} ~ 5 kHz
$R_{Data-WF}$ (500ns time window) ~ 2 MB/s

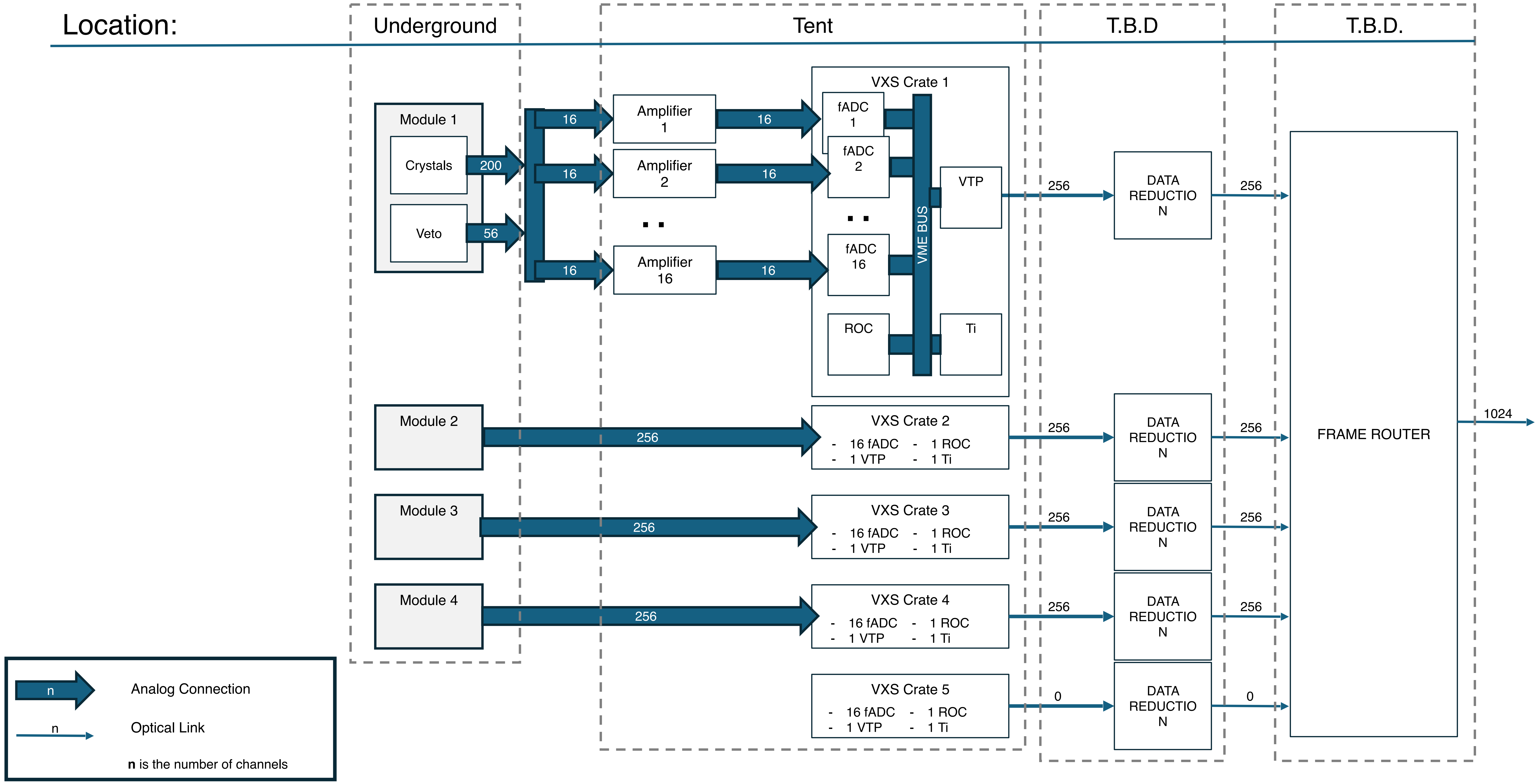
1 Module (64 chs)
R_{Events} ~ 320 kHz
$R_{Data-WF}$ (500ns time window) ~ 130 MB/s

Whole Veto (~256 cos)
R_{Events} ~ 1.3 MHz
$R_{Data-WF}$ (500ns time window) ~ 0.5 GB/s

BDX DAQ

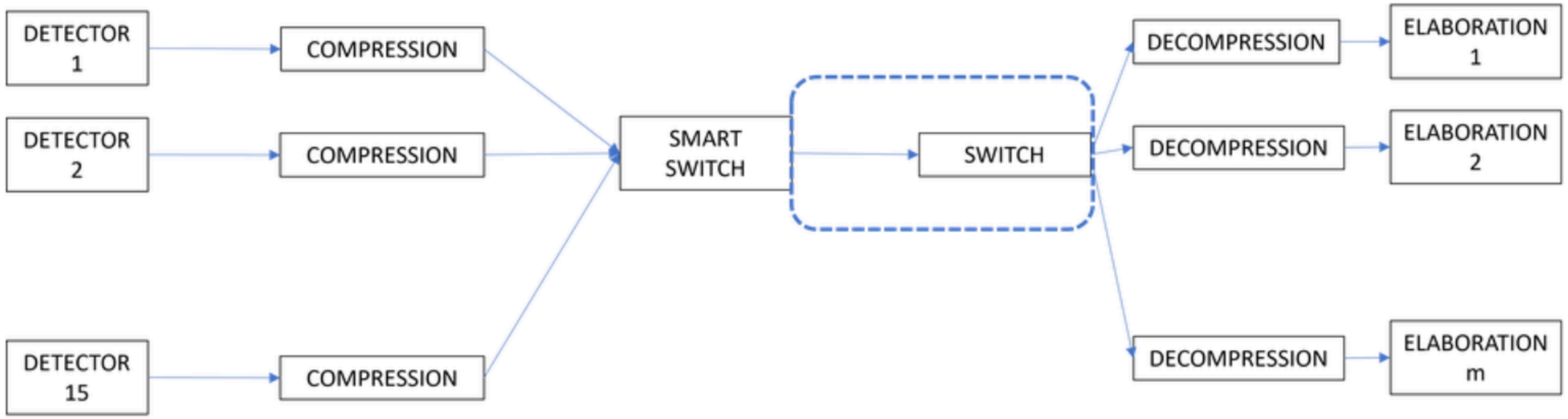
Detector Module I
3M-3M
MCX-MCX
MCX-3M
256 chs preamps/ bias
Lemo-Lemo
256 chs (fADC250) 1 VME-VXS crate
• 200 chs calorimeter (CsI(Tl) Thr ~5 MeV
• 56 chs Vetos (plastic scent) The ~1-2 pe
opt fiber
~1 Gb/s
+0.1 Gb/s
data reduction
DAQ
~4 Gb/s
+0.4 Gb/s
(no data red)
Frame routing
• Minimum bias trigger?
• local coincidence?
• Node I
• Node II
• Node III
Data Rec
• Node N
Detector Module IV
3M-3M
256 chs preamps/ bias
Lemo-Lemo
256 chs (fADC250) 1 VME-VXS crate
• 200 chs calorimeter (CsI(Tl) Thr ~5 MeV
• 56 chs Vetos (plastic scent) The ~1-2 pe
opt fiber
~1 Gb/s
+0.1 Gb/s
data reduction
PIT
Tent
CC

BDX Test Setup



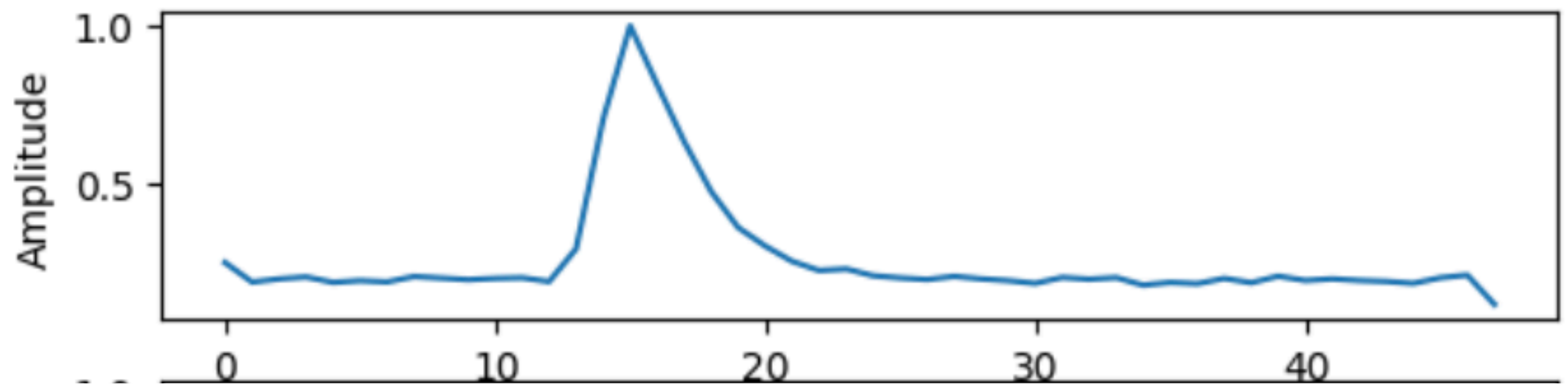
F.Rossi

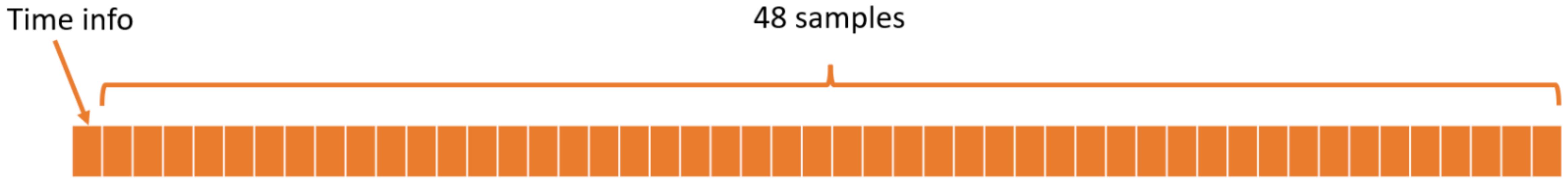
Autoencoder for data reduction

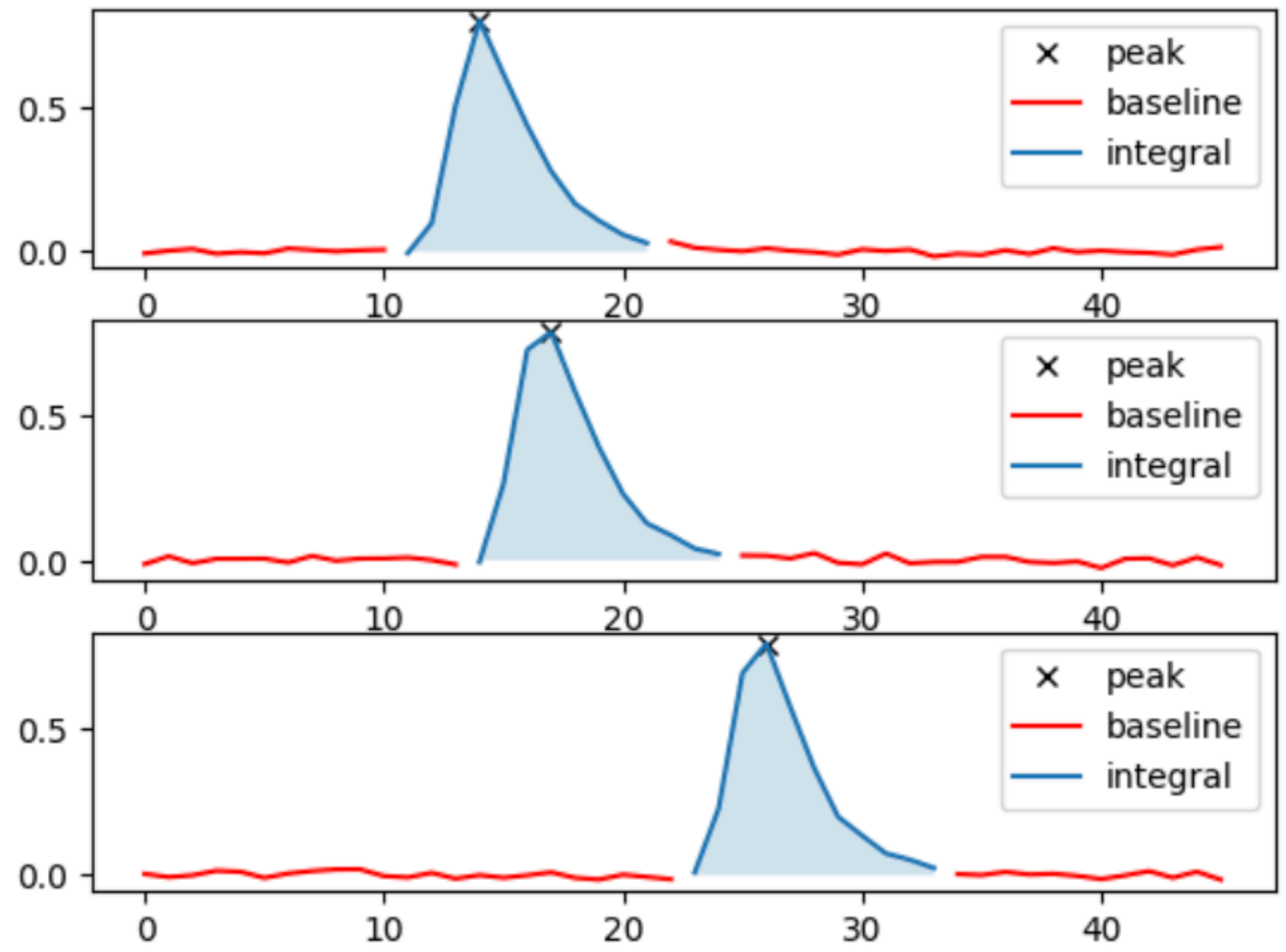


Goal: compress the waveform preserving the information (and eventually retrieve it)

Digitized waveform (48 samples)








```
Integral:[3.12319, 0] Baseline:0.19383583333333
Integral:[3.256660000000001, 0] Baseline:0.2101
Integral:[3.1619400000000004, 0] Baseline:0.200
array([[3.12319, 0.     ],
       [3.25666, 0.     ],
       [3.16194, 0.     ]])
```

Integral and peak position is a possible data reduction (lossy) algorithm

Autoencoder for data reduction

ML NN: FF Autoencoder with dim of latent feature space < dim input layer
(48,96,48,12 - 12, 48, 96, 48)
dim [48] → [12]

- REQUIREMENTS: the NN should be implemented on an FPGA (only integer numbers): as small as possible (regularisation) and weights need to be INT (quantized)
- DATA SET: 25k waveform digitized by a fADC250 at JLab
- Training/Validation/Validation/Test: 48%/15%/12%/25%
- LOSS function: MSE
- ADAM optimizer ($\eta = 10^{-3}$)
- EPOCHS: ~100
- ACTIVATION: ReLU
- WEIGHTS SETTING: training sample, test (during training on validation sample)
- PERFORMANCE: Quality based on comparison of WF integral TRUE/MODEL
- After training, the final weights are transferred to the FPGA for fast inference

```
Model: "Baseline_Model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_4 (InputLayer)        [(None, 48)]              0

 dense_21 (Dense)            (None, 96)                4704

 dense_22 (Dense)            (None, 48)                4656

 dense_23 (Dense)            (None, 12)                588

 dense_24 (Dense)            (None, 12)                156

 dense_25 (Dense)            (None, 48)                624

 dense_26 (Dense)            (None, 96)                4704

 dense_27 (Dense)            (None, 48)                4656

=================================================================
Total params: 20088 (78.47 KB)
Trainable params: 20088 (78.47 KB)
Non-trainable params: 0 (0.00 Byte)
```

Baseline model
- best results
- Starting point for further optimization

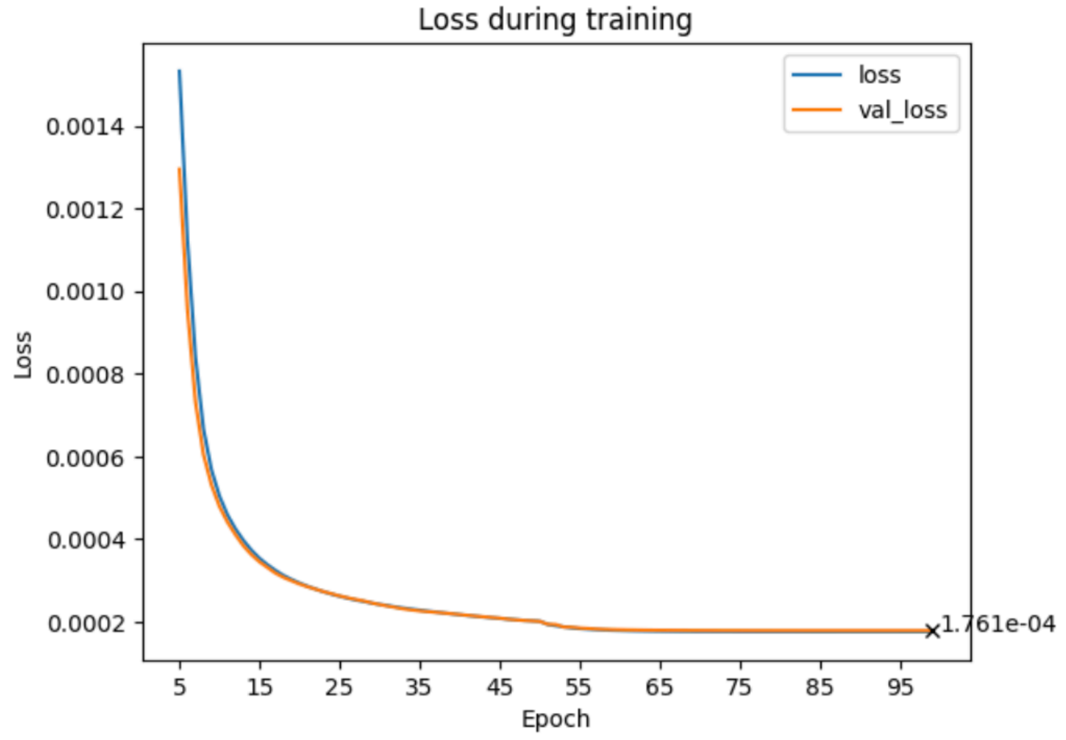
Pruned model
- zero (low values w_i) uppression
- re-trained starting from baseline
- sparsity evaluation

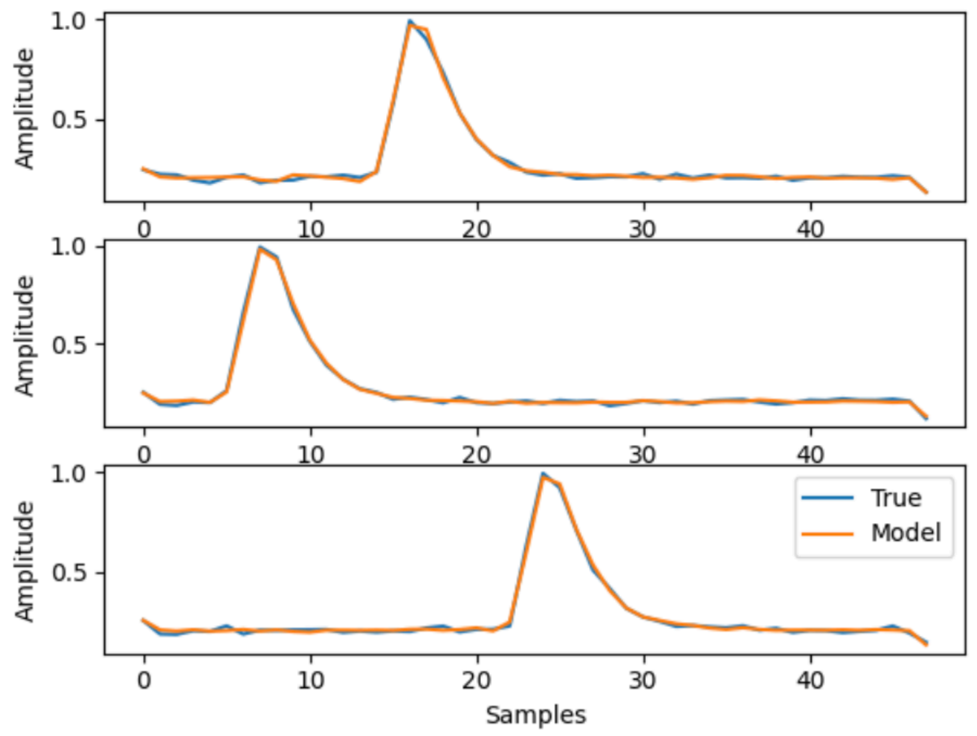
Quantized model
- re-trained starting from pruned (keeping pruned model)

Autoencoder for data reduction

Results

Baseline model





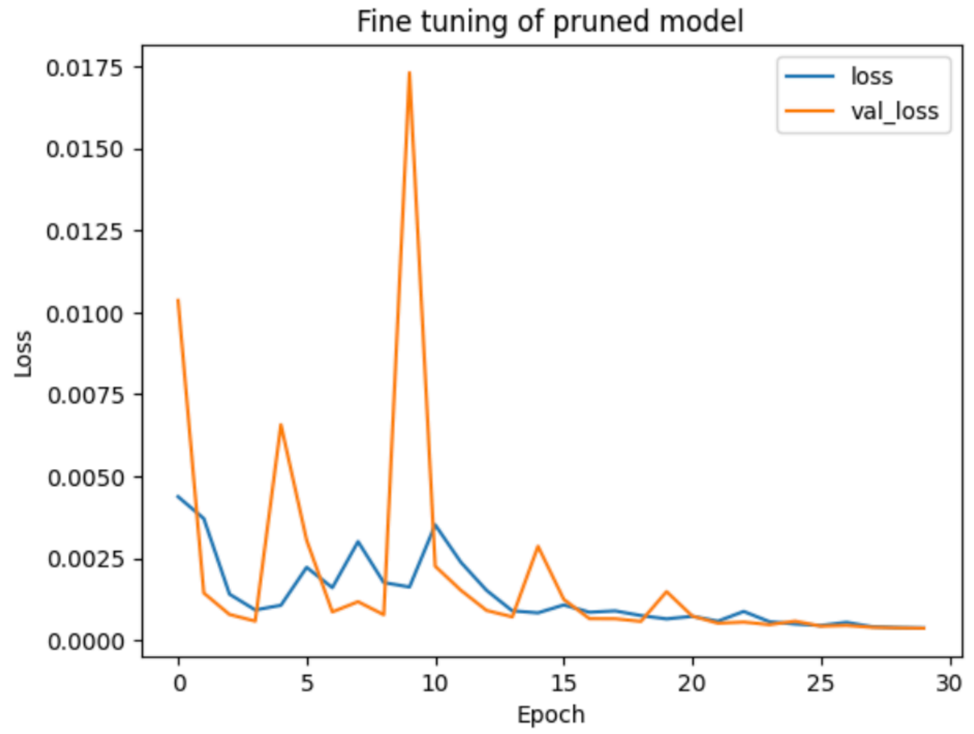
Original	Recon	Error%
3.154	3.151	0.093
3.245	3.224	0.656
3.220	3.225	-0.141

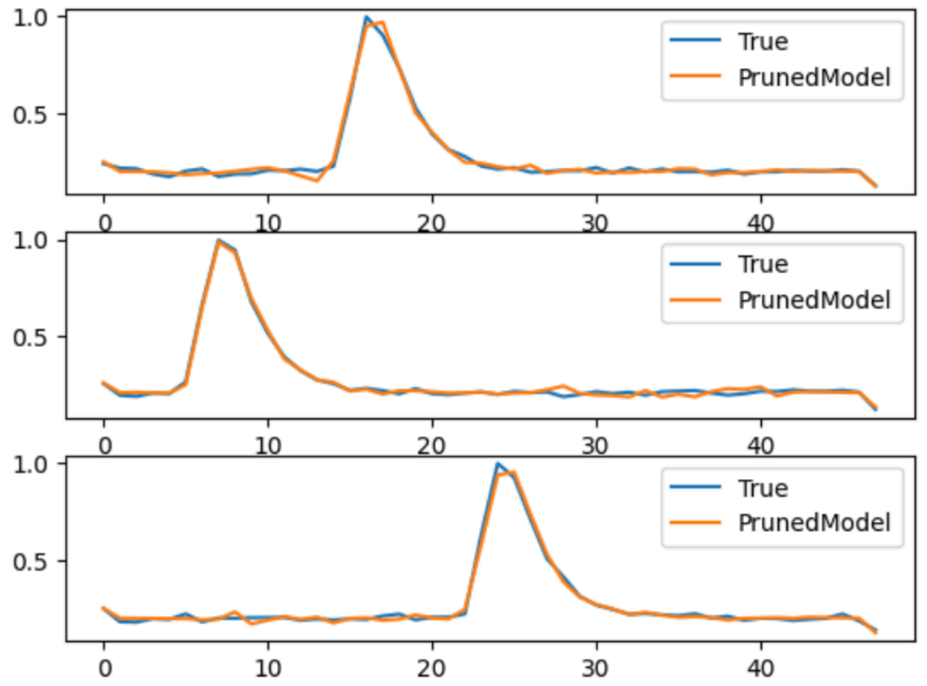
```
dense_21/kernel:0: 79.99% sparsity  (3686/4608)
dense_21/bias:0: 2.08% sparsity  (2/96)
dense_22/kernel:0: 79.99% sparsity  (3686/4608)
dense_22/bias:0: 0.00% sparsity  (0/48)
dense_23/kernel:0: 80.03% sparsity  (461/576)
dense_23/bias:0: 0.00% sparsity  (0/12)
dense_24/kernel:0: 79.86% sparsity  (115/144)
dense_24/bias:0: 8.33% sparsity  (1/12)
dense_25/kernel:0: 80.03% sparsity  (461/576)
dense_25/bias:0: 4.17% sparsity  (2/48)
dense_26/kernel:0: 79.99% sparsity  (3686/4608)
dense_26/bias:0: 0.00% sparsity  (0/96)
dense_27/kernel:0: 79.99% sparsity  (3686/4608)
dense_27/bias:0: 0.00% sparsity  (0/48)
```

```
Original model
[[ 2.56448984e-05  2.04816848e-01 -2.20985472e-01 ...  1.59392953e-01
   1.78266406e-01 -1.47754893e-01]
 [-1.90256640e-01  3.36886868e-02 -1.01705797e-01 ... -1.12358101e-01
  -2.99690175e-03 -1.91902950e-01]
 [-5.07703274e-02  7.48560280e-02  4.45862524e-02 ... -6.69857115e-02
   7.05242679e-02  7.37352520e-02]
 ...
 [ 1.14938974e-01 -1.74160168e-01 -1.65666074e-01 ... -1.08603626e-01
   1.09323412e-01  1.56354651e-01]
 [ 1.82403296e-01 -1.98742032e-01  1.27638802e-01 ... -9.61249769e-02
   1.73485637e-01  1.09817624e-01]
 [-1.05654188e-01 -1.86228636e-03 -1.85805395e-01 ...  1.30987376e-01
   1.46889716e-01  6.92100003e-02]]

Pruned model
[[-0.          0.22588937 -0.22141045 ...  0.          0.17150576
   0.        ]
 [-0.19826248  0.          0.         ...  0.          0.
  -0.1860043 ]
 [-0.          0.          0.         ...  0.          0.
   0.        ]
 ...
 [-0.          0.          0.         ...  0.          0.
   0.        ]
 [ 0.17366754 -0.1748717   0.         ...  0.          0.
   0.        ]
 [-0.          0.         -0.18634865 ...  0.          0.
   0.        ]]
```

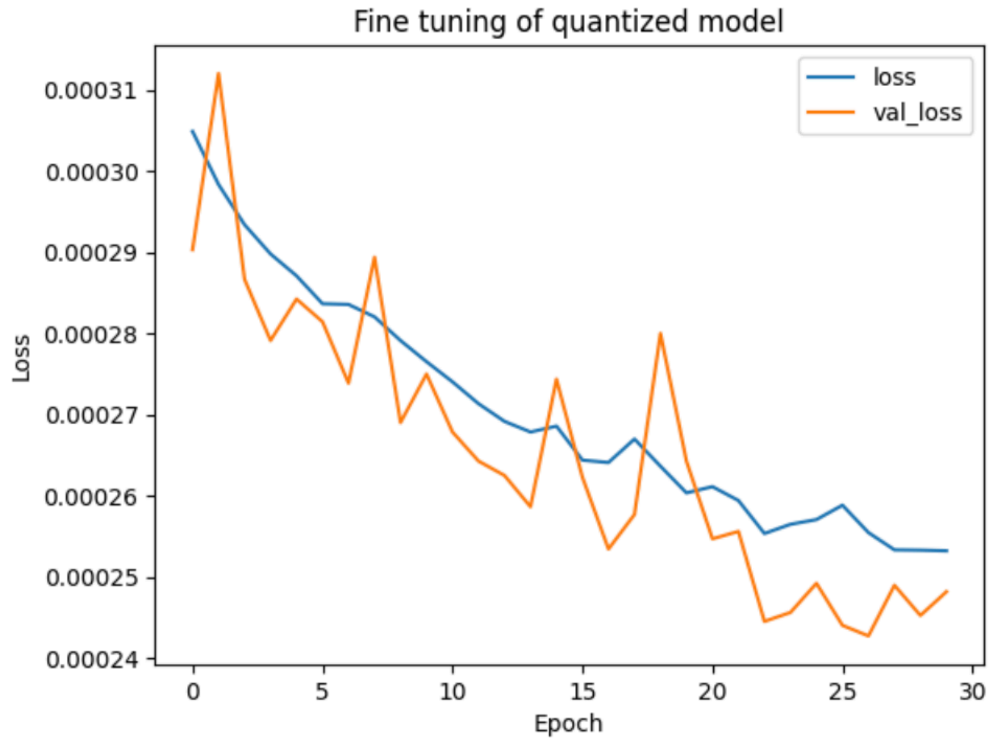
Pruned model

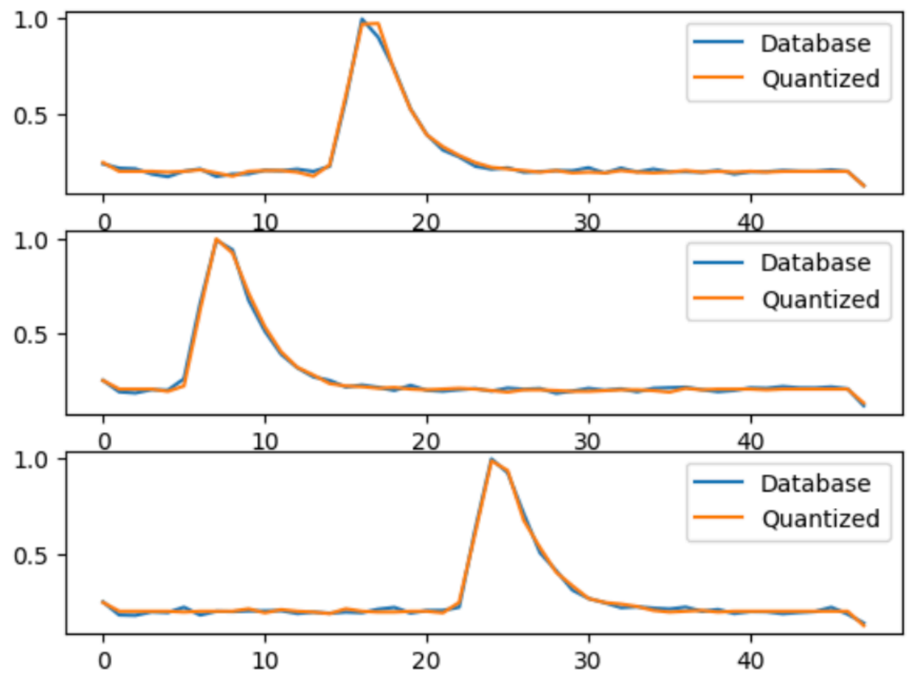




Original	p_Recon	Error%
3.154	3.247	-2.934
3.245	3.207	1.183
3.220	3.229	-0.278

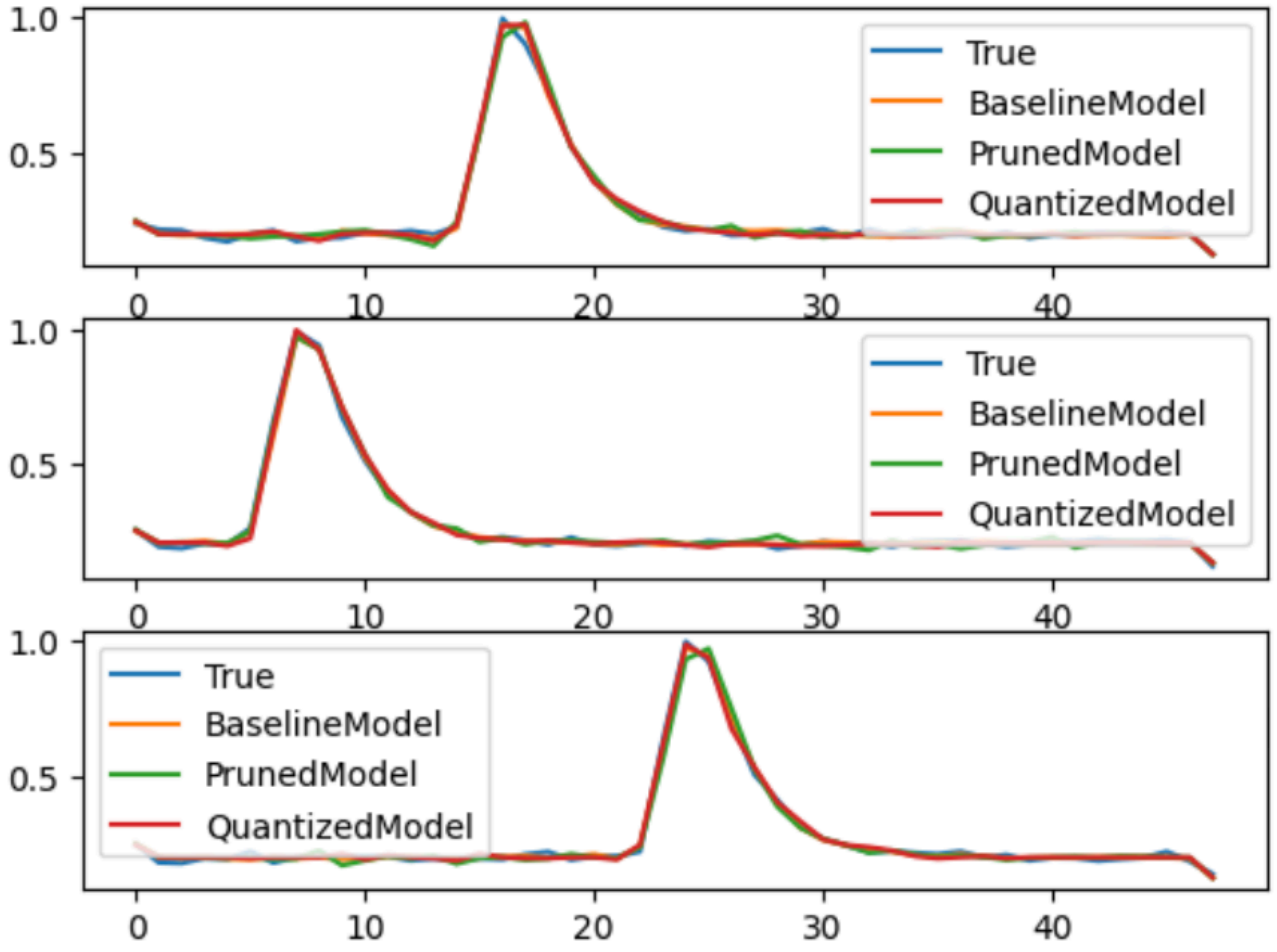
Quantized model





Original	q_Recon	Error%
3.154	3.310	-4.936
3.245	3.248	-0.094
3.220	3.237	-0.538

Autoencoder for data reduction



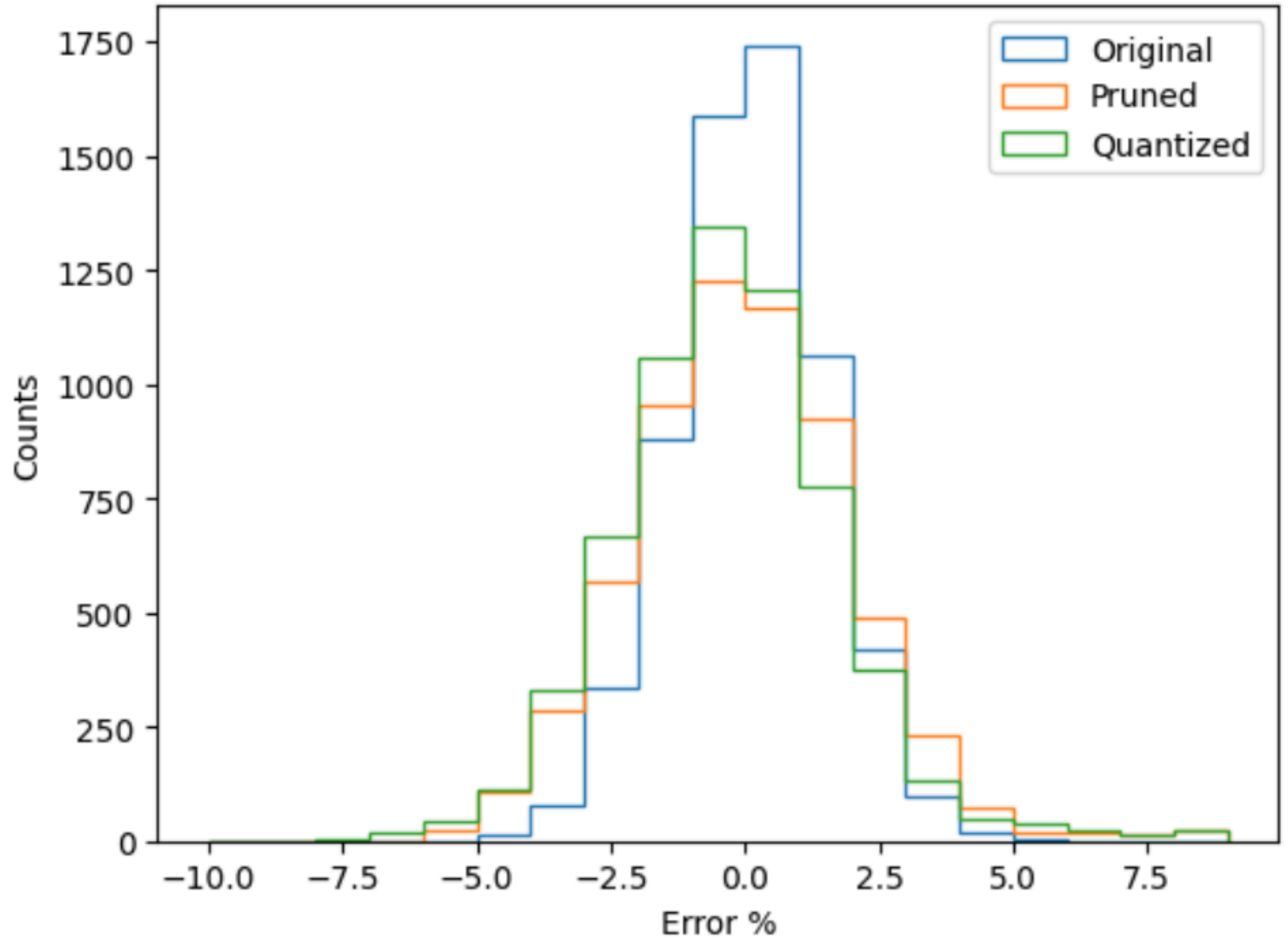

```
Original Recon Error%
3.154    3.151  0.093
3.245    3.224  0.656
3.220    3.225  -0.141
Original p_Recon Error%
3.154    3.247  -2.934
3.245    3.207  1.183
3.220    3.229  -0.278
Original q_Recon Error%
3.154    3.310  -4.936
3.245    3.248  -0.094
3.220    3.237  -0.538
```

Results

```
Comparison between models
################################
Baseline model:
  Mean: 0.12 Std: 1.41
  Elapsed time: 284 us
  Float model in kb: 82.0

Pruned model:
  Mean: 0.23 Std: 2.77
  Elapsed time: 279 us
  Pruned model in kb: 82.0

Pruned+Quantized model:
  Mean: -0.24 Std: 2.24
  Elapsed time: 396 us
  Quantized model in kb: 26.2
```

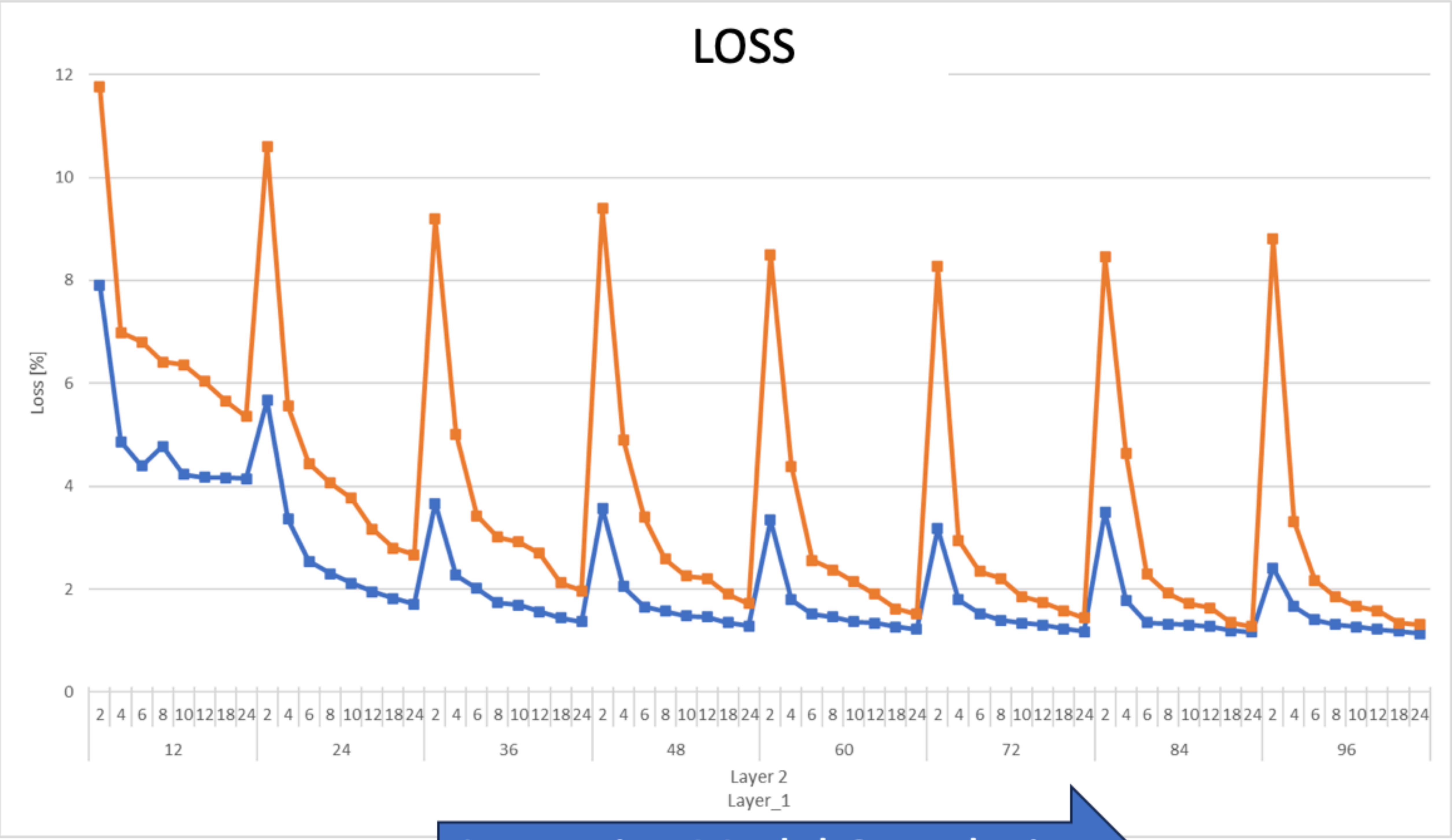


Compression 1:4

Autoencoder for data reduction

- Performance considerations

Signal Compression: Execution time vs Loss



Increasing Model Complexity

Model: RASPBERRY PI4
Quad core Cortex-A72 (ARM v8)
C code without any optimizations