

October 21st, 2026

Lossless TPC Hardware Compression at ICARUS

HARRY HAUSNER

CPAD 2026 — Seattle

FERMILAB-SLIDES-26-0235-PPD



U.S. DEPARTMENT
of **ENERGY**

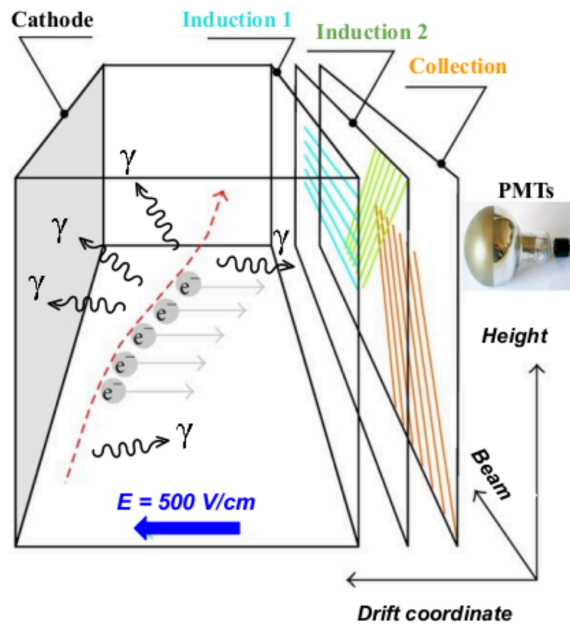
Fermi National Accelerator Laboratory is managed by
FermiForward for the U.S. Department of Energy Office of Science



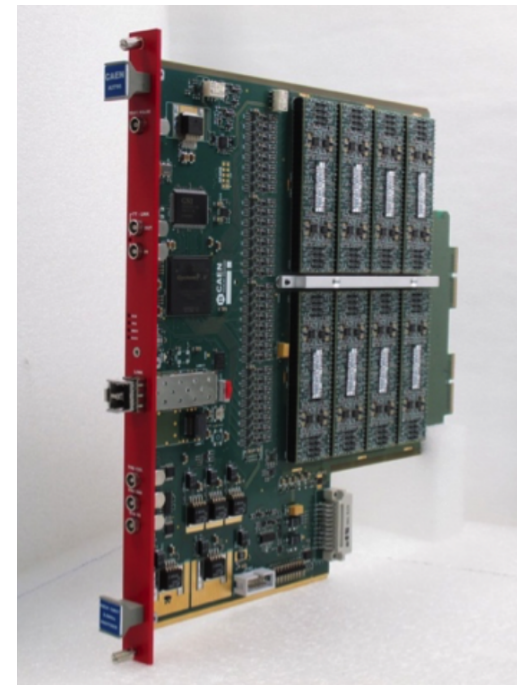
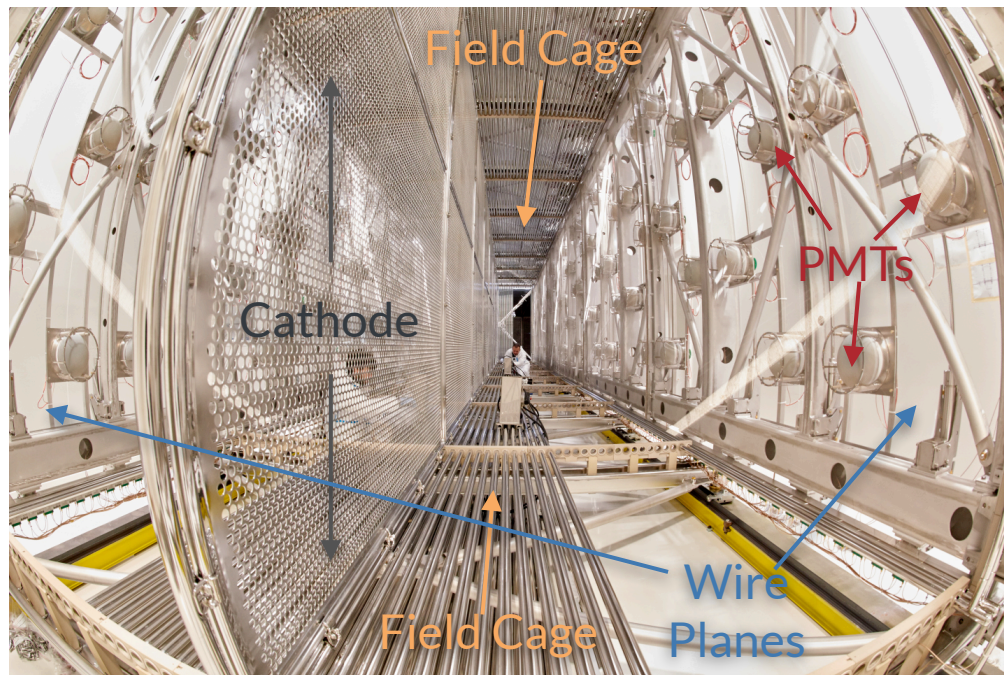
ICARUS

Far detector of the Short-Baseline Neutrino program at Fermilab

- 760-ton liquid argon detector, two cryostats for four wire TPCs
 - Refurbished at CERN after LNGS running
 - Moved to FNAL for short baseline neutrino oscillation physics
- 600 m from the BNB target, 800 m from the NuMI target
 - At the surface, so cosmic rays are abundant
- 53,248 TPC wires sampled at a 2.5 MHz rate
- 832 CAEN A2795 boards with analog and digital processing
- DAQ with [artdaq](#)



doi:10.1140/epjc/s10052-023-11610-y



[L. Bagby et al 2018 JINST 13 P12007](#)





Every trigger reads out the full detector

The detector is BIG and charge readout SLOW

218,103,808 ADC samples

4096 ADC samples per wire per trigger
for the full ~1 ms drift

Event number			
Trigger time stamp			
	Ch 1, S 0		Ch 0, S 0
	Ch 3, S 0		Ch 2, S 0
⋮			
	Ch 1, S 1		Ch 0, S 1
	Ch 3, S 1		Ch 2, S 1
⋮			
	Ch 1, S 4095		Ch 0, S 4095
	Ch 3, S 4095		Ch 2, S 4095
⋮			
Trailer			

64 ch × 4096 samples × 16 bit

436 MB

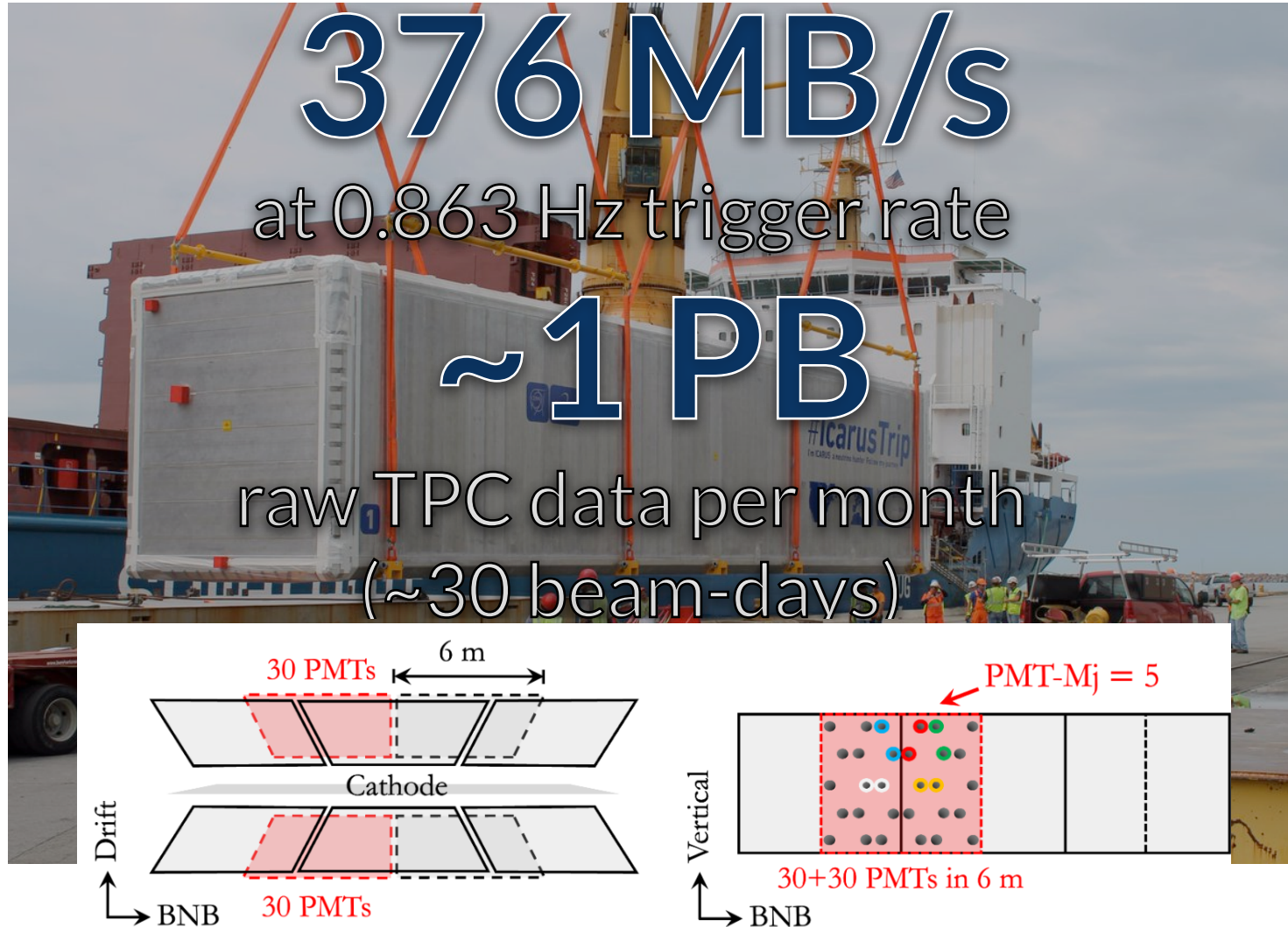
12-bit ADC + headers per trigger





Every trigger reads out the full detector

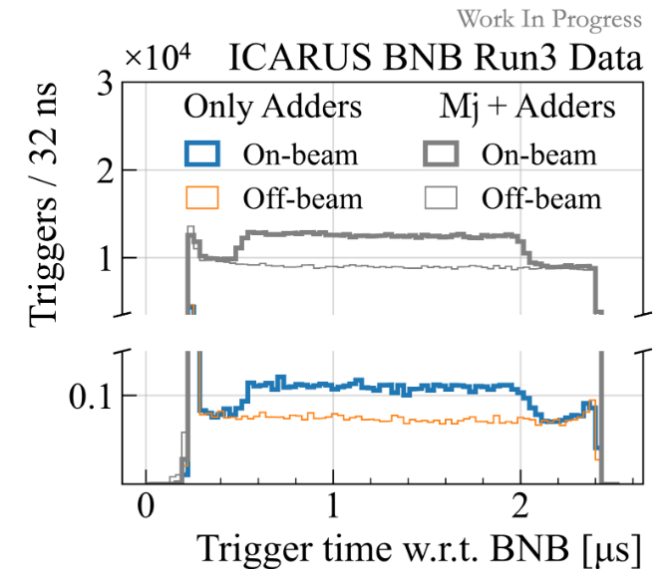
The detector is BIG and charge readout SLOW



F. Abd Alrahman et al 2025 JINST 20 P10044

The ICARUS trigger logic reads out when a configurable number of PMTs in a 6 m TPC slice have an LVDS signal over 13 PE ($M_j = N$)

An "Adder" signal also triggers when the total signal from PMTs in a 3 m TPC slice go above a 60 mV thresholds

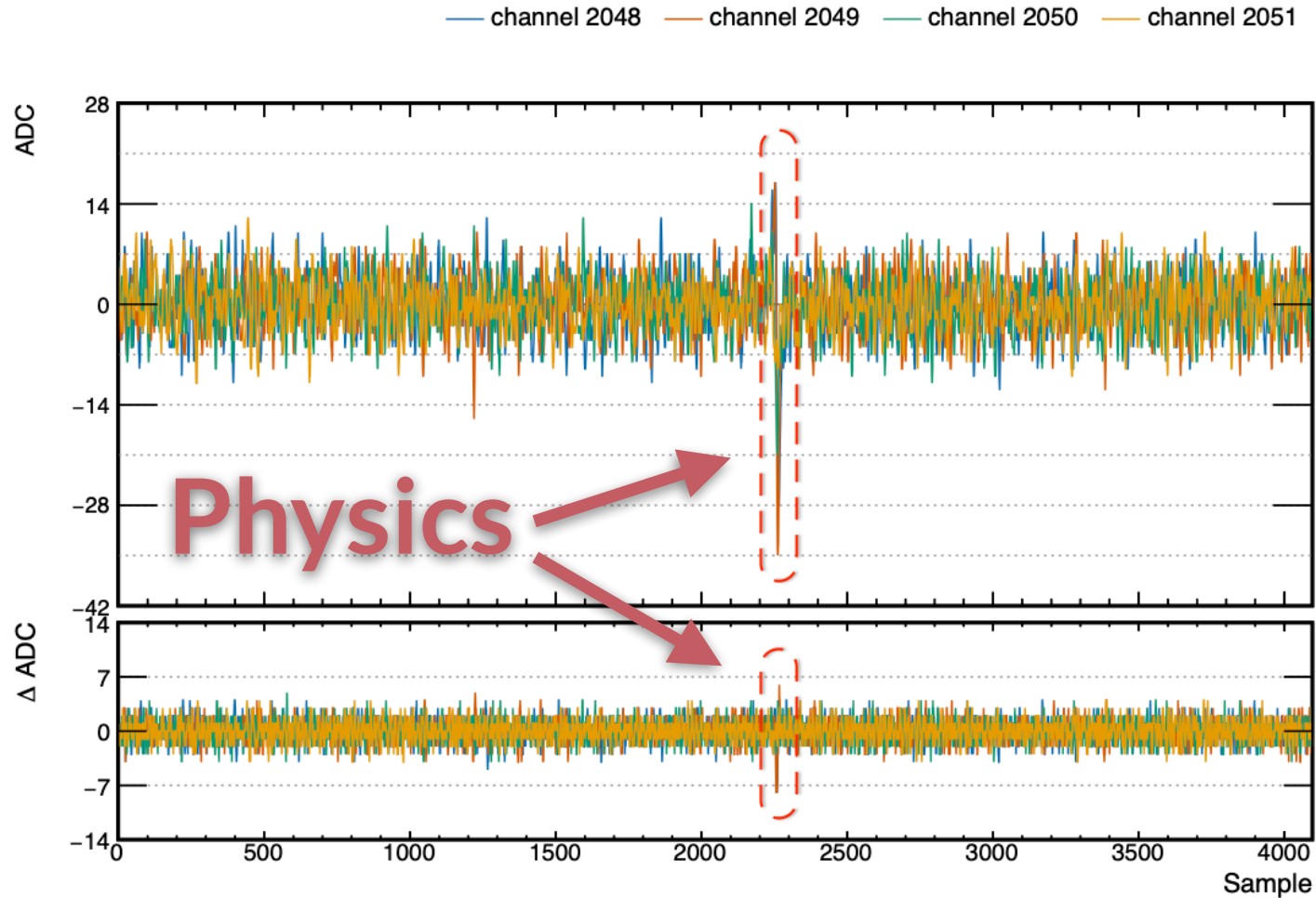


Collecting on-beam and off-beam data we get a 0.863 Hz global trigger rate



What Does That TPC Data looks Like?

Four adjacent channels over a full 1.64 ms readout window



Flat baseline, sparse signal

Raw ADC near zero apart from ionization signal

Tick-to-tick difference are small, and on the order of a few ADC

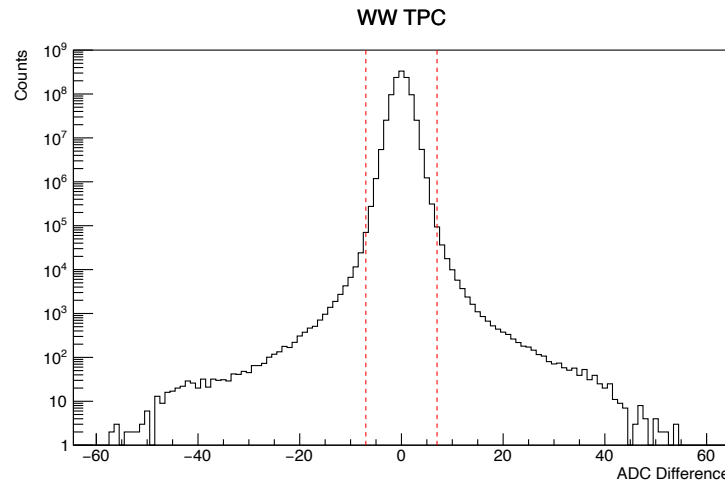
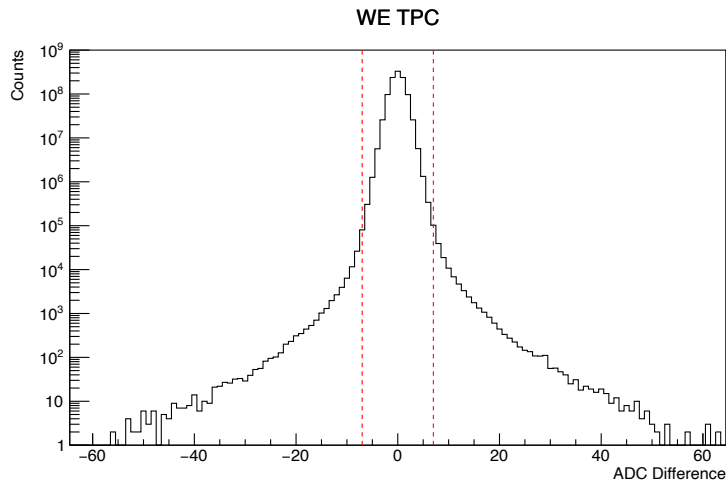
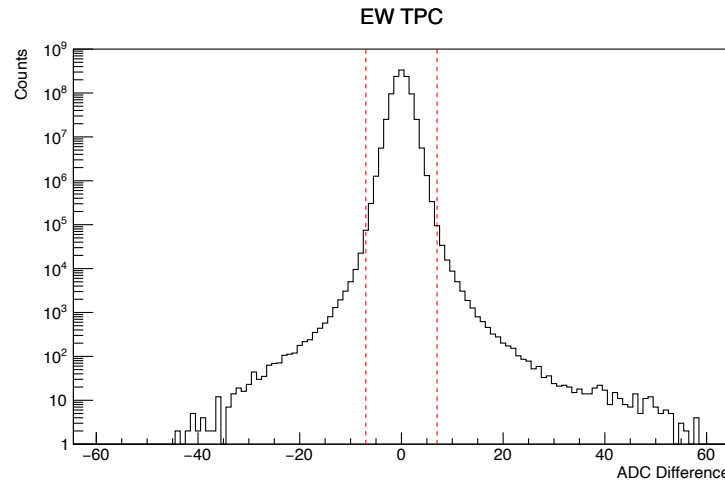
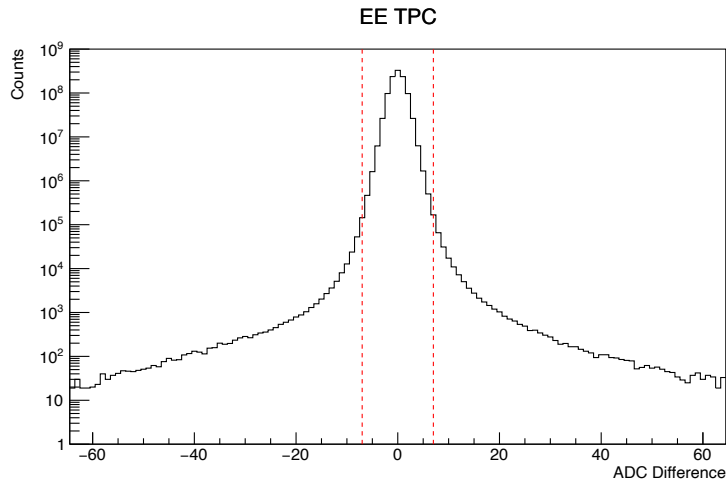
We don't need to 12-bits to store these numbers

→ store the difference, not the value



ADC differences

Sample-to-sample ADC difference are highly peaked near 0



~93%

of differences within ± 7 counts
in every TPC for cosmic data

→ a 4-bit signed delta
covers almost every
sample

Only ~7% fall in the tails



Encoding

Lossless per-sample ADC encoding



Raw

12-bit ADC in a 16-bit word: the top 4 bits are spacers

Compressed

Four channels' differences in one word: 4 bits per sample instead of 16

Overflow

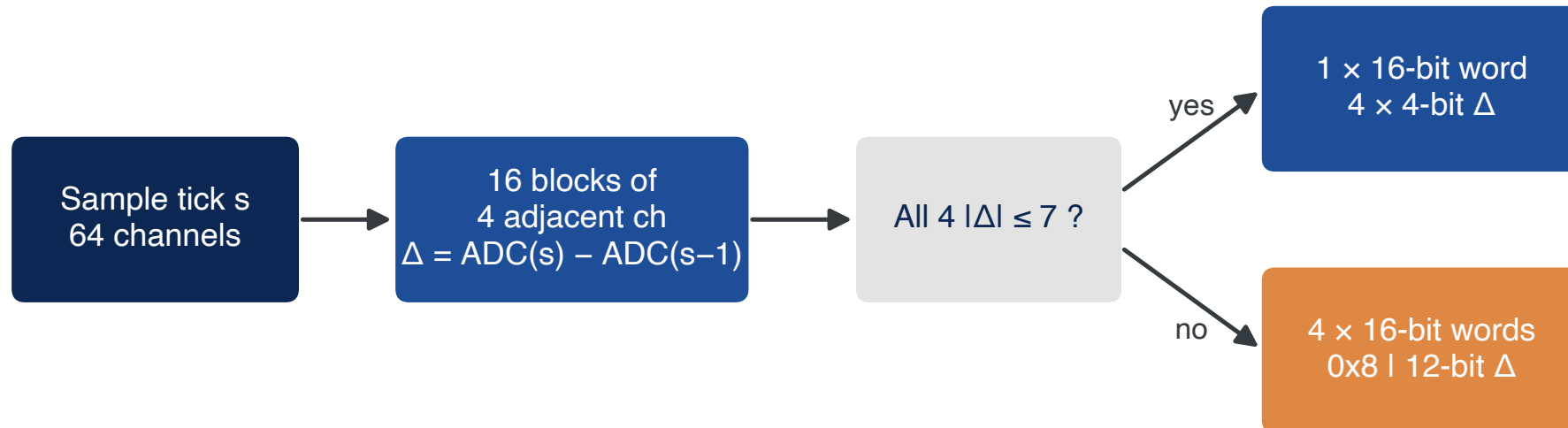
0x8 marker + 12-bit value: the same size as a traditional word



Firmware

Compression on the hardware

- Compression enabled/disabled in control register 0x1004
- Overflows are checked in blocks of 4 adjacent channels
 - One overflow in a block means the whole block is uncompressed to preserve the fragment's 32-bit word structure
- Sample 0 is stored as absolute ADC value for baseline

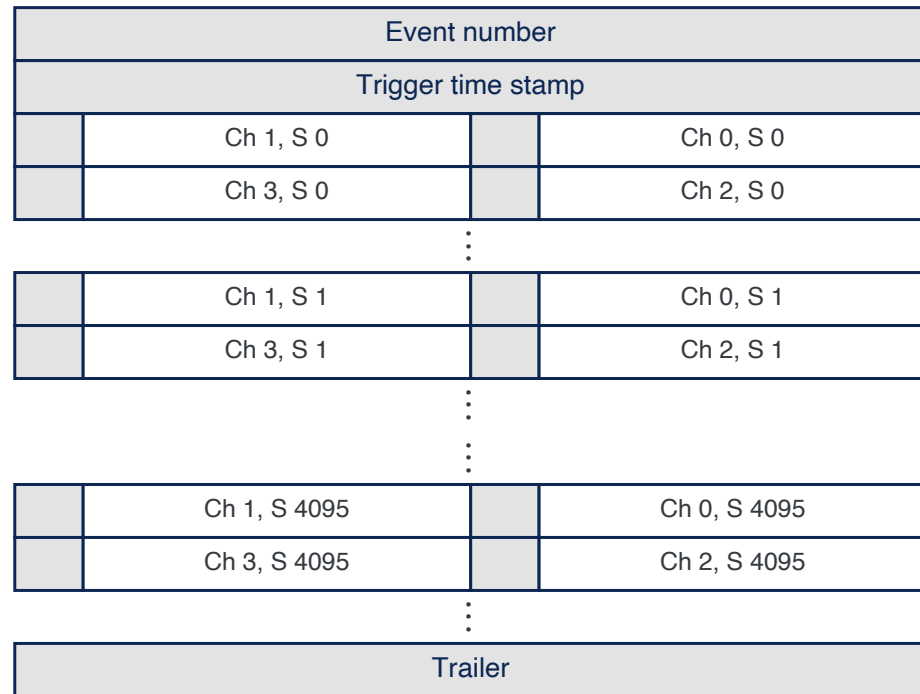




How the data is stored

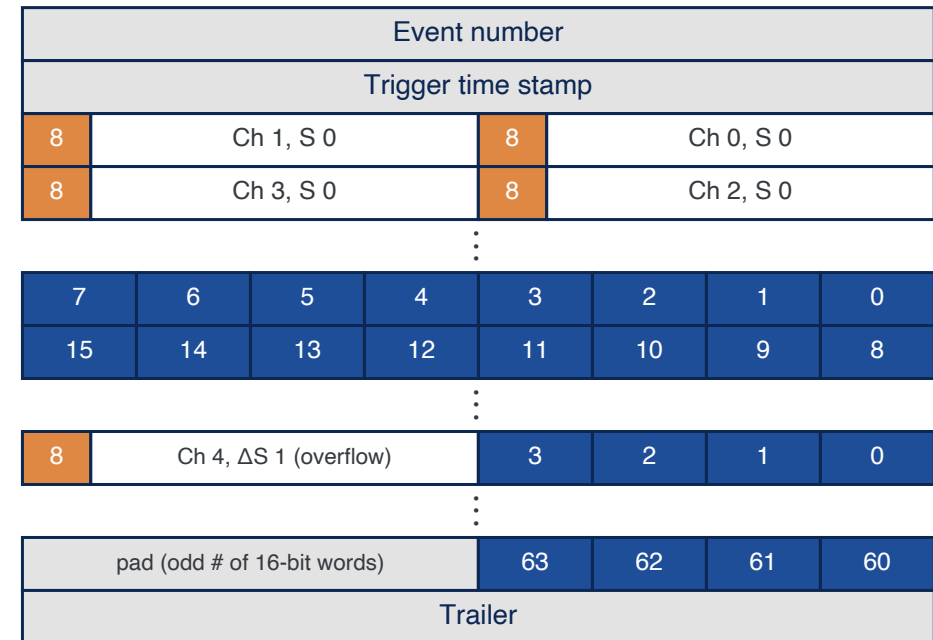
64 channels $\times$ 4096 samples per board, read out as 32-bit words in *artdaq*

Uncompressed: 2 channels / 32-bit word



64 ch $\times$ 4096 samples $\times$ 16 bit

Compressed: up to 8 channels / 32-bit word



blue cell n = 4-bit Δ of channel n



Bandwidth

How much does this save us

- Uncompressed: $64 \text{ ch} \times 4096 \text{ samples} \times 16 \text{ bits} \approx 524 \text{ kB}$ per board (with header and trailer); 832 boards at 0.863 Hz $\rightarrow$ 376 MB/s
- Compressed: $\sim 7\%$ of sample differences fall outside ± 7 ; conservatively assume 7% of every 4-channel block is stored uncompressed $\rightarrow$ 3.31 $\times$ smaller fragments
- Real fragments vary, so 114 MB/s is an upper bound on the throughput

3.3 $\times$

smaller TPC fragments
(conservative estimate)

132 MB

TPC data per event
down from 436 MB

114 MB/s

aggregate readout
down from 376 MB/s

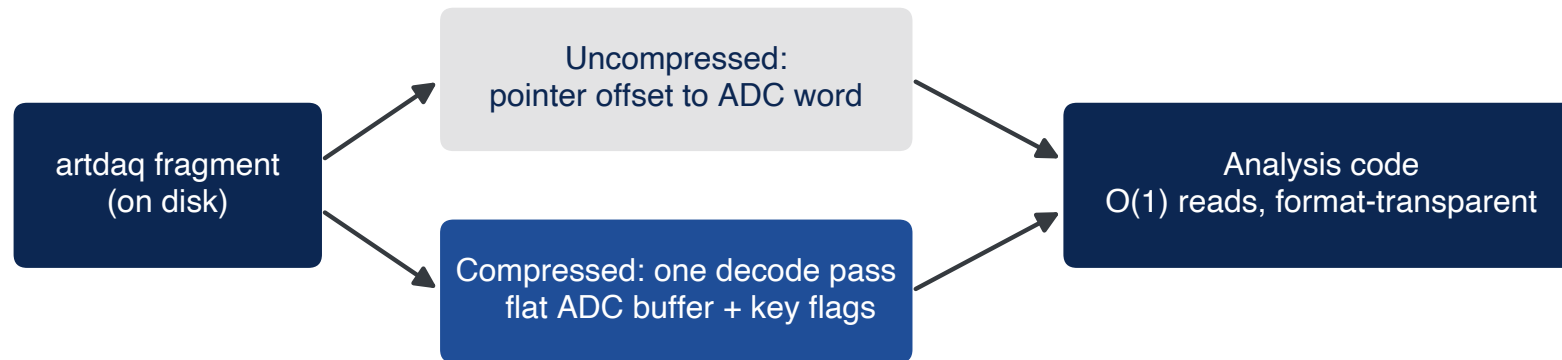
Relaxes event-builder budget $\rightarrow$ room for higher trigger rates without data loss



Software

How to handle indeterminate data blocks

- Payload length depends on the data, so there is no fixed offset to a given ADC value
- Walk each board 16 bits at a time: high nibble 0x8 → one channel's 12-bit value or Δ ; otherwise → four 4-bit Δ s for the next block



Compressed fragments are decoded once, when the overlay is built (~17 ms per fragment)

After that, sequential and random reads are $O(1)$ from a flat, contiguous buffer

Downstream code never needs to know whether the data were compressed



Disk Space

Hardware and ROOT compression are complementary

- Data stored in ROOT files on disk, which has it's own software compression scheme
- To validate the compression files with 10 neutrino events were evaluated with/without the hardware compression and the ROOT compression
- The two are complementary and total space savings are a combined factor of 5.4

10 neutrino events	File size (GB)	Storage / day (TB)	Storage / month (TB)
Uncompressed	4.876	36.4	1091
ROOT only	1.634	12.2	366
Hardware only	1.668	12.4	373
Both	0.896	6.68	200

Same 10 events with hardware and ROOT compression on/off; storage assumes 0.863 Hz and ~30 beam-days/month. Both together save ~0.89 PB per month.



Why it matters

Today at ICARUS, and for future LArTPCs

- Cosmic overlays
 - Real cosmic TPC and PMT waveforms on top of simulation to reduce uncertainties
 - Requires large off-beam samples at a high trigger rate and which take up disk
- First hardware compression of TPC readout bandwidth in a surface accelerator-neutrino experiment
 - Compression effectiveness survives higher occupancy events with cosmogenic activity
- DUNE near detector: ~55 interactions per spill, so bandwidth must be controlled and compression must be lossless
 - DUNE ND hardware is different, but methods can be adapted



Lossless firmware compression: 3× less readout bandwidth, 5.4× less disk

4-bit deltas per 4-channel block in the CAEN A2795 · bit-exact ADC values · 376 MB/s → 137 MB/s at nominal beam trigger · ~0.89 PB saved per month · no offline access penalty



Fermi**FORWARD**



U.S. DEPARTMENT
of ENERGY



B

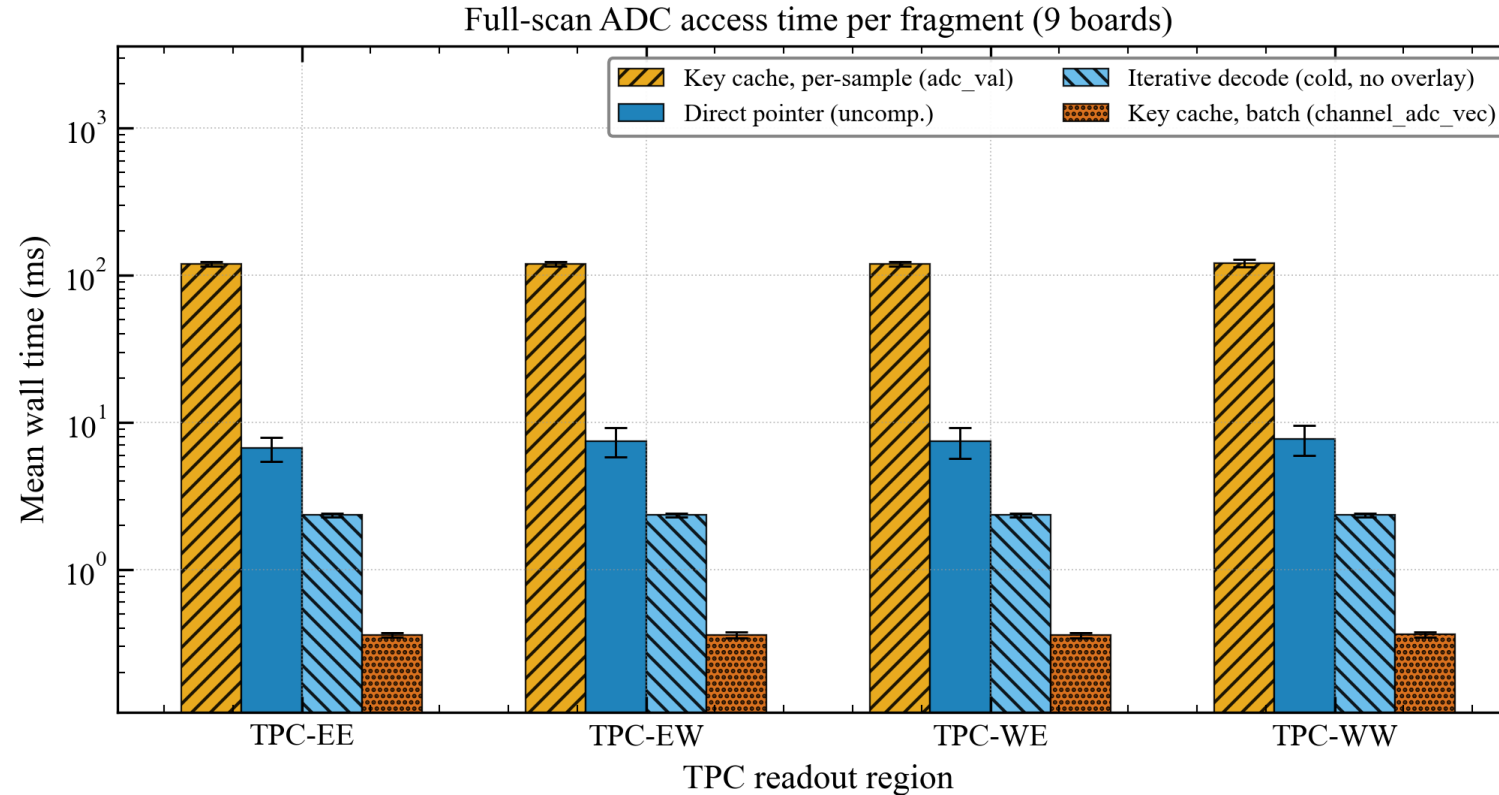
Backup



Offline data reads

Full-scan ADC access time per fragment (9 boards, log scale)

- Key cache, per-sample: the random-access call used in a sequential loop (worst case)
- Direct pointer: uncompressed baseline, fixed offset into the raw fragment
- Iterative decode: walks the compressed bitstream from scratch, without the overlay
- Key cache, batch: production path, reads whole waveforms from the decoded buffer



Sequential scan

0.36 ms vs. 7.34 ms
uncompressed: ~20× faster from
the decoded buffer

Random access

10⁵ queries: 6.2 ms vs. 1.6 ms,
~4× slower but negligible next to
reconstruction

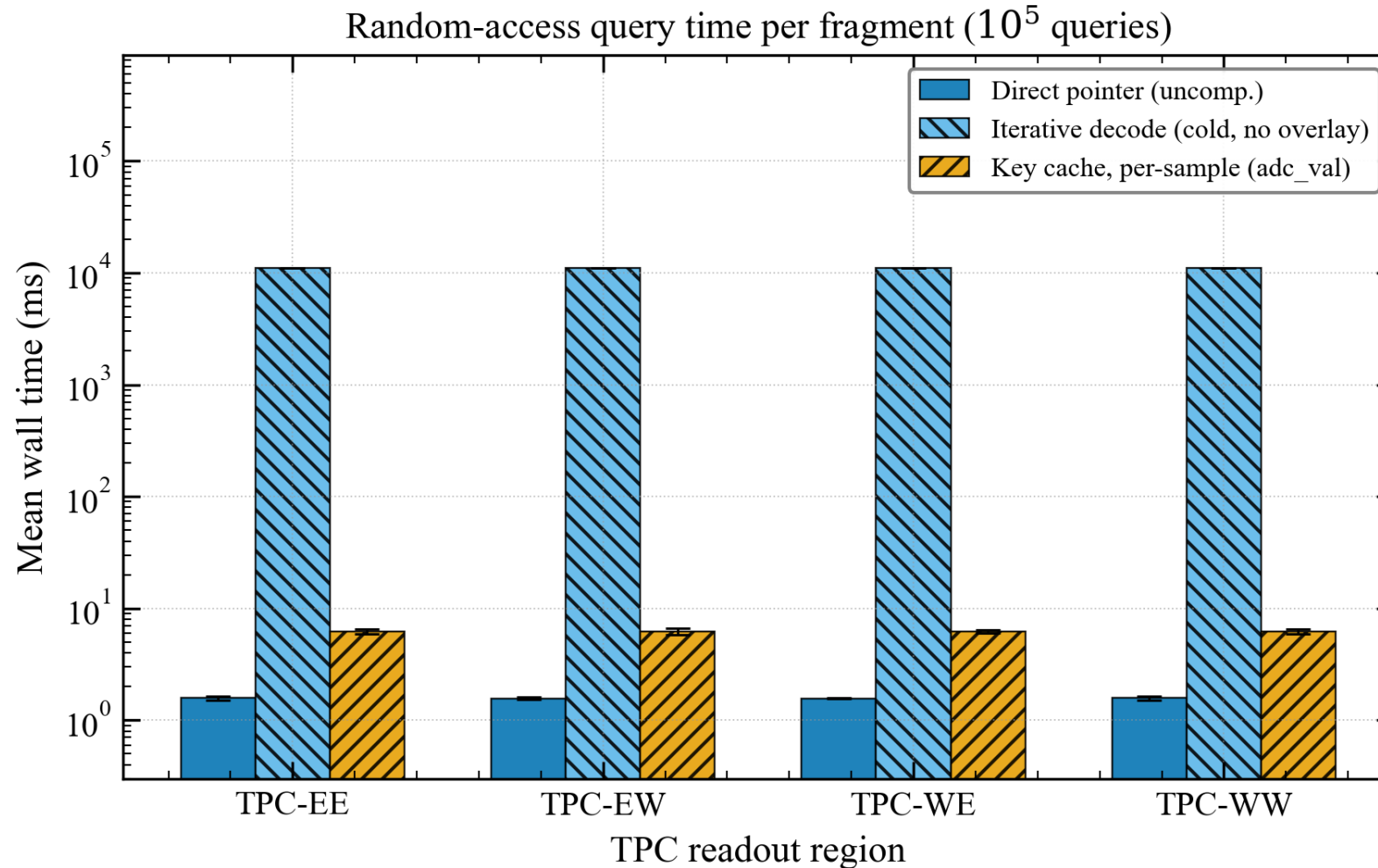
Decode pass

~17 ms per fragment, paid once



Random-access performance

10^5 uniformly distributed (board, channel, sample) queries per fragment



Uncompressed pointer: ~1.6 ms

Compressed key cache: ~6.2 ms

Cold bitstream walk per query: ~11 s

→ the one-time decode pass is essential



A2795 compression registers

Compression is toggled run by run with a single bit

Address	Name	Function
0x0000	Output Buffer	BLT read window for event data
0x1004	Acquisition Control	Bit 5 enables compression mode
0x1008	Acq. Control (bit set)	Set bits without rewriting the register
0x100C	Acq. Control (bit clear)	Clear bits without rewriting the register

Full register map: CAENA2795 user manual (UM2580). Mode is latched at run configuration and applies to every board.