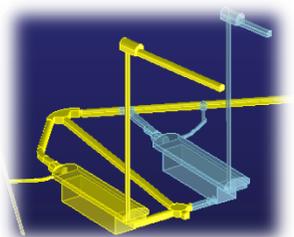


DAQ Design Concept

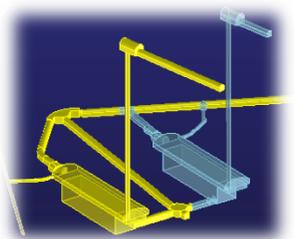
DAQ Breakout Session
1 November 2010

The Long-Baseline Neutrino Experiment



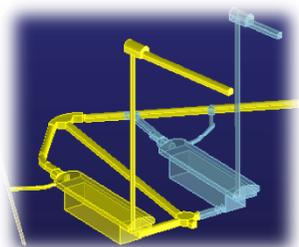
Outline

- General Considerations
- Rates summary
- Data Flow Diagram & Component Counts
- Technical Info on Components/Subsystems
 - Data Concentrator Module
 - Timing System
 - Ethernet Switch Array
 - Data Farm: Triggering & Event building
 - Network Management: Supervisor & Routing Master
 - Power Requirements
- Examples of Costs
- Risks
- Work remaining for CD-1

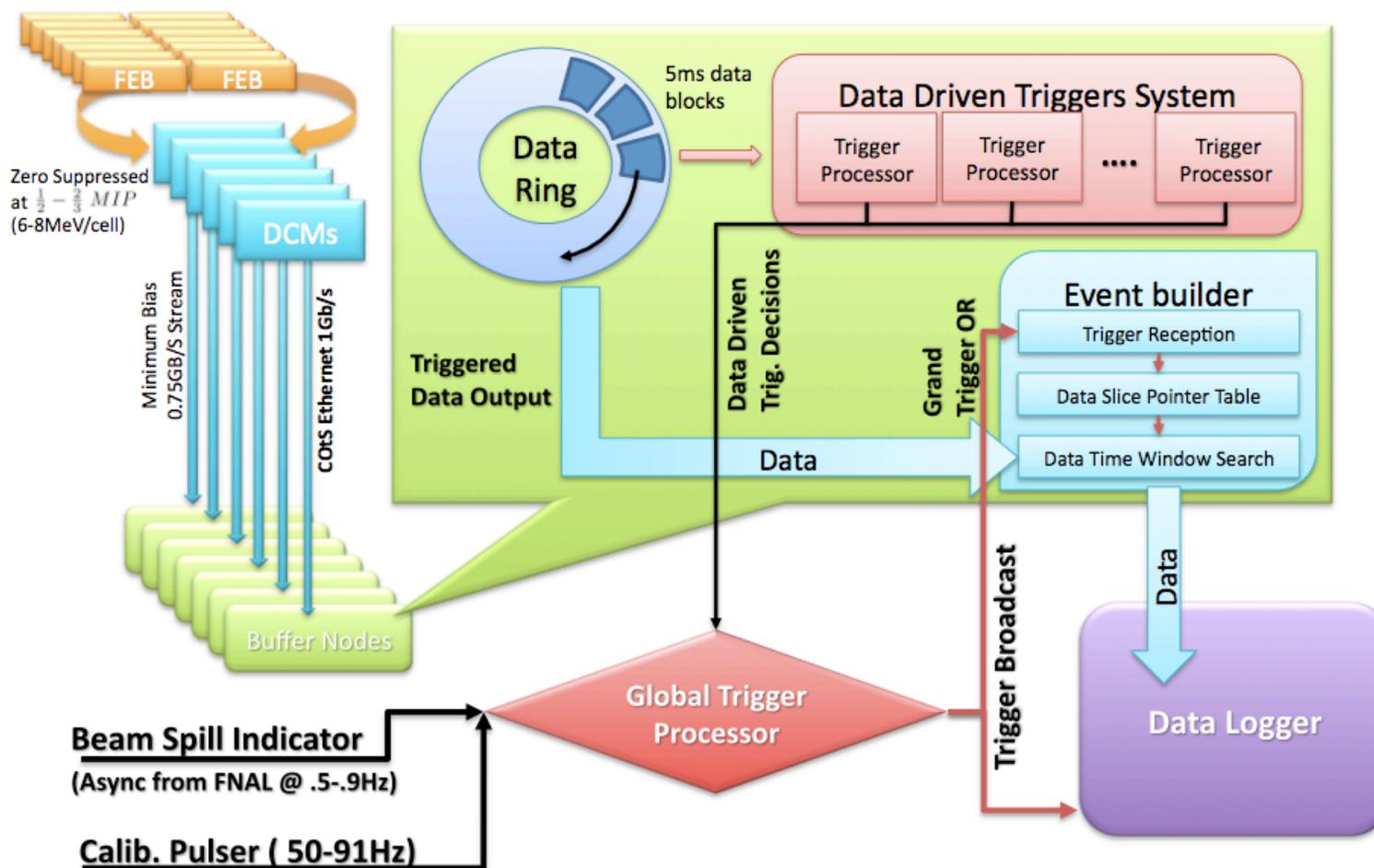


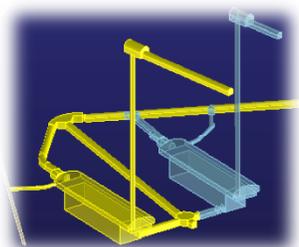
General Considerations

- With sparsification in the front-ends, expect modest data rates through the DAQ despite high channel count (645k) and sampling rate (2MHz).
- Suggests to exploit commodity commercial components & experience from other expt's as much as possible. Use of ethernet switch arrays (as in D-Zero, CMS) is an obvious application.
- NOvA provides a good model to consider for much of the system.
 - Comparable channel count (360k for 14kt NOvA, vs 645k for LAr20)
 - Similar beam signal timing properties & handling requirements
 - 2 MHz sampling at front-end, same as for LAr20, but event duration is short (scintillation signal) – few μs vs. 1.6 ms.
 - On the other hand, per channel activity from cosmics & noise is high (surface detector) – ~ 100 Hz (mostly cosmics) vs. ~ 10 Hz (^{39}Ar)



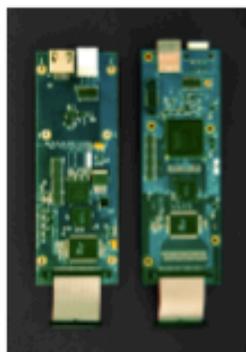
NOvA Data Flow & Triggering





NOvA Data Flow & Triggering

Front End Boards



Nanoslice stream

(single sample 16MHz clock)

Data Concentrators



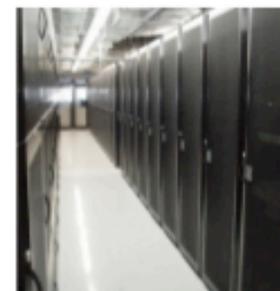
microslice blocks = 5 μ s window

Stage 1 Event Builder (DCM)

millislice blocks = 5ms window

180x DCM

Buffer Farm (Data Center)



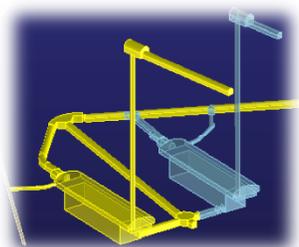
NOvA Data Buffer = 150ms

Total Buffer time = 150ms x N nodes
= 20-30s baseline

July 2010

A.Norman, FNAL-CD

12



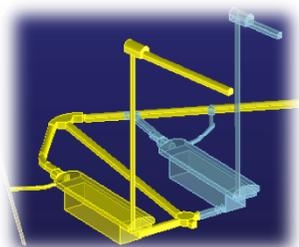
LAr20 Rate Estimate Summary

Rates/data sizes of dominating processes

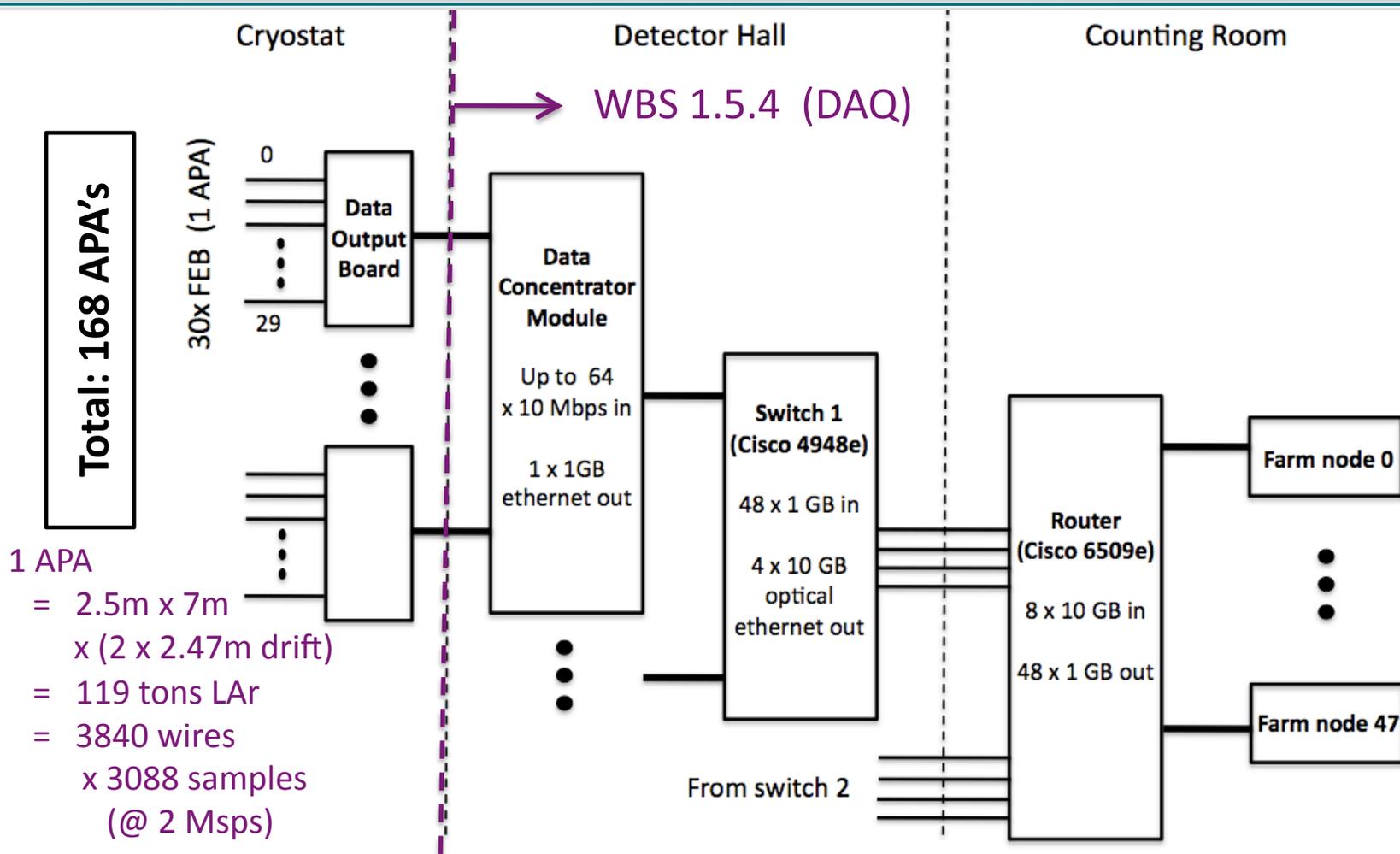
- All numbers are per APA
- Zero-suppression assumed, unless noted otherwise
- “Instantaneous data rates” are for 1.6ms timeframe
- “Avg data rates” factors in rate for ‘rare’ processes

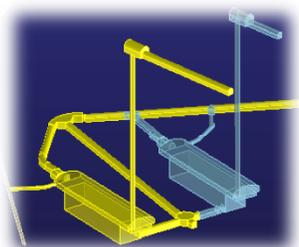
1 APA (out of 168)
 = 2.5m x 7m
 x (2 x 2.47m drift)
 = 119 tons LAr
 = 3840 wires
 x 3088 sample (@2 Msps)

Process	Rate (kHz/APA)	Samples (per APA)	Inst. Data Rate (Mbps)	Avg. Data Rate (Mbps)
Generic 1.6 ms interval (not zero-suppressed)	0.63	1.2×10^7	92,000	92,000
Cosmic ray muons	0.0006	5×10^4	400	0.4
10 GeV EM shower (*Avg. at 0.6 Hz CR μ rate)	—	1×10^6	9,000	9*
Radioactivity: U/Th (γ 's)	~ 1	40	0.48	0.48
$^{39}\text{Ar}/^{85}\text{Kr}$ (β 's)	42	24	12	12
Electronics noise (not common mode)	2	15	0.4	0.4



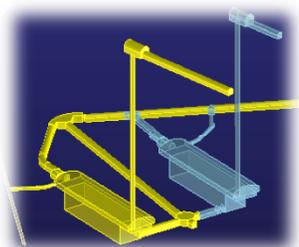
DAQ System Data Flow Diagram





Component Count

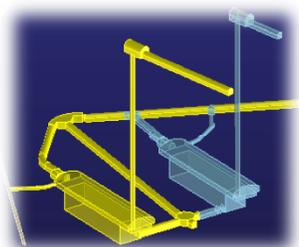
Quantity	Description
84	Data Concentrator Modules
2	Ethernet Switches (Cisco 4948E or similar)
1	Ethernet Switch Chassis (Cisco 6509E or similar), with
1	8-port input optical 10-GB ethernet interface module, and
1	48-port output 1-GB ethernet interface module
48	Data Farm compute nodes
1	Readout Supervisor compute node (not shown in Figure)
1	Routing Master compute node (not shown in Figure)
1	Master timing unit + GPS receiver (not shown in Figure)
14	Slave timing units (not shown in Figure)
1	Run Control compute node (not shown in Figure)
1	Slow Controls compute node (not shown in Figure)



Data Concentrator Module

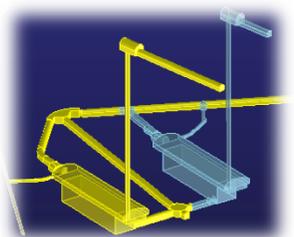
- Based on NOvA Design
 - Up to 64 inputs, each 4 twisted pair (RJ-45) @ 200 MB/s. We could eliminate final 30-1 Mux stage in the cold volume, and deploy 1 DCM for every 2 APA's.
 - This would give 84 DCM's total (NOvA has 180), distributed geographically.
 - FPGA combines raw data from inputs
 - PowerPC processor (running linux) for configuration, buffering & routing
 - Processor generates ethernet packets
 - 1 x 1GB (copper) ethernet out
- NOvA DCM's now being commissioned for 'Near Detector on Surface' operations.





Timing System

- Also Based on NOvA Design
 - Note NuMI beam spill signal:
 - Spill length $10 \mu\text{s}$, rep rate 2 s, similar to LBNE
 - Transmission to Minnesota via internet: typ latency 600 ms, 99% transmission within 5s. Again can assume similar for LBNE. NOvA uses signal downstream of DCM's in trigger farm/buffer nodes.
 - Above implies free-running (untriggered) readout through DCMs.
 - All FEBs must have synchronized clock signals
 - Timing system central facility: GPS receiver & Master timing unit
 - Distributed Components: slave 'Timing Distribution Units'
 - Distribute synchronized clock signal to front-ends via DCMs.
 - TOF/2 to account for propagation delay over timing backbone
 - Probably want 1 TDU for every 6 DCMs / 12 APAs, 14 TDU's total.
 - We may need to worry about propagation delay within the cryostat.



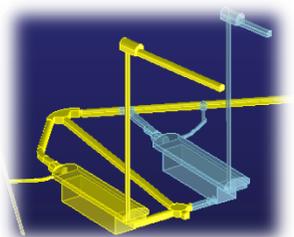
Cisco 4948E Ethernet Switch



Cisco 4948E Ethernet Switch

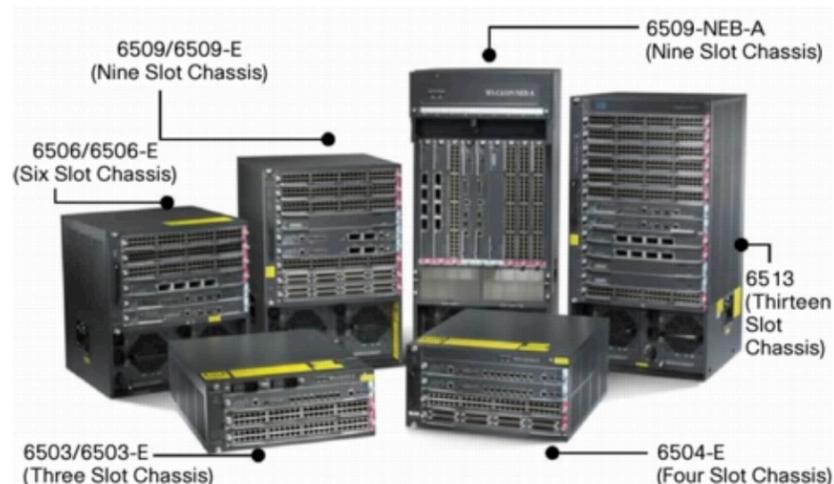
- 48 x 1GB in
- 4 x 10GB optical out
- 175 MB memory:
automatically buffers data if link is in use

- Located in Detector Hall
- Concentrates data from 42 DCMs
- DCMs just send packets
- Optical uplink to Counting Room

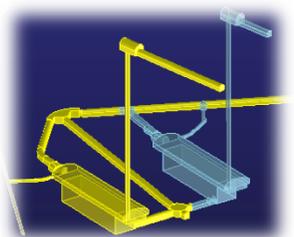


Cisco 6509 E Router System

- Located in Counting Room (on surface).
- Utilize Cisco 6509 E chassis
- 1 x 8-port 10GB blade
 - Input from 2 x 4948E's
- 1 x 48-port 1 GB blade
 - Output to data farm

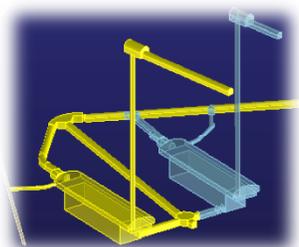


Cisco 6500 Router Chassis Series



Data Farm & Online Computing

- CPU:
 - 48 compute nodes for event building/triggering
 - Output to data logger / disk farm
 - To account for events spanning multiple time blocks, will need to send data for a given block to two (or three) separate nodes, so that a given node has data for blocks (N-1), (N) and (N+1).
- Disk farm:
 - 100 TB staging disk system in preparation for transfer of data off-site via internet.
 - Should be sufficient to store triggered data for 5 days (even if no prescaling of CR muons)

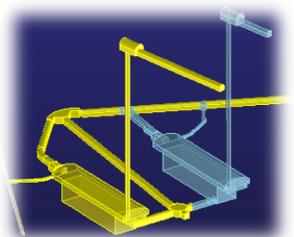


Bandwidth Table

- Assume 15 Mbps out from each APA (^{39}Ar)
 - So 1 Mbps per FEB into each DCM input port

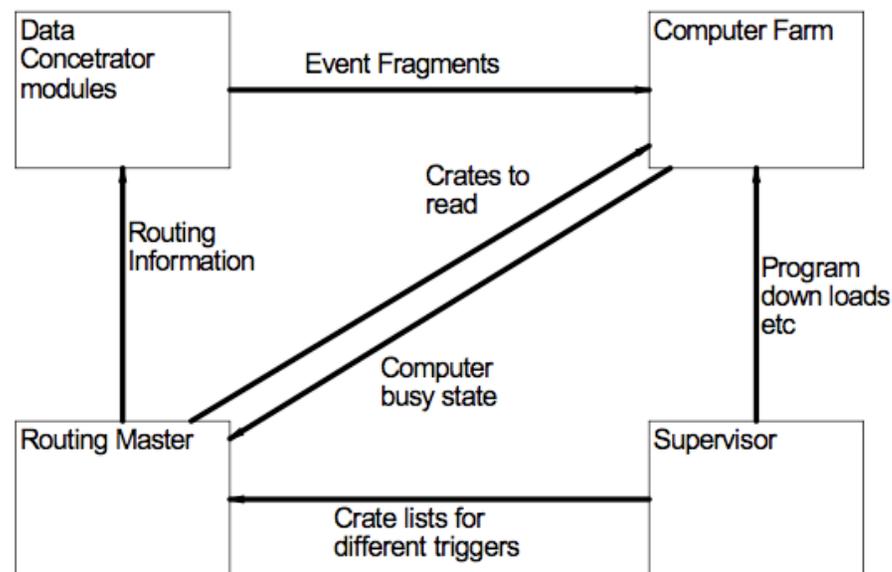
Component	Rate In	Bandwidth In	Rate Out	Bandwidth Out
DCM	60 × 0.5 Mbps	64 × 200 Mbps	30 Mbps	1 Gbps
Cisco 4948E switch	42 × 30 Mbps	48 × 1 Gbps	1.3 Gbps	40 Gbps (4 × 10 Gbps)
Cisco 6509E router	2 × 1.3 Gbps	80 Gbps (8 × 10 Gbps)	2.6 Gbps	48 Gbps (48 × 1 Gbps)

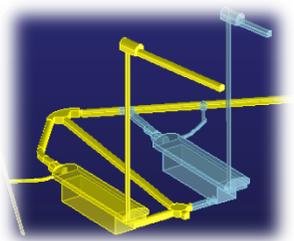
- Comfortable headroom at every stage



Network/Data flow management

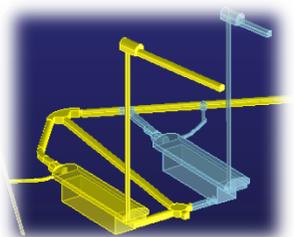
- Routing Master
 - Tracks availability of farm nodes & provides routing info to DCMs
- Supervisor
 - Interface with farm nodes: downloads, collation of statistics, results from processing of data.





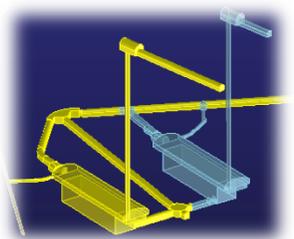
Power Requirements / Racks

- Power in the Detector Hall:
 - DCM's: $<50\text{ W}$ each $\times 84 \rightarrow 4.2\text{ kW}$.
 - Cisco 4948E's: 275 W max $\times 2 \rightarrow 550\text{ W}$.
 - Assume power supplies w/ 75 % efficiency
 - Brings power for these items to 6.3 kW overall
 - ~Double this to account for other parts: \rightarrow **15 kW**
- Need ~ 14 racks: (1 Timing unit + 6 DCMs)
 - See M. Johnson talk (docdb-3020) for discussion of rack locations, plan for grounding/shielding



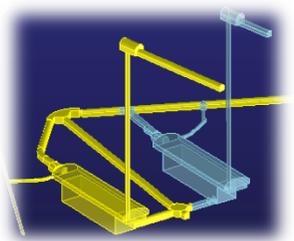
Examples of Costs

- Data Concentrator Modules:
 - NOvA costs: 2.2 k\$ per DCM
 - 84 modules needed + spares → **~220 k\$**
- Networking:
 - Cisco 4948 E: 2 @ 12 k\$ → **24 k\$**
 - Cisco 6509 + 2 blades: → **43 k\$**
- Disk Farm:
 - A 100 TB system today would be **100-150 k\$**
 - Looking ahead, conceivably such a system would be packaged as 2 x Raid-6 arrays of 8 x 10 Tbyte disks (incl. parity & spares), and would be cheaper by a factor of 2-3 (or more).



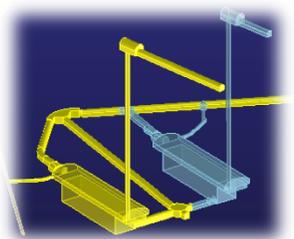
Risks, Mitigations & Prototyping

- Rates due to radioactivity, cosmics higher than expected.
 - Mitigation: Design for comfortable bandwidth margin.
- Sparsification at front-end not as effective as expected.
 - Mitigation: Put more processing capacity into DCM's.
 - Prototyping: The ~800-ton instrumented membrane cryostat prototype will be extremely helpful.
- DUSEL LAN/WAN not reliable.
 - Mitigation: augment local data storage capacity.



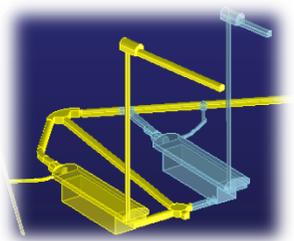
Work Remaining for CD-1

- Refine estimates of rates and data size.
 - Tune DAQ architecture to expected data flow.
 - Get more reliable component counts.
 - Provide more technical detail for the following systems:
 - Timing system
 - Run Control
 - Slow Controls
 - Interfaces with Veto, Photon & Cryogenics systems
 - Refine to cost & schedule estimates.
-

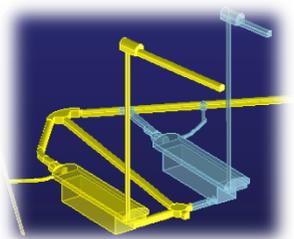


Summary

- The DAQ system design depends on detector/front-end electronics configurations. The design is in early phases relative to systems upstream.
- Evaluation of physics/background rates & data sizes shows modest throughput requirement, despite high channel count/sampling rate.
- A system exploiting commercial ethernet switch arrays is viable.
- And, the main custom components (Data Concentrator Modules, Timing system) can be based on NOvA modules now being constructed/commissioned.
- The LAr20 DAQ reference design is conservative, and could be built using components available today.

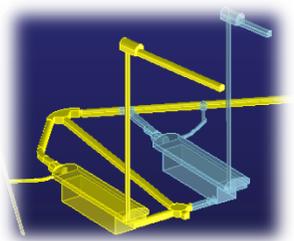


Backup Slides



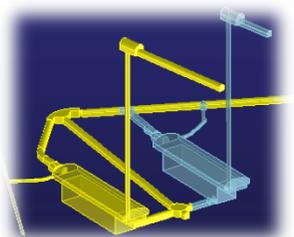
DAQ System Scope

- Configuration, control of detector operations
- Generation & distribution of timing signals
- Collection of data from TPC, Photon & Veto systems
- Event building, filtering, fast processing
- Slow controls, including experiment conditions monitoring
- Operator interface (Run Control)



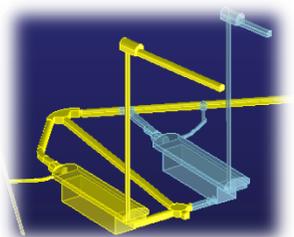
Requirements – I

- Signals to and from detector elements
 - Provide configuration/control signals to TPC et al.
 - Collect raw data via xface w/ front-end electronics
 - 2 MHz sampling rate at the front-ends.
 - 1.6 ms maximum drift time.
- Timing
 - Distribute synchronized clock signals
 - Incorporate LBNE beam spill signal information



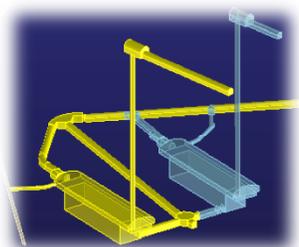
Requirements – II

- Event building, filtering and storage
 - Unpack, assemble data into single event records
 - Note free-running data flow from front-ends
 - Note 1.6ms maximum drift & approx. readout block times
 - DAQ must handle events spanning multiple readout blocks.
 - Event record format must be amenable to analysis
 - DAQ must compute & apply event selection criteria
- Pulsing/testing of front-end electronics
 - Initiate, collect, record detector checkout data
 - DAQ must be operational during electronics installation & commissioning



Requirements – III

- Run Control and Data Monitoring
 - Allow experimenters to operate detector
 - Monitor data collection integrity & data quality
- Slow Controls
 - Interface w/ hardware operating in support of detector systems (power supplies, etc.).
 - Monitor detector, subsystems, environmental conditions
 - Interface w/ Cryogenics control systems to provide information/interlock signals for detector operations



DAQ System Overview

Key Features

Costs

\$2.8M DI, 36% Contingency

Design

Draws from existing/designed NOvA, MicroBooNE, D-Zero, etc., systems

Exploits commodity commercial components

48 Gbps throughput capacity for 1.5 Gbps average rate (whole detector).

Acquisition

In-house construction + commodity computing/networking components

Risks

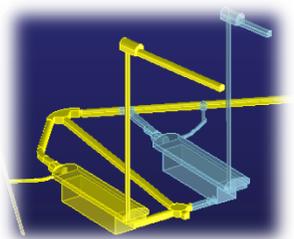
Rates, Sparsification, DUSEL network bandwidth

Prototypes

NOvA components

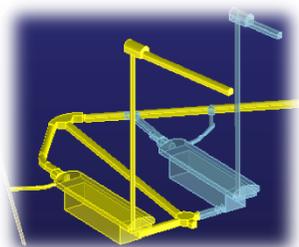
Critical Milestones

None



Alternatives

- Put more processing at interface w/ cold electronics. Could carry out additional sparsification/data compression if needed.
- Build separate event records from each APA.
 - Would make network/switch middle stage much simpler
 - Would make offline analysis simpler for events contained within 1 APA
 - Would make offline analysis more complex for events spanning multiple APA's.



Contributors

Argonne – xface w/ Veto System

Jon Paley

BNL – xface w/ Front End

Veljko Radeka

Gianluigi De Geronimo

Craig Thorn

Bo Yu

Columbia U. – DAQ Architecture

Leslie Camilleri

Cheng-Yi Chi

Mike Shaevitz

Bill Sippach

Bill Willis

FNAL – DAQ Architecture

Mark Bowden

Marvin Johnson

Indiana U. – Project Management

Jon Urheim

Oxford U. – DAQ Architecture

Giles Barr

SMU – Optical drivers

Tiankuan Liu

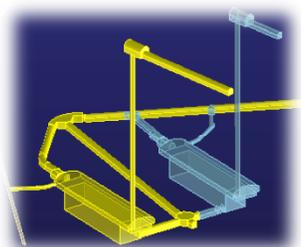
Jingbo Ye

Ping Gui

Yale U. – Online/Offline Software

Eric Church

Note: Organizational structure currently under development



Schedule
