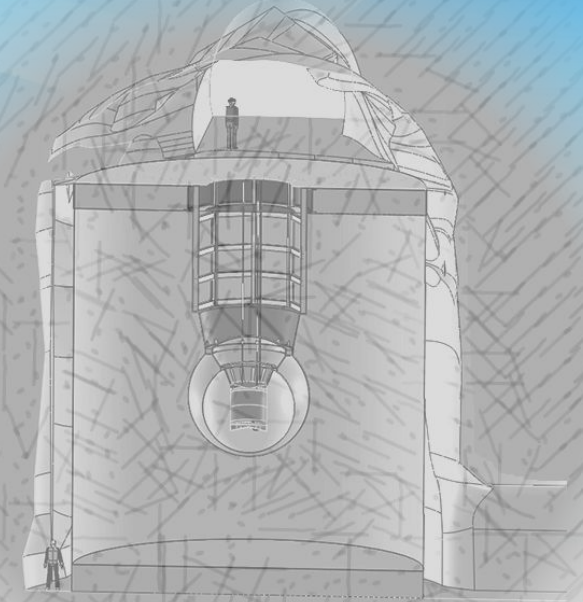




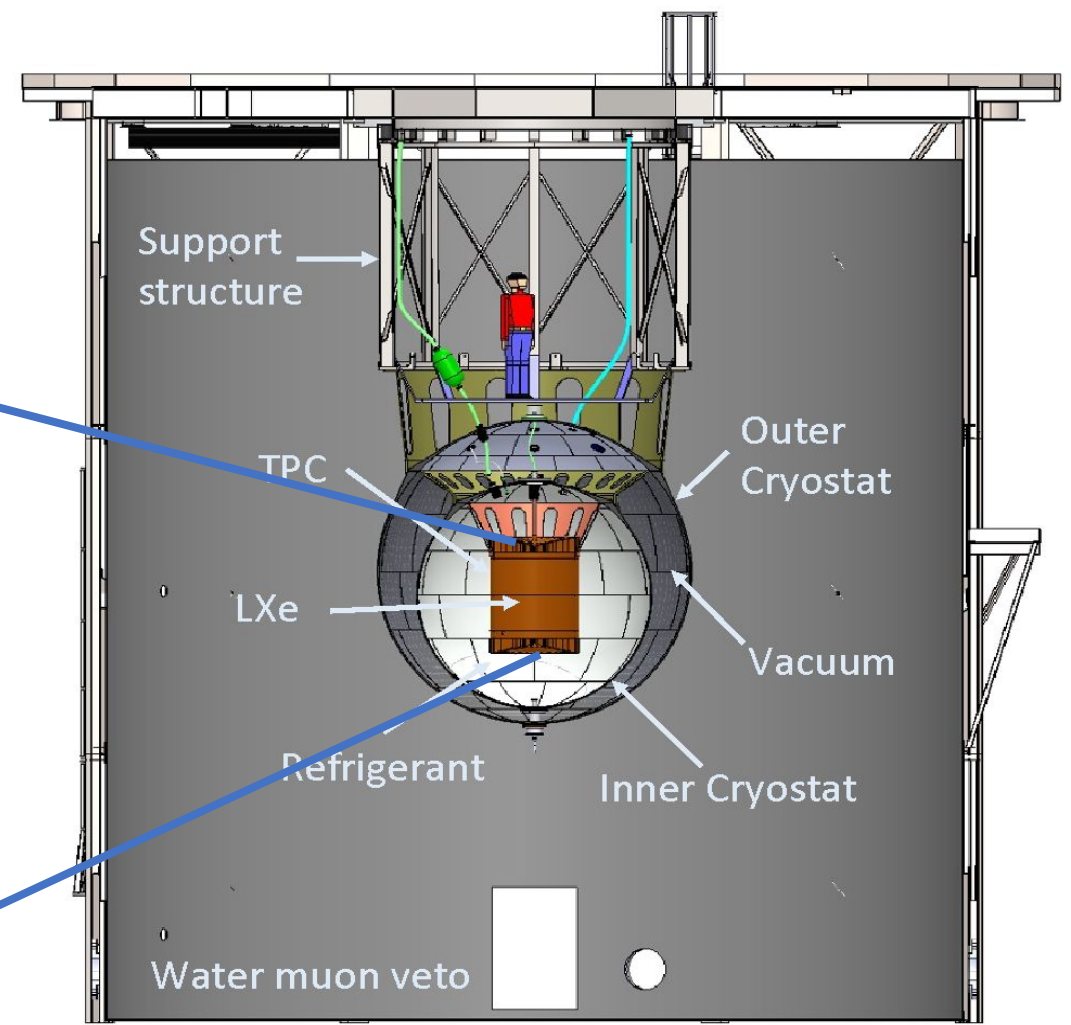
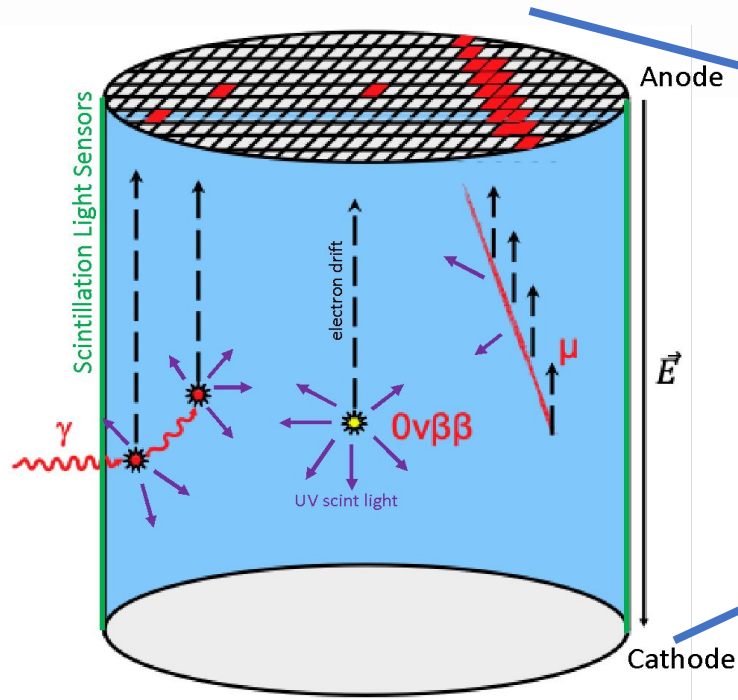
# Conceptual Design of the DAQ System for the nEXO Experiment

Larry Ruckman on behalf of nEXO Collaboration  
Nov. 21, 2024 - RDC 05 Parallel Session



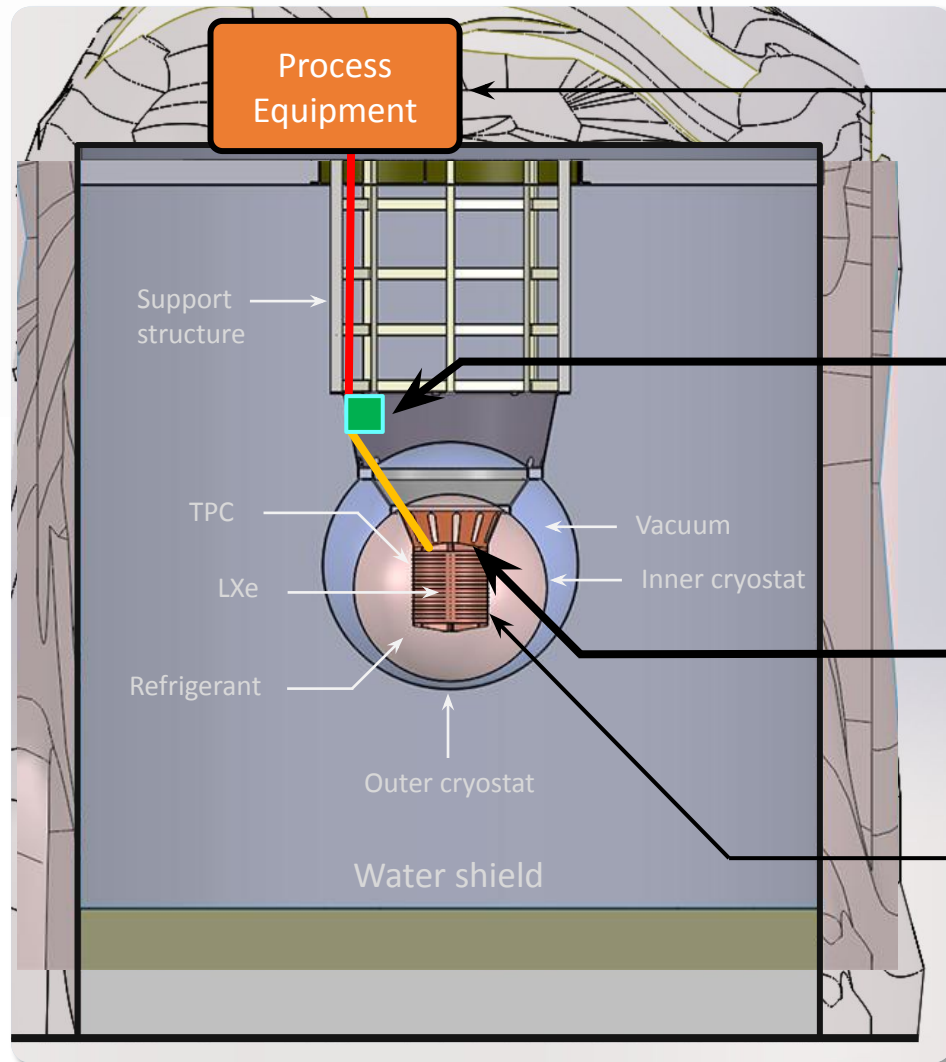
# Technical Overview

- Scintillation - Charge Readout Electronics (**CRE**)
- Ionization - Photon Readout Electronics (**PRE**)
- Muon Veto - Outer Detector (**OD**)



neutrinoless double-beta decay in xenon-136 to explore the fundamental properties of neutrinos and determine whether they are their own antiparticles

# Technical Overview



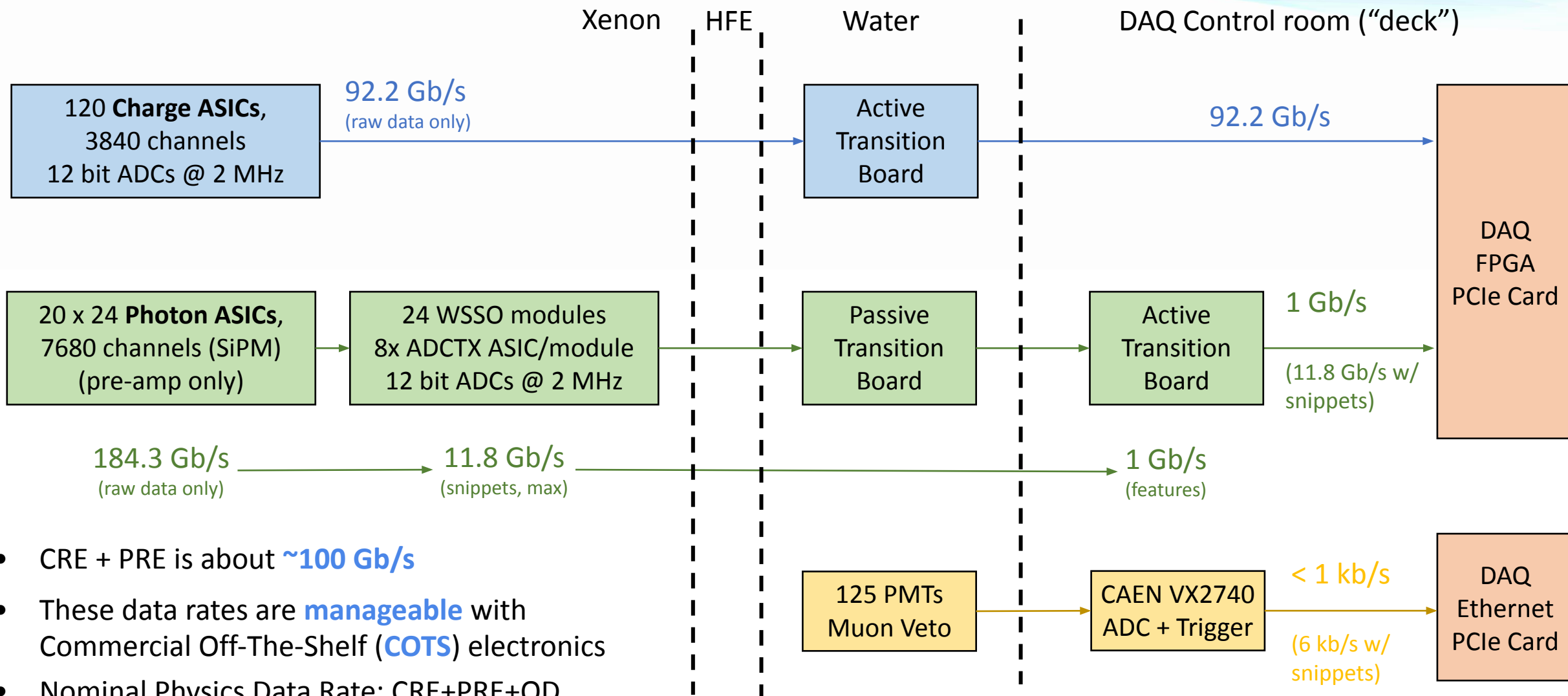
**Backend readout for CRE & PRE:**  
DAQ, slow controls, trigger processing, clock distribution, power supplies, etc.

**Warm transition cards (LVDS to Fiber) for CRE & PRE:**  
Multiplex LVDS to fiber, LV + bias handling/distribution, slow controls processing, etc.

120 Charge ASICs, 3840 channels  
12 bit ADCs @ 2 MHz **Charge**

480 Photon ASICs, 7680 channels (SiPM)  
12 bit ADCs @ 2 MHz (waveform snippets) **Photon**

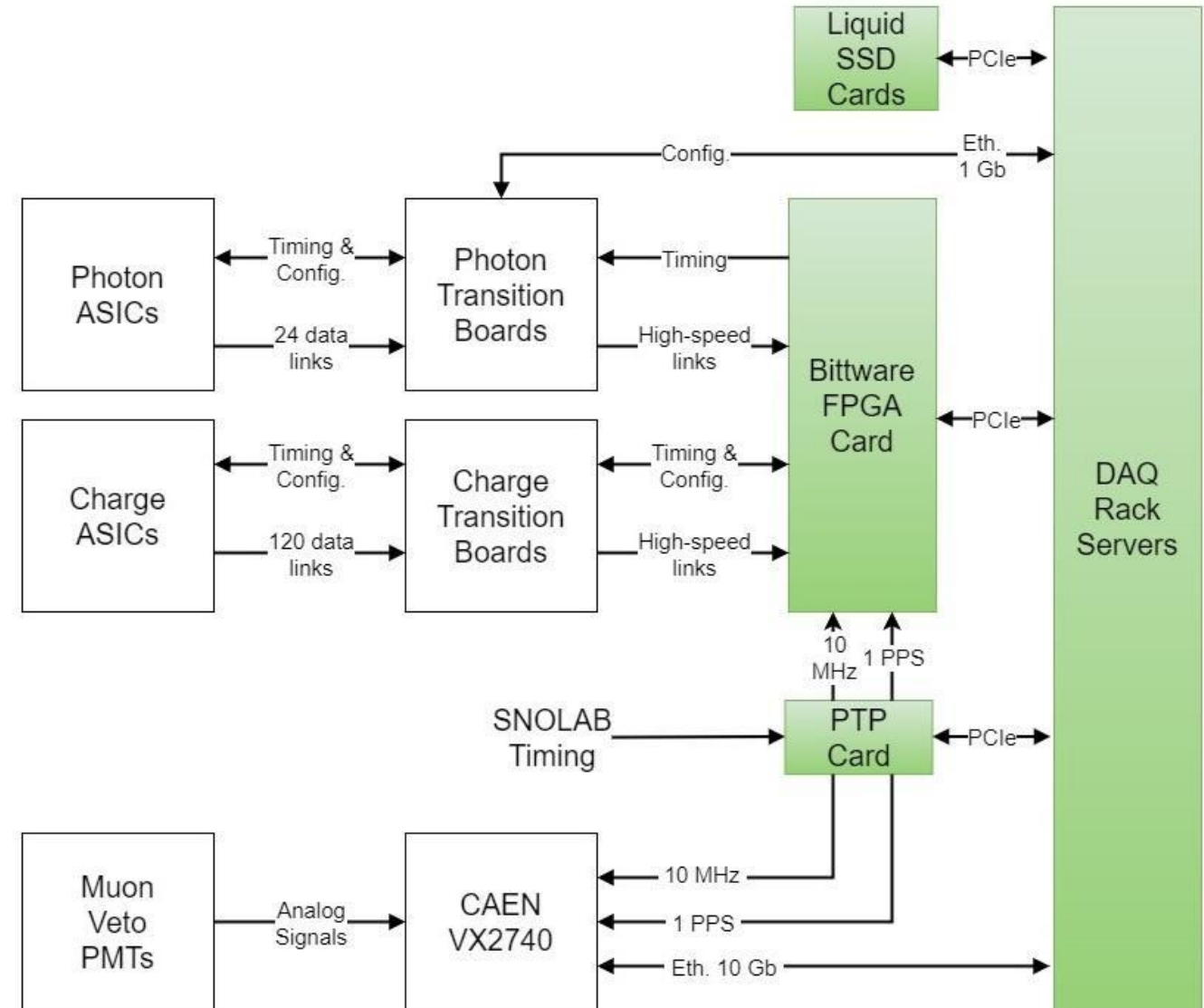
# System Data Rates



- CRE + PRE is about **~100 Gb/s**
- These data rates are **manageable** with Commercial Off-The-Shelf (**COTS**) electronics
- Nominal Physics Data Rate: CRE+PRE+OD
  - **4 MB/s (127 TB/year)**

# DAQ Hardware Structure

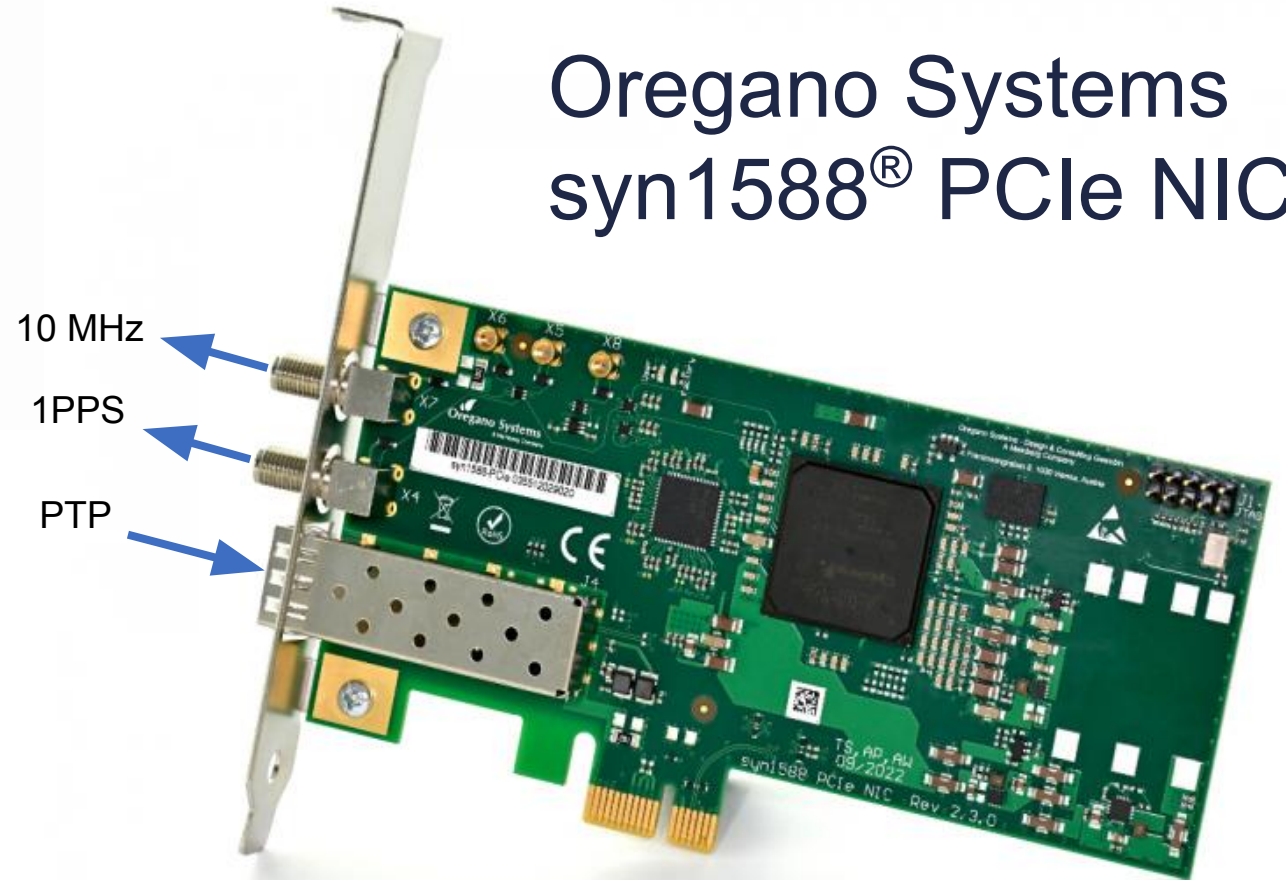
- **Compact design:**
  - 1 rack server for all DAQ hardware (and run control)
- COTS-based design to remove risk of custom electronics development
- Reuse of well understood, mature, open source firmware to reduce the development time/cost
- “Green” blocks are in our DAQ scope



# SNOLAB Timing Receiver

- DAQ is responsible for receiving the SNOLAB timing and distributing to all subsystems
- SNOLAB offers many options for timing:
  - Network Time Protocol (NTP): 15  $\mu$ s RMS
  - **Precision Time Protocol (PTP): 100 ns RMS**
  - Inter-Range Instrumentation Group (IRIG): 10 ns RMS
  - All RMS measurements above provided by SNOLAB
- **PTP was selected**
  - SNOLAB said “Easy” to provide
  - Able to use COTS receivers to generate **1PPS/10MHz SNOLAB timing signals**
- PTP can be used to satisfy the < 500 ns requirement

## Oregano Systems syn1588<sup>®</sup> PCIe NIC



<https://www.oreganosystems.at/products/syn1588/hardware/syn1588r-pcie-nic>

Example of a COTS PTP(IEEE1588) PCIe Card

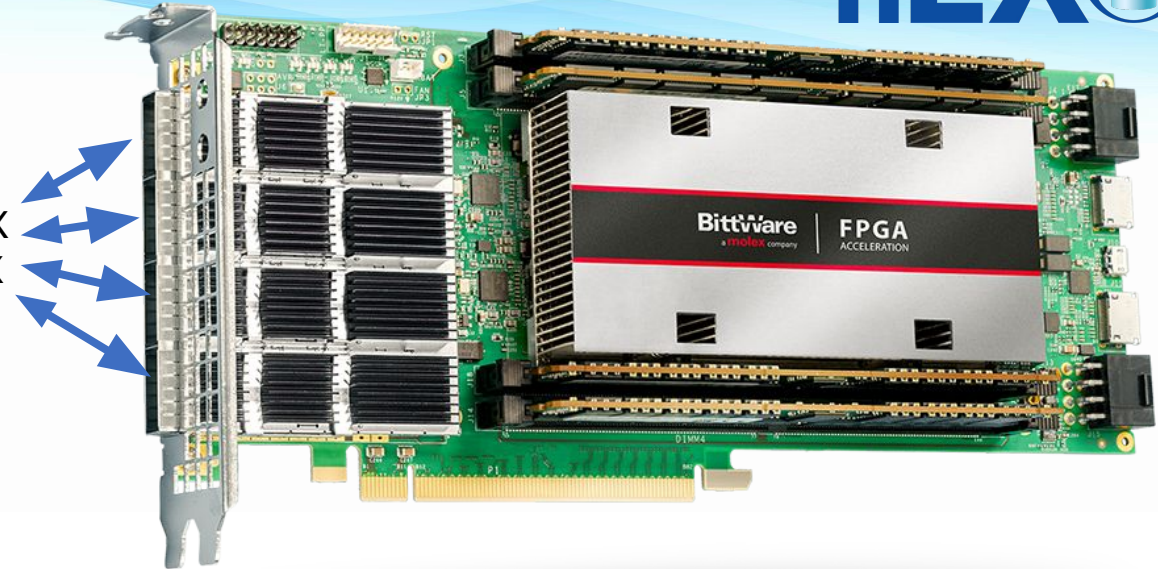
# BittWare FPGA PCIe Card: COTS FPGA Solution



## • BittWare XUPVV8-0049

- Targeting 64 GB DDR4 Memory
  - 16GB/DIMM, 2400 MT/s
- PCI Express Gen3 x 16 lanes
  - > 100Gb/s data bandwidth
- Front panel SMAs for 1PPS/10 MHz
- Passive Cooling Solution

Up to 32 Fiber RX  
Up to 32 Fiber TX

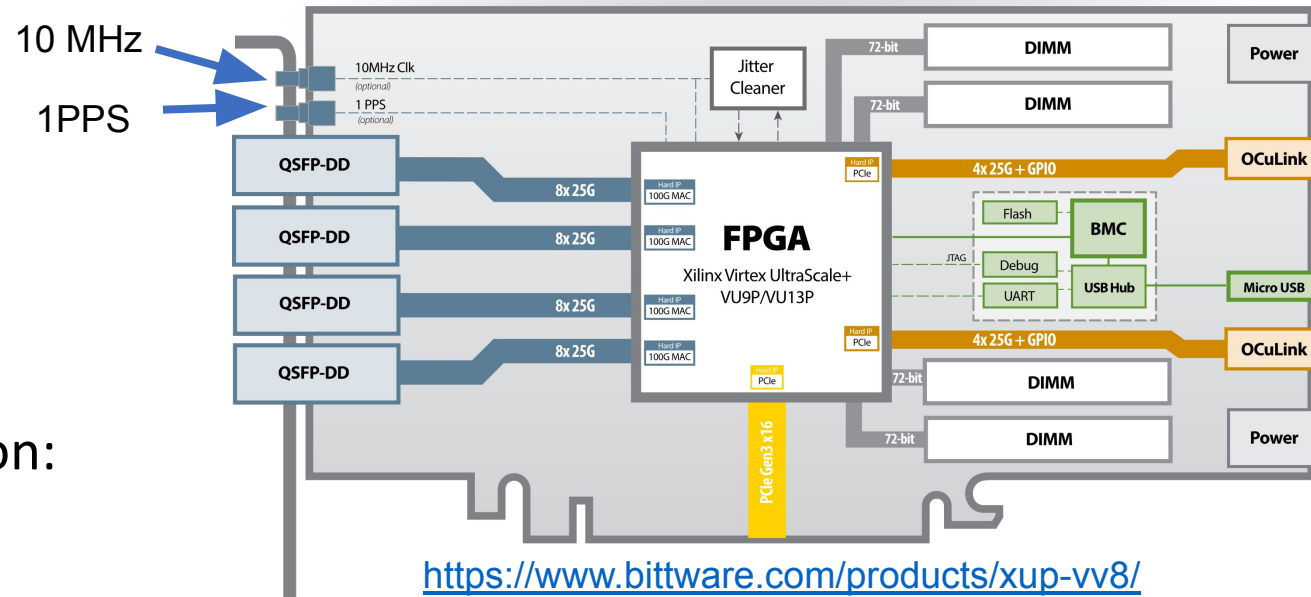


## • 4 x QSFP-DD Cages

- 32 fiber input and 32 fiber outputs
- Up to 25Gb/s per lane switching (NRZ only)
- 8 fiber TX/RX pairs per QSFP-DD

## • Xilinx Virtex Ultrascale+ V13UP loading option:

- 3.4M Flip-Flops
- 1.7M Look Up Tables (LUTs)
- 12k DSP Slices

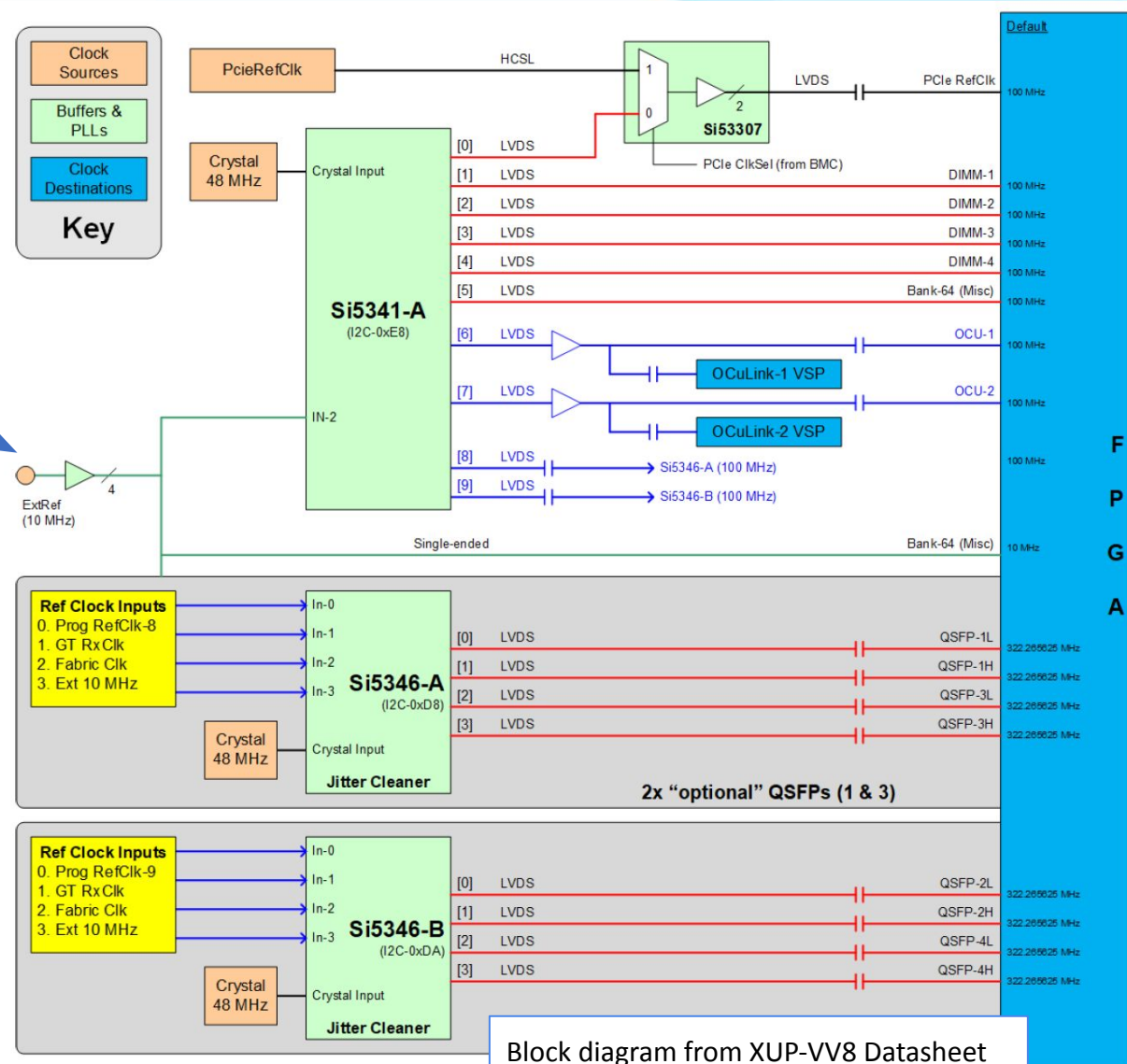


<https://www.bittware.com/products/xup-vv8/>

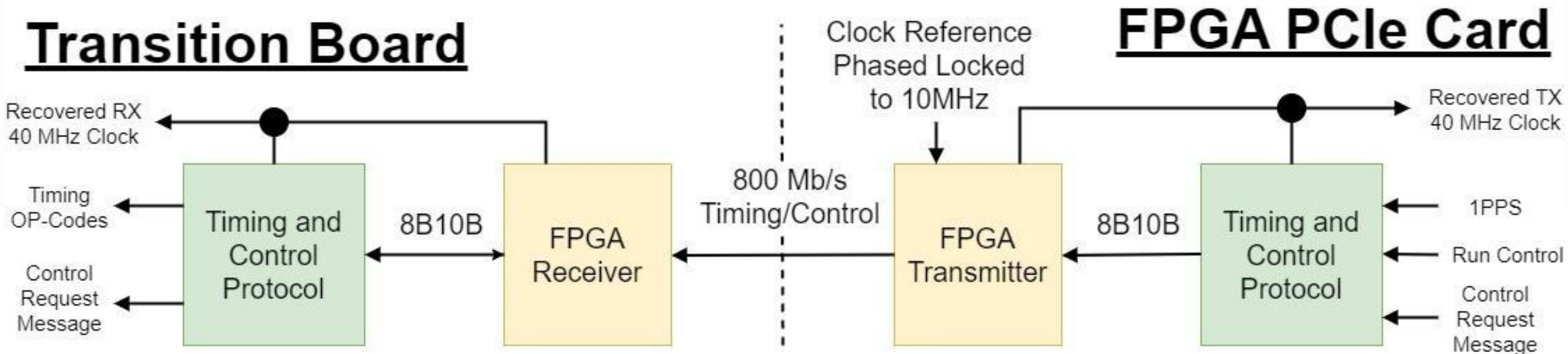
# BittWare FPGA PCIe Card: Clock Circuitry



- 1PPS input is signal conditioned then connected to a general-purpose FPGA input (not shown in diagram)
- PTP's 10 MHz clock is used as a reference input clock for the jitter cleaner Phase Lock Loop (PLL)
- This PLL is used to generate the faster clock references required for the FPGA's multi-gigabit transceivers
- All fiber optic transmission is synchronous to the PTP's 10 MHz clock



# Timing Distribution to a Relative Clock

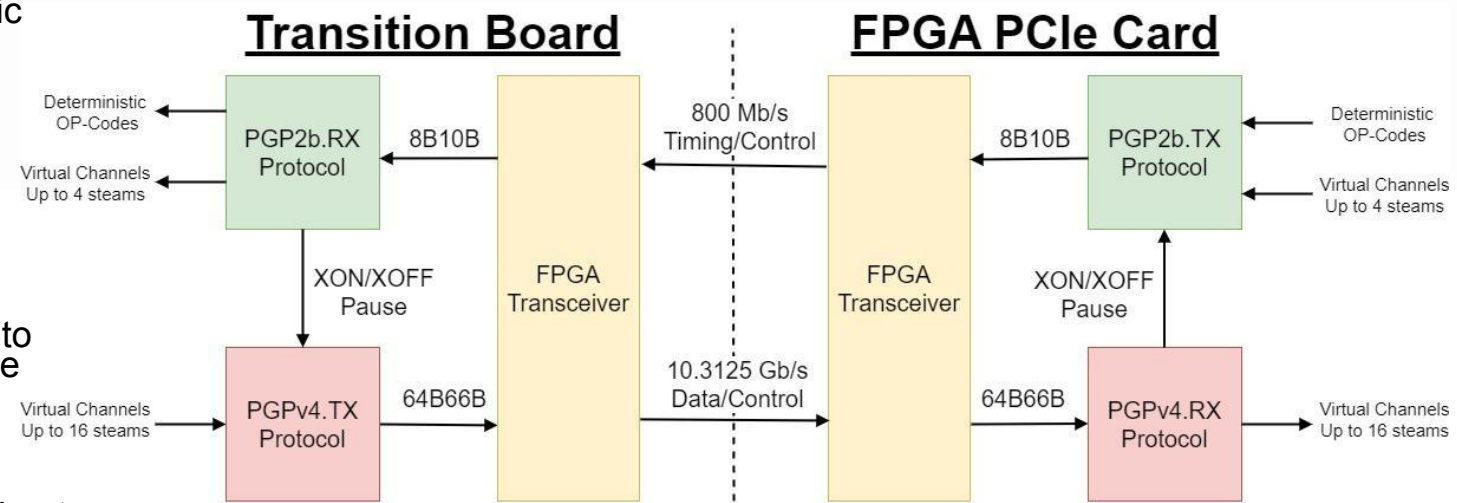


- Transition Board's recovered RX clock is phased locked to PTP 10MHz clock and synchronous to the PCIe card's recovered TX clock
- Able to synchronize timing between light and charge channels to a relative clock

# Timing/Control/Data Link Protocol



- **PGP2b**: Pretty Good Protocol Version 2b
  - Similar to Aurora 8B/10B
  - 10+ years old, open source
  - Developed to interleave data streams and transmit operational codes (OP-codes) with deterministic latency all on same link
- **PGP4**: Pretty Good Protocol Version 4
  - Similar to Aurora 64B/66B
  - 5+ years old, open source
  - Developed for increased bandwidth compared to PGP2b and drops deterministic latency op-code support (low latency instead)
- This **PGP2b/PGPv4 hybrid solution** lets us get the best features out of each protocol type
  - **Distributing the 40 MHz clock** via RX clock recovery at the transition board
  - **Deterministic latency op-codes for executing run control**
  - Interleave data/control streaming (**reduce cabling**)



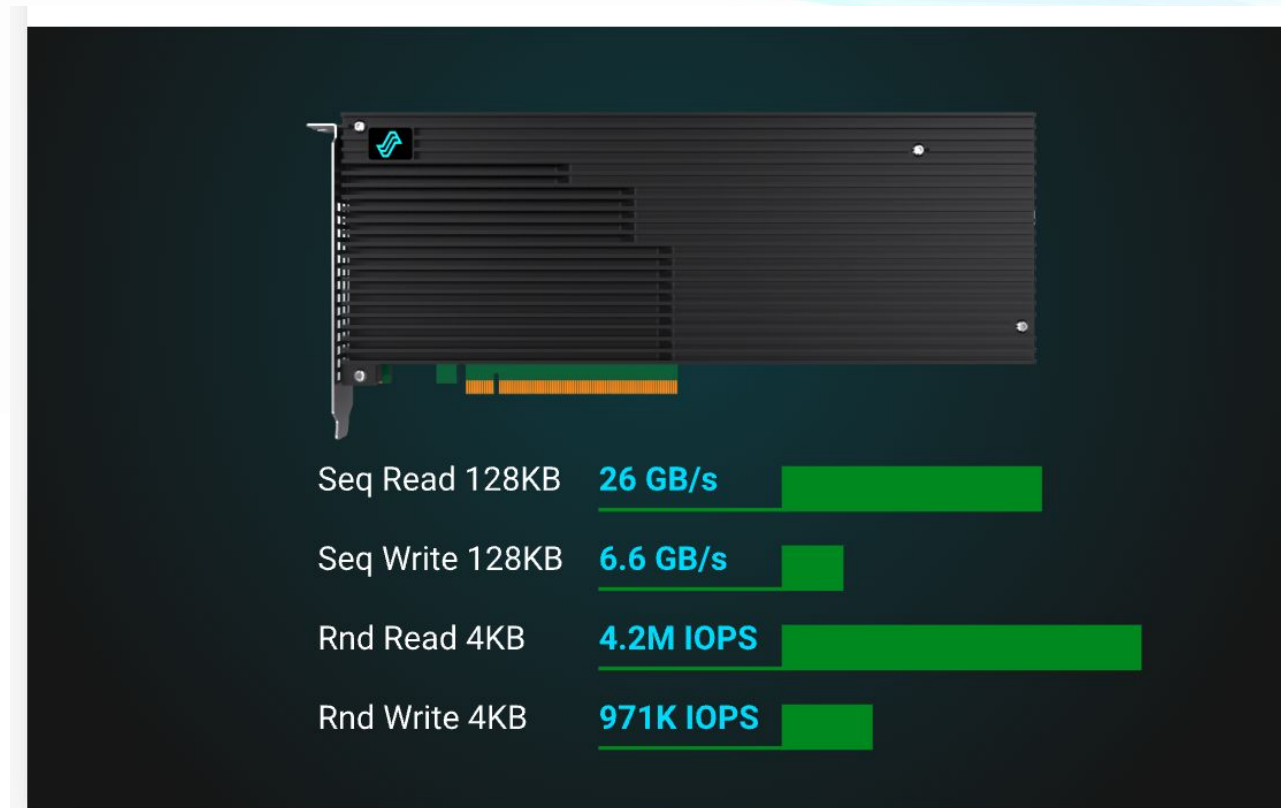
<https://github.com/slaclab/surf/tree/master/protocols/pgp/pgp2b>

<https://github.com/slaclab/surf/tree/master/protocols/pgp/pgp4>

# Calibration data sent to SSDs



- **LQD4500 IO Accelerator**
  - PCI Express Gen4 x 16 lanes
  - Read/Write Raw Bandwidth: ~24 GB/s
  - Raw capacity: Up to 32 TB
  - Endurance: Up to 61.53 PBW (**PetaBytes Written**)
  - **Passive Cooling Solution**
- **Two SSDs should be adequate for the calibration runs**
  - Assume a compression factor of 3 (“worst case”)
  - $100 \text{ Gb/s} / (3 * 8) = 4.2 \text{ GB/s}$
  - 2 hours is 30 TB
  - **Factor of 2 margin** in storage capacity and write bandwidth
- **Modular design architecture**
  - Additional SSD could be added to increase system calibration write performance



<https://www.liquid.com/products/io-accelerators>

# Rack Server with Passive Cooling Support

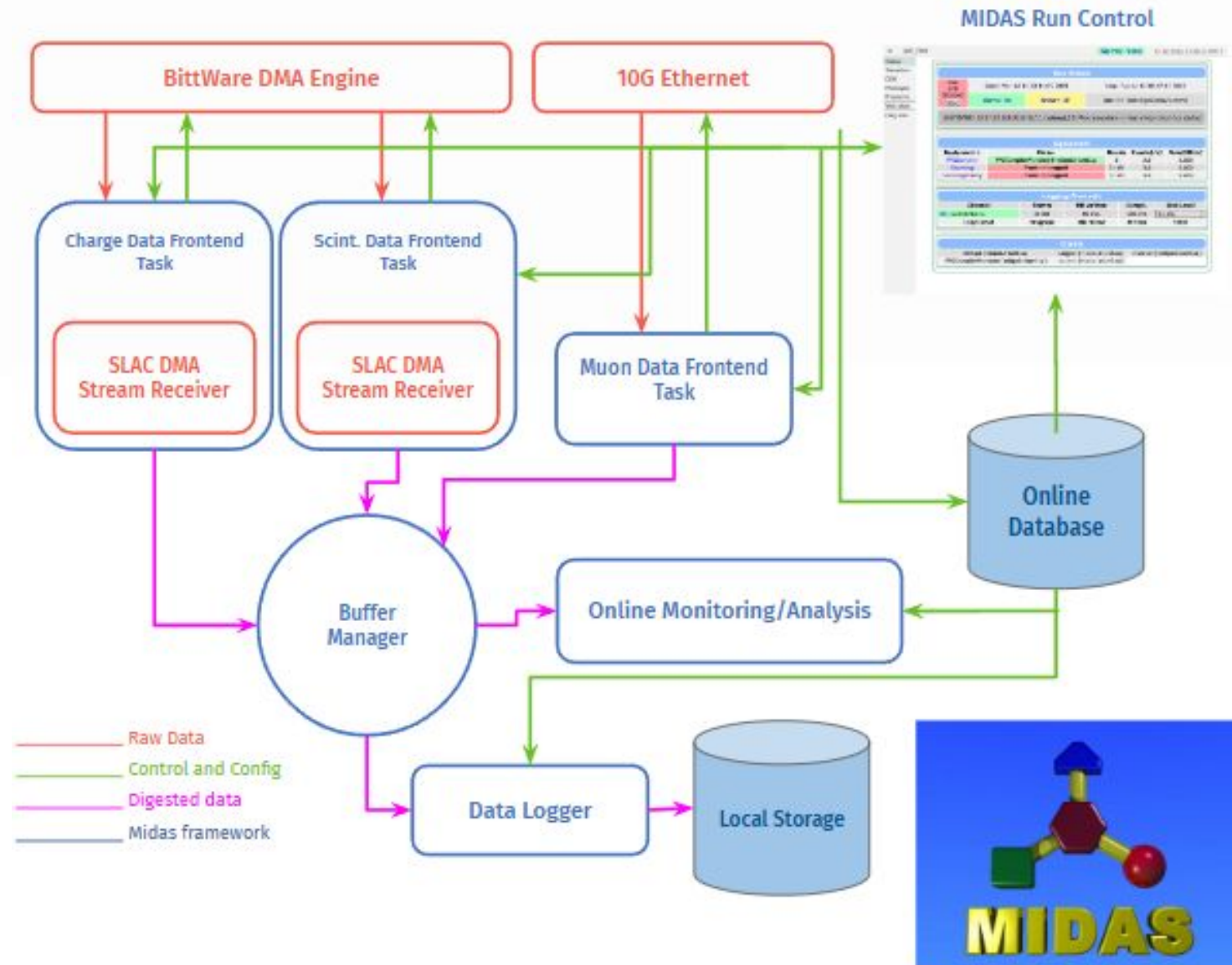


- Plan to use Supermicro AS-4125GS-TNRT for **initial prototyping**
  - **Form Factor:** 4U Rackmount
  - **CPU:** Dual AMD EPYC™ 9004 Series Processors
  - **RAM:** 24 DDR5 DIMM Slots (up to 6TB)
  - **PCIe:** GEN5 x16 lane (FHFL), 9 slots available
  - **Integrated server fans for passively cooled PCIe cards**
    - Required for BittWare and SSDs
  - 4x 2000W Redundant power supplies
- Probably **not** the server for production/commissioning
  - Likely will purchase the latest and greatest by that time



<https://www.supermicro.com/en/products/system/gpu/4u/as-4125gs-tnrt>

# DAQ Software Structure

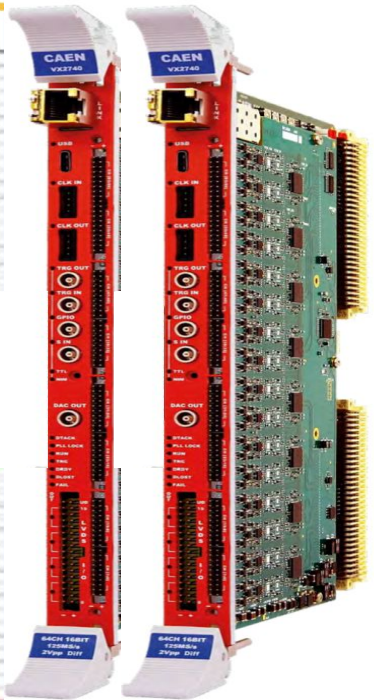
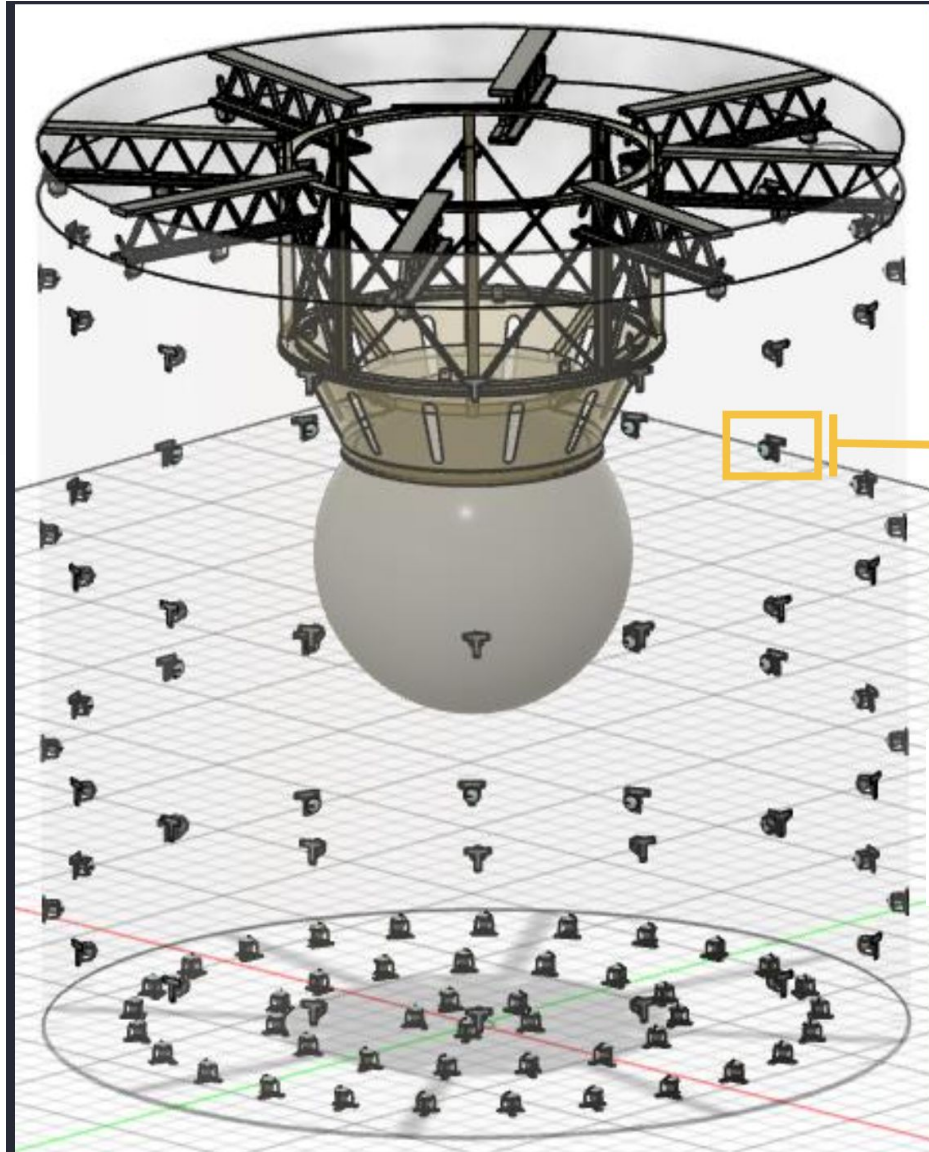


**MIDAS is a C++ based DAQ framework developed and maintained by TRIUMF**

- Tried and tested by several groups including DEAP, TK2 and **SuperCDMS at SNOLAB**
- **Web based run control** that defines the hardware sequence for initialization and running of a detector
- **Web based monitoring**
- **Centralized database** used to archive all hardware and run information
- Built in frontend framework allows users to define how data is acquired from the hardware
- **Event building & Alarm system**
- Data persistency to different formats including binary and ROOT
- **“MIDAS is SNOLAB’s DAQ framework of choice”**



# Outer Detector – Signals and Electronics



- Outer Detector (**OD**) is a water Cherenkov muon veto
- 125 x 8" PMTs from Daya Bay
- 2 x CAEN VX2740 digitizers collect and trigger PMTs data
- DAQ's scope is to take signals out of the digitizers from OD sub-system
- **MIDAS already has VX2740 driver support**

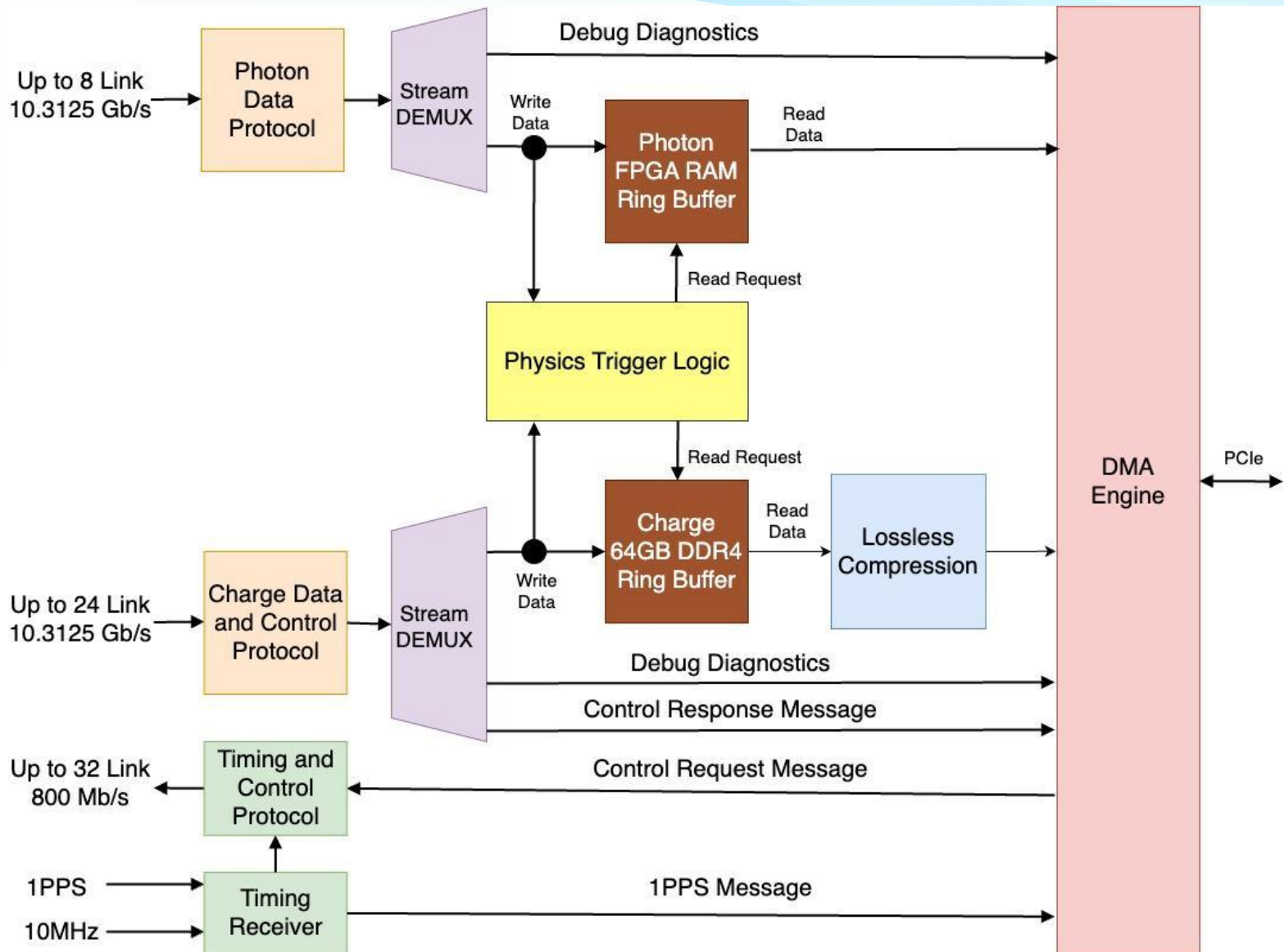
- A solid conceptual design for the DAQ exists, which is based on requirements that flow down from the top-level nEXO science and operation requirements
- A successful CoDR and CD-1 Director Review was conducted in FY2024, where there were no significant findings that would prevent the team from advancing to CD-1 in FY2025
- Looking forward to continue this work starting FY2026 (post CD-1)

# Backup Slides

# BittWare Firmware Block Diagram



- Receive the photon's trigger primitives messages and store in the FPGA's BRAM/URAM
  - > 100 ms buffer
- Physics trigger logic
  - Capability to determine a trigger event based on combinations of charge and photon signals
- Receive the charge's raw data and stored into a 64GB ring buffer
  - > 100 ms buffer
- Using "lossless" compression for charge data
  - Both physics and calibration mode



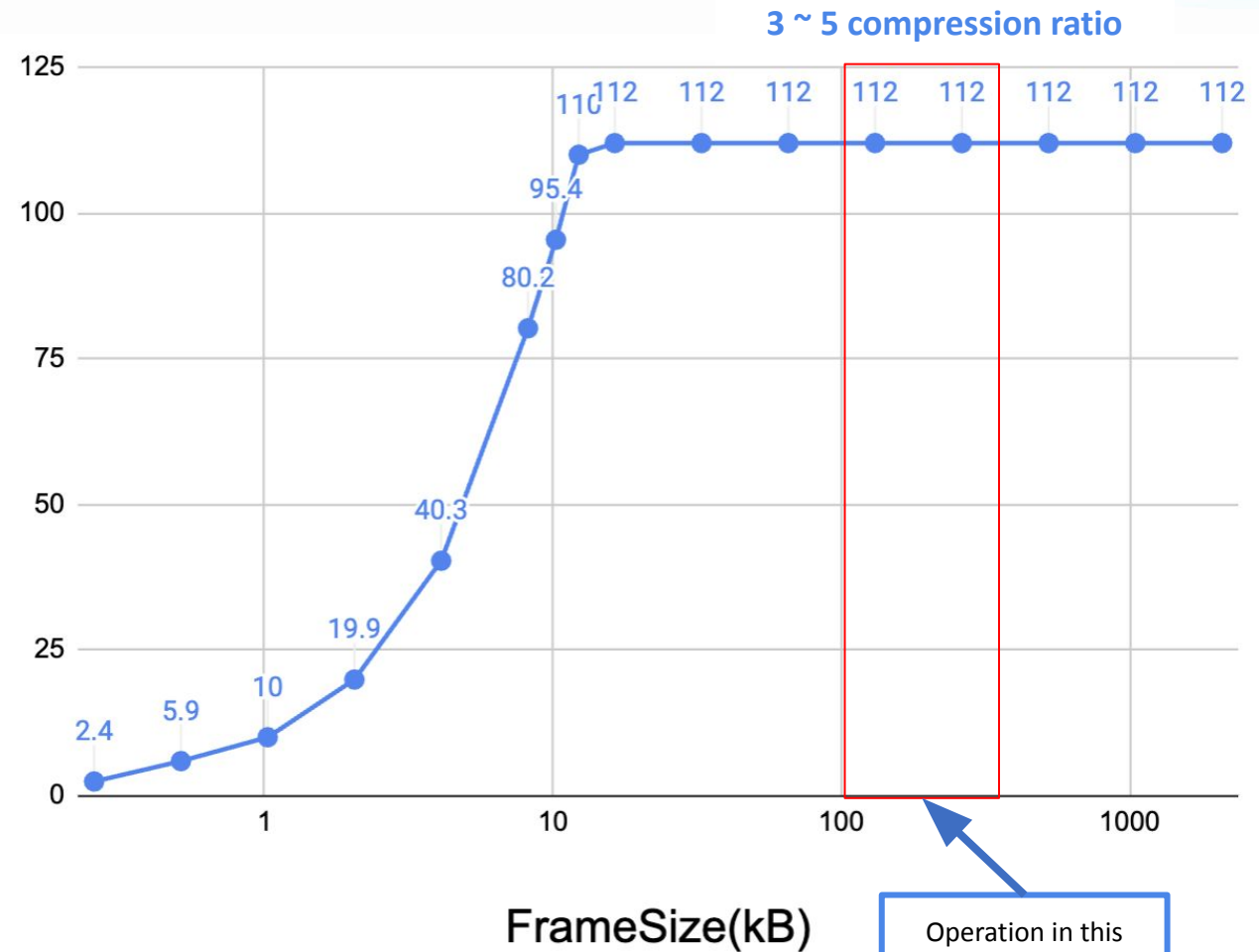
- The compression technique is **Arithmetic Probability Encoding (APE)** vs **Huffman on EXO-200**
- Advantages of APE over Huffman
  - **APE is deterministic**
    - Huffman involves a non-deterministic sort
  - **APE achieves the entropy limited compression factor**
    - More effective/stable when entropy is very low
  - Achieved a **4-8 compression factor on EXO-200 using Huffman**
    - Advertised nEXO noise is similar
  - **Expect to get >3.5 with nEXO using APE**
    - Meets requirements to stream calibration data to the SSDs
- **APE was deployed and demonstrated in protoDUNE**
  - FPGA implementation using AMD/Xilinx High-Level Synthesis (HLS)
  - Parallel compression engines handling groups of 128 channel per engine
  - “lossless” compression of data to storage

# DMA Firmware and Kernel Driver



- **aes-stream-driver**: DMA firmware and kernel driver
  - 5+ years old, open source
- Developed for multiple DAQ experiments to support:
  - interleaved data (up to 2048 destinations)
  - zero-copy memory buffer handling
  - >1 MHz frame rate
    - no batching, small data size
- **Benchmarked the BittWare card's DMA bandwidth with kernel driver**
  - BittWare PCIe Gen3 x 16 lanes
  - x670 Motherboard (PCIe Gen5 capable)
  - Ryzen 9 7900X + 64GB DDR5 4800 MT/s
- Greater than **30.7 Gb/s requirement (= 92.2/3)**

Bandwidth(Gb/s)



Operation in this range for compressed data

# Charge Ring Buffers



- BittWare has 4 x 16 GB DIMMs
  - 64GB total (> 4 second buffer)
- **Benchmarked the BittWare card's DDR DIMM bandwidth**
  - 1 memory controller per DIMM
  - 4 memory controllers in total
- When doing both reads and writes:
  - Benchmarked 57 Gb/s bandwidth per DIMM
  - This is the mode that calibration runs in
- **228 Gb/s total memory bandwidth**
  - 57 Gb/s x 4 DIMMs
  - Greater than **92.2 Gb/s requirement**

