Streaming Grand Challenge Overview

Graham Heyes

February 12th 2019

Jefferson Lab

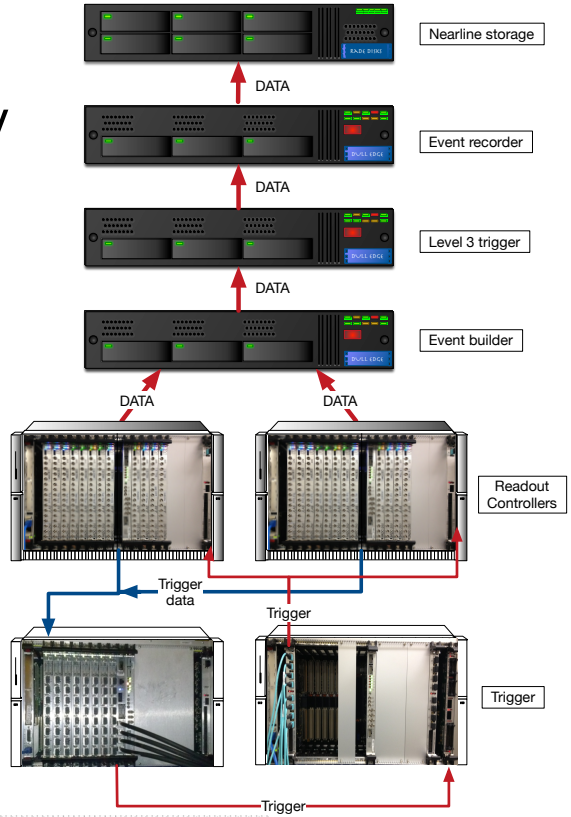
Where are we now?

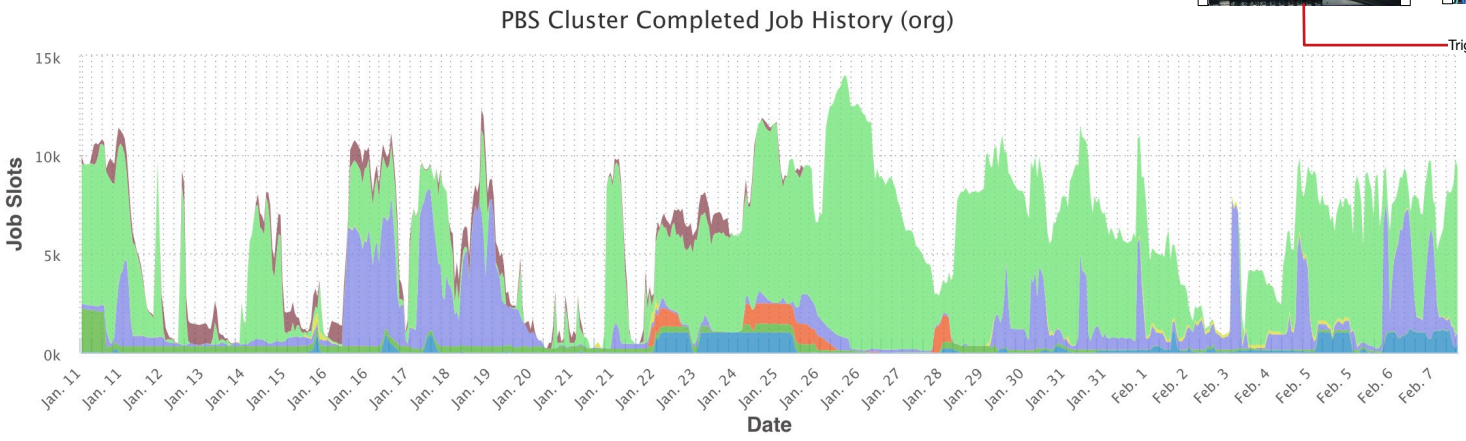
Online :

- Triggered, pipelined readout systems build events online. Sequentially store events in files ordered by event number.

Offline :

- Files of events are processed in steps: monitoring, calibration, decoding, reconstruction, analysis.
  - Data is passed between stages in flat files.
  - Pauses of days/weeks/months between steps.
  - Very little integration between the various steps.
  - Batch farms of fairly homogeneous architecture.





Jefferson Lab

Where is everyone headed?

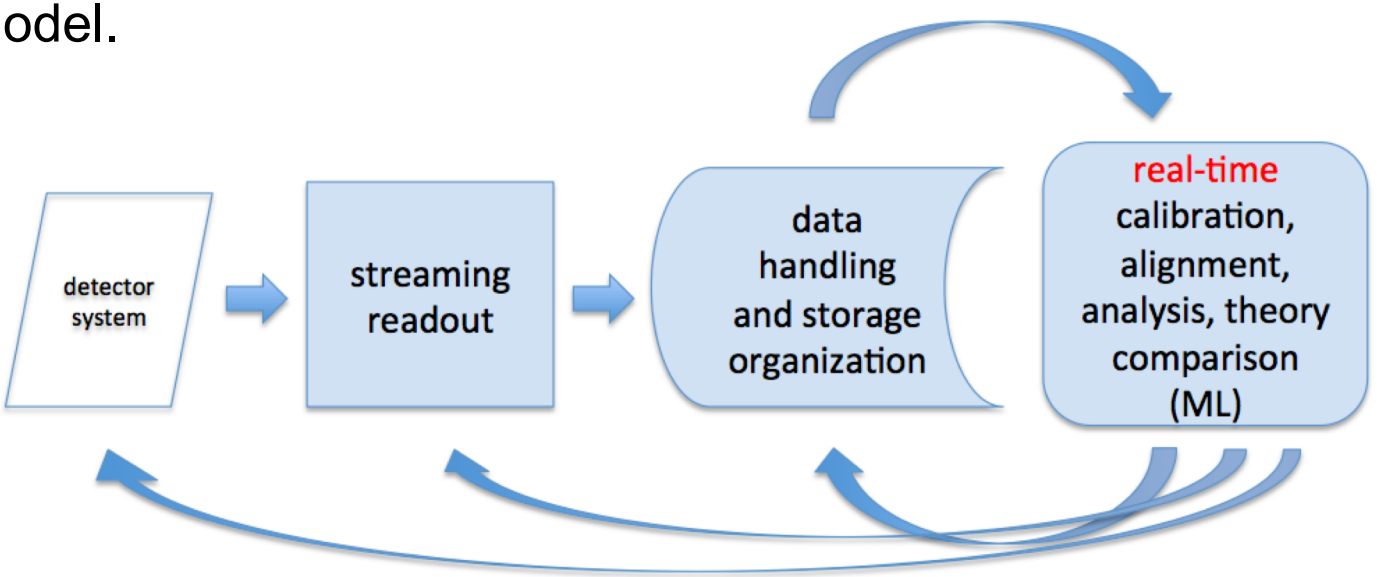
- Several workshops in recent years have explored this topic.
  - Micro-electronics and computing technologies have made order-of-magnitude advances in the last decades.
  - Statistical methods and computing algorithms have made equal advances.
- Online
  - Much interest in triggerless or minimal trigger readout.
  - Streaming readout – parallel data streams all the way from detectors to storage.
  - Rapid online monitoring, data processing (i.e. calibration) and even reconstruction.
- Offline
  - Heterogeneous, distributed computing hardware architectures.
  - Service oriented software architectures.
  - Use of ML, AI and other modern data processing methods.
- The distinction between offline and online is increasingly blurred.

Where do we want to go?

- Several experiments are adding elements such as streaming readout, AI, and real time processing as upgrades to existing systems.

- This approach of “adding on” does not lead to an integrated system that is consistent in approach from DAQ through analysis.
  - LHCb is the closest approximation but stops at online.

- We aim to remove the separation of data readout and analysis altogether, taking advantage of modern electronics, computing, and analysis techniques in order to build an integrated next generation computing model.



Key Elements

- An integrated whole-experiment approach to detector readout and analysis will take advantage of multiple existing and emerging technologies. Amongst these are:
  - "Streaming readout" where detectors are read out continuously.
    - A "stream" is a time ordered sequence of data. It can be real, i.e. network link or backplane, or virtual, i.e. in a database or file system.
  - Continuous data quality control and calibration via integration of machine learning technologies.
  - Task based high performance local computing.
  - Distributed bulk data processing offsite using, for example, supercomputer centers.
  - Modern, and forward looking, statistical and computer science methods.

How do we get there?

- Several of the current LDRD proposals as well as separate on-going efforts naturally fit into the framework of the integrated whole-experiment model of data handling and analysis. They are :
  - Jefferson Lab EIC science related activities.
    - Web-based Pion PDF server.
  - Jefferson Lab and EIC related (as part of the Streaming Consortium proposal to the EIC Detector R&D committee).
    - Crate-less streaming prototype.
    - TDIS TPC streaming readout prototype.
    - EM Calorimeter readout prototype.
    - Computing workflow - distributed heterogeneous computing.
  - LDRDs.
    - JANA development 2019-LDRD-8.
    - Machine Learning MC 2019-LDRD-13.
    - Streaming Readout 2019-LDRD-10.

What is the “Grand Challenge’

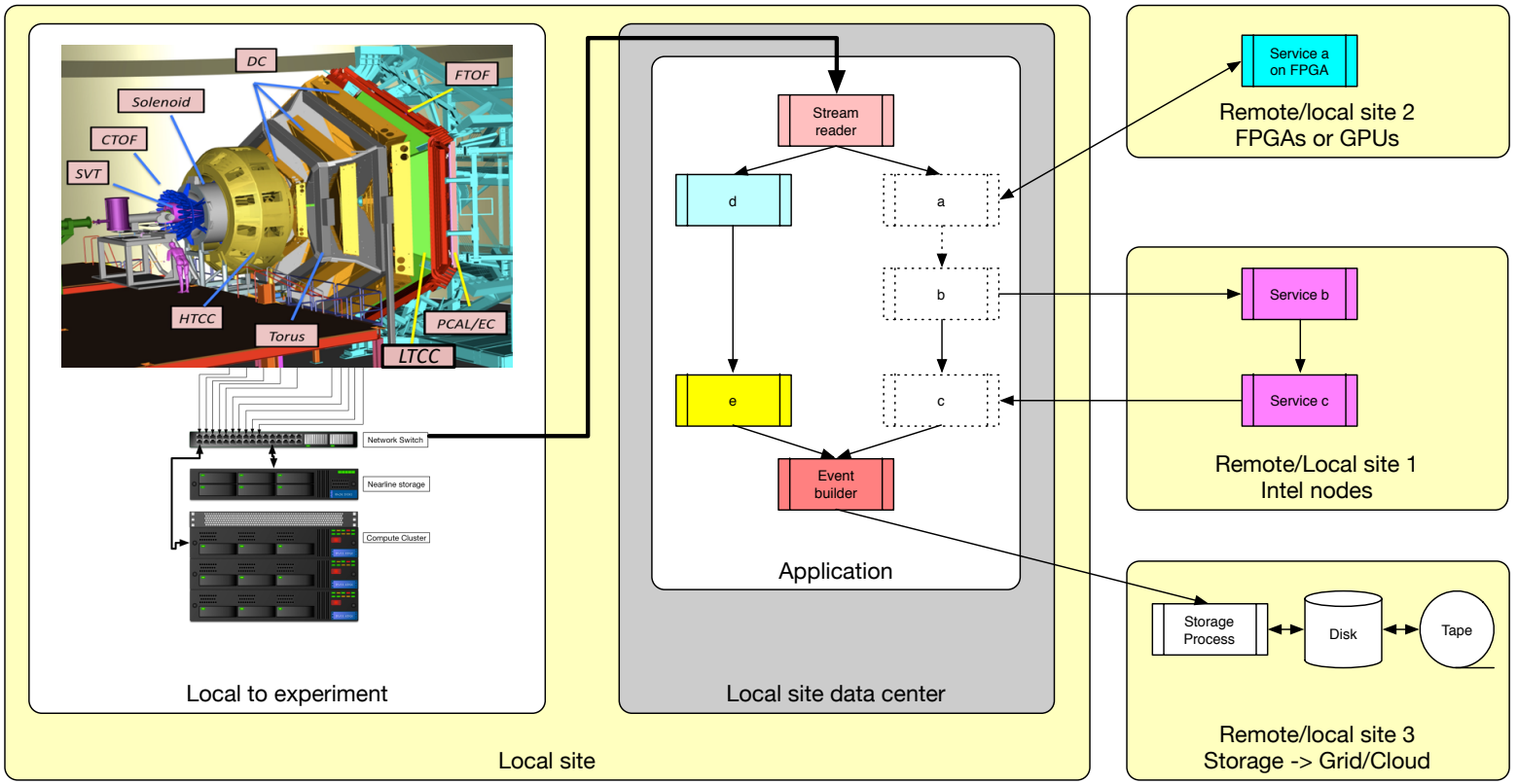
- To develop a proof of concept integrated readout and analysis based on modern and forward looking techniques in disciplines such as electronics, computing, AI, algorithms and data science.

- Long term aim is to develop production systems suited to CEBAF experiments and the Electron Ion Collider.

- We will begin by organizing some of the LDRD proposals and other exploratory work around these themes to achieve proof of concept.

A concept

- Reimagine applications and workflows as nets of “services” processing streams of data.
  - Services can be implemented as software, on traditional CPUs or GPUs, or in firmware on FPGAs.
  - Develop a toolkit of standardized application building blocks – one data type in, another out.
  - Streams route data between services running on appropriate hardware.
  - Services can be local or distributed.
    - Currently we ship whole applications plus associated data to OSG or NERSC in containers.
    - Can we instead deploy services at remote sites and connect them with streams?



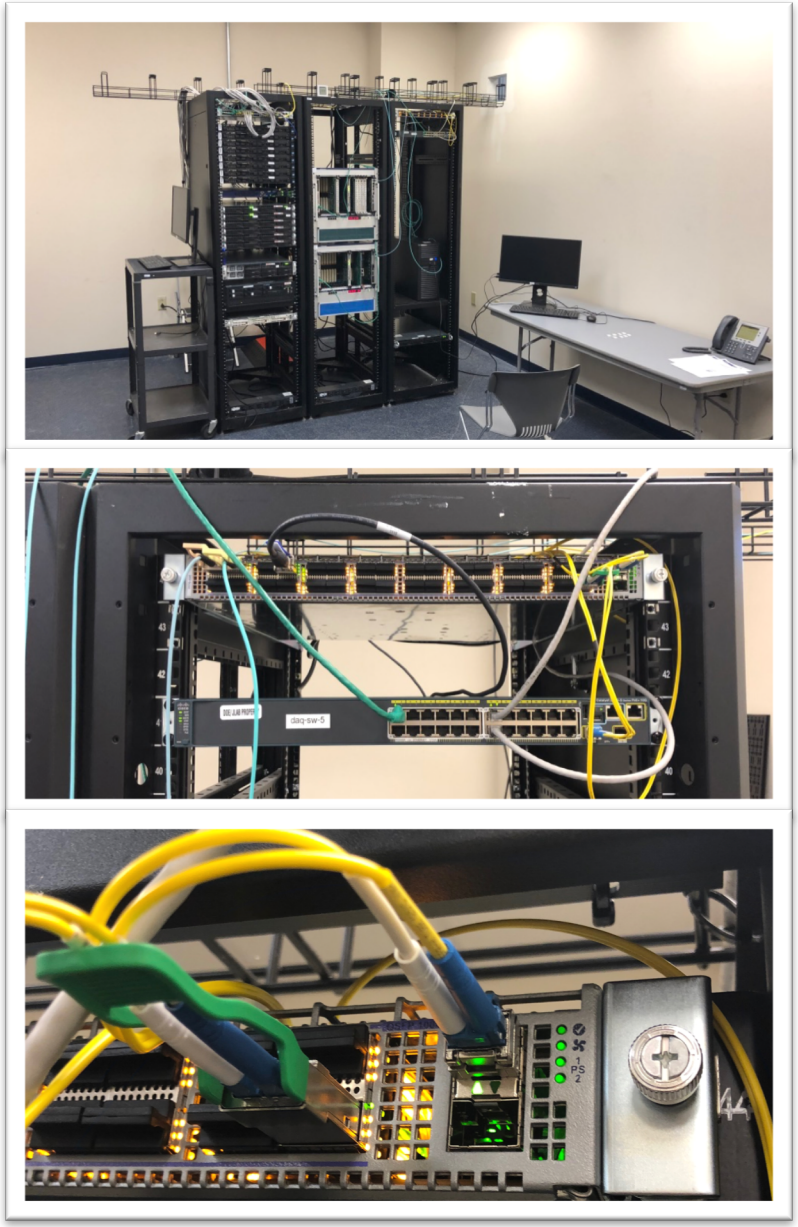
Jefferson Lab

Resources

- Facility for Innovation in Nuclear Data Readout and Analysis (INDRA).
- Located on the ground floor of F-wing, next to DAQ group lab.

- The INDRA facility is taking shape.
  - DAQ group server cluster.
  - "streaming capable" user programmable network switch, linked to the datacenter via a 100 Gb/s data link.
  - A fast PC with several full size PCI slots for testing high speed data links, GPU and FPGA boards.
  - A fast server machine with multi cores and ample memory - 100 Gb/s link to switch.
  - Two VXS crates for R&D with "legacy" boards.
  - Coming soon, fast server with SSDs to allow high rate data storage R&D.

- Open for business if people have projects!



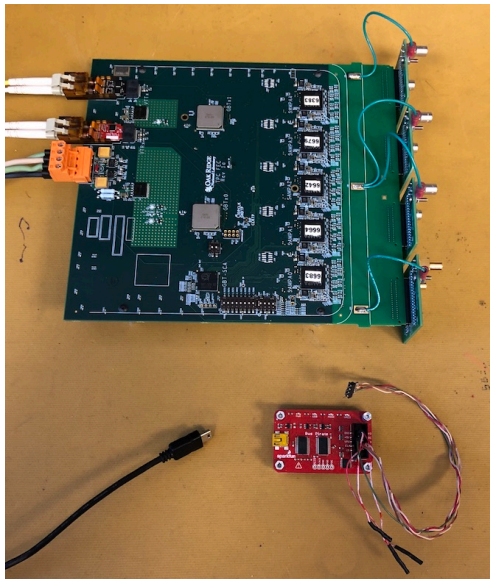
Jefferson Lab

FPGA and data handling R&D

- XILINX FPGA evaluation and test board.
  - Allows testing of data processing firmware on XILINX FPGA.
  - Can take fiber inputs compatible with our existing front end boards.
  - Same board being used by SLAC for testing firmware for HPS readout.
  - We have tested 5 Gbyte/s data transfers between board and host PC.
- EXAR DX2040 data compression board.
  - Compresses data streams at up to 12 Gbyte/s

DAQ projects : TPC readout prototype

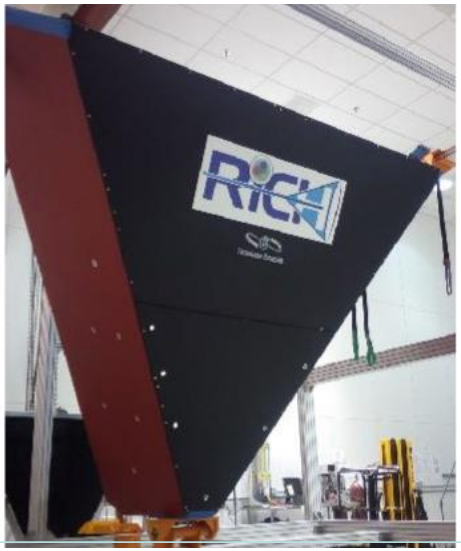
- Test of proposed readout for TDIS experiment and TPCs in general.
- Start out with an existing design from ALICE that has five SAMPA readout chips.
  - It was an effort to identify and procure all the parts as well as to find the right people to ask for help.
  - Now up and running and being tested in the DAQ lab.
- Firmware installed via USB using small adapter card.
- Data over fiber to Felix PCI card in a PC.
- Can see signals from the board but there is more noise than we would like to see.
- Talking to the board designers to come up with a solution.



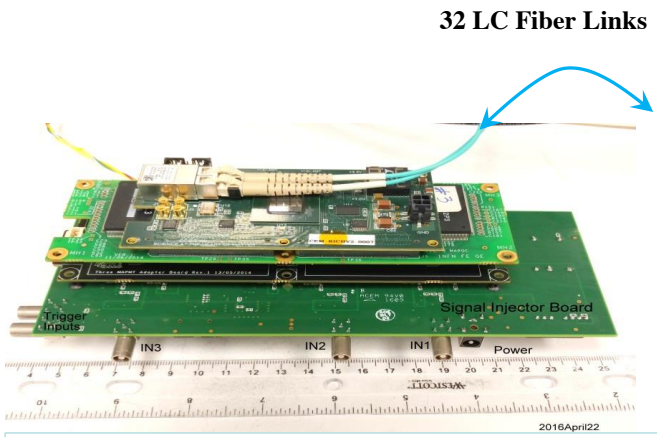
Jefferson Lab

DAQ projects : crateless and streaming DAQ

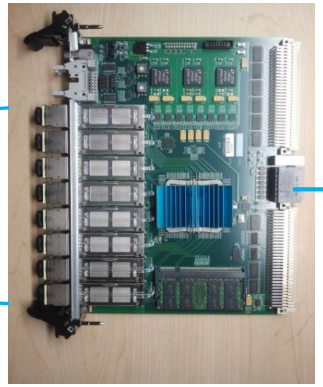
- CLAS12 RICH detector is instrumented with FPGA boards on the detector.
  - These are read out via fiber to Sub-System Processor (SSP) boards in VXS crates.
  - SSPs are read out over VXS serial backplane by a VTP.
  - VTP read out over VME - limits readout bandwidth.
  - Same setup is used by GlueX DIRC.
- Project :
  - Can we send the data out out using the fibers on the front panel of the VTP?
  - Can we modify the firmware on the three types of board to operate this system in a streaming mode?
  - Can we remove the SSP and VTP?
    - Run fiber links to a switch and process data on a generic FPGA board.



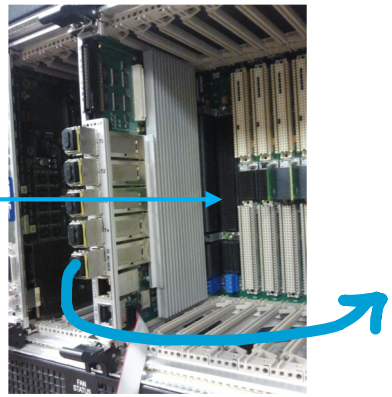
391 -- H12700 Hamamatsu
64-anode PMT
Total anodes: 25,024



On Board 192 channel FPGA Readout Board
MAROC3 ASIC mates to maPMT
Artix 7 FPGA drives LC fiber optic transceiver



VXS Sub-System Processor
32 - 2.5Gbps links to RICH
FPGA Readout Boards

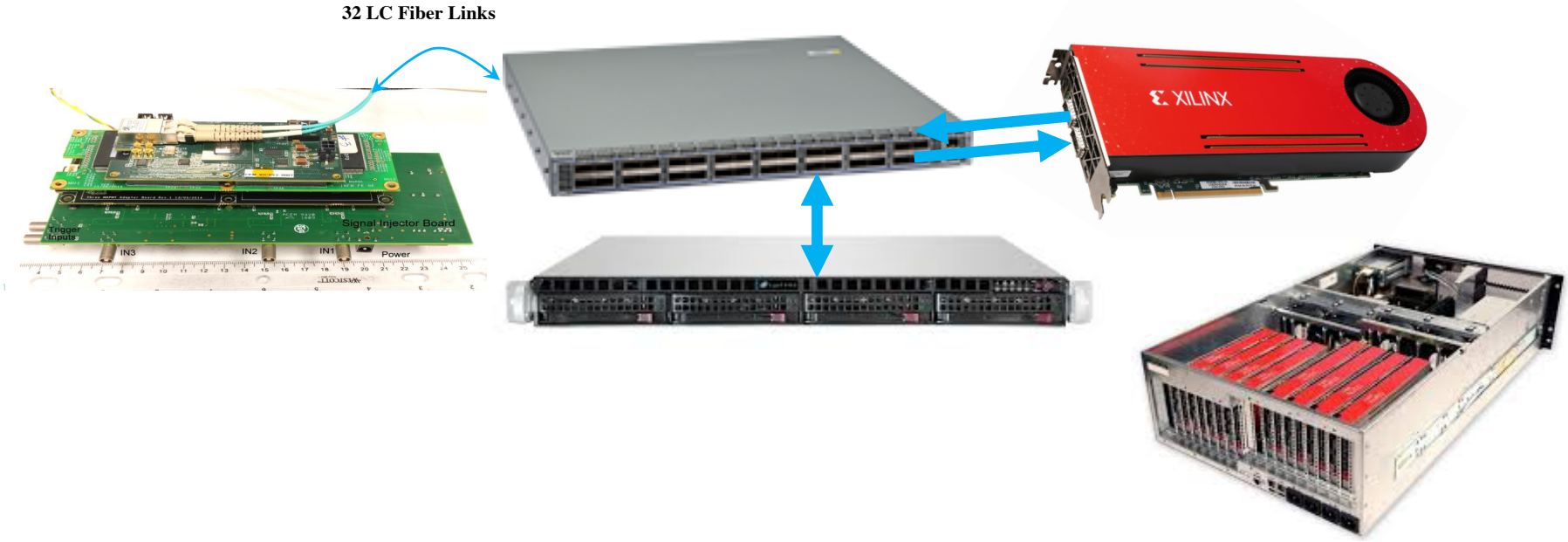


VTP processor in VXS
switch slot.
Output 40 Gbit/s fiber

Jefferson Lab

DAQ projects : Streaming through commercial hardware

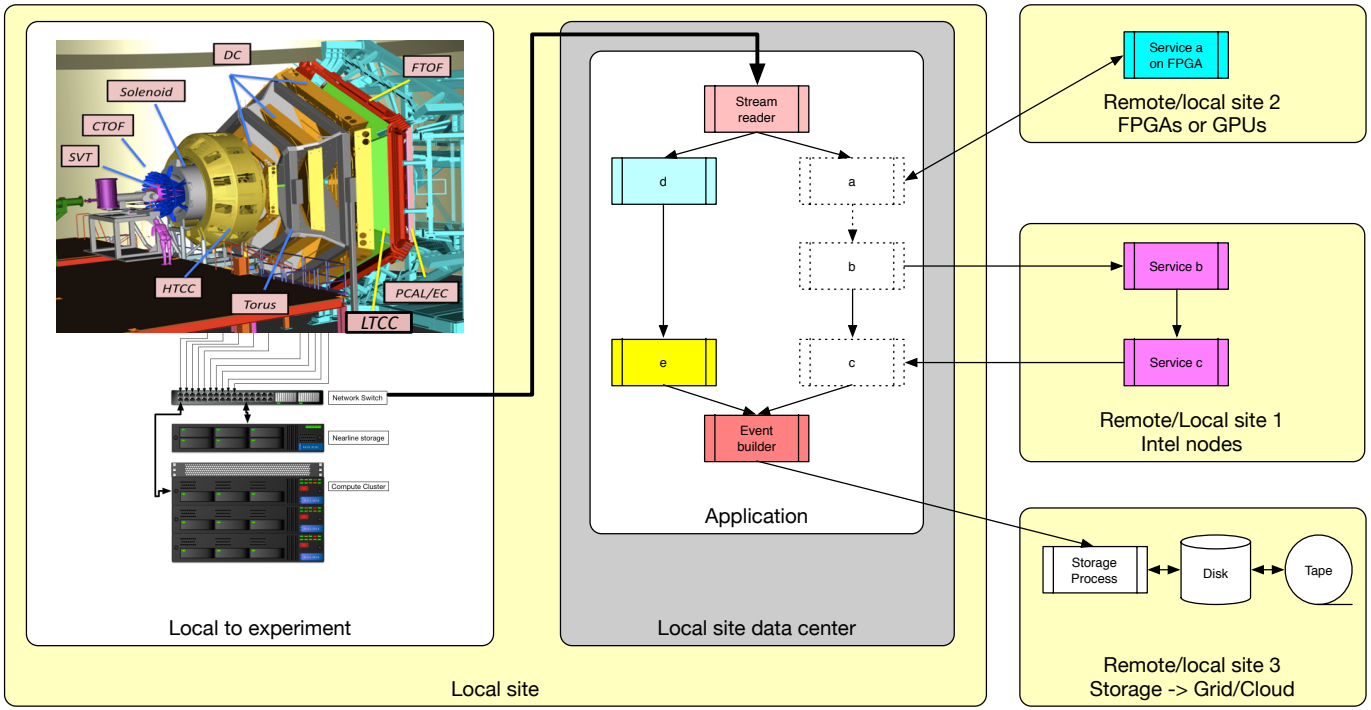
- Can we replace the majority of the streaming readout system with commercial hardware?
  - Route the data through a network switch instead of the SSPs and VTPs.
  - The SSPs and VTPs also run firmware to process the data from the front end cards. Replace this functionality with generic FPGAs in PCIe.



Jefferson Lab

The rest of the picture

- The previous slides cover most of the left side of the concept diagram and get the data as far as short term storage.
  - 2019-LDRD-10 covers what happens next – how to handle time ordered data streams from a streaming readout.
  - The JANA related LDRD, work on the next generation of CLARA, and on Machine Learning cover the remaining areas.
- Much work left to do.



Jefferson Lab

Summary

- The Streaming Grand Challenge is an amalgamation of various projects into a strategic initiative to develop a proof of concept advanced, integrated, readout and analysis for future experiments.

- The Grand Challenge is relatively new and ideas are evolving.

- We would like to invite anyone who is interested to participate either through working on projects or sharing ideas or concerns.