ISBN: 978-3-95450-200-4 doi:10.18429/JACoW-PCaPAC2018-FRCB1

DATA ACQUISITION SYSTEM FOR ELI BEAMLINES

P. Bastl†, P. Pivonka, B. Ploetzeneder, O. Janda, J. Trdlicka, V. Gaman, J. Sys,
ELI Beamlines/Institute of Physics of the ASCR, Prague, Czech Republic

Abstract

The ELI Beamlines facility is a Petawatt laser facility in the final construction and commissioning phase in Czech Republic. In fully operation phase, four lasers will be used to control beamlines in six experimental halls.

In this paper we describe the ultrafast and distributed data acquisition system of the facility. First, we discuss the requirements and demands on the system and introduce the environment. Then we discuss the structural divide in top- and low-level system. The unique idea of this system is the extensive fibre-based low-latency communication infrastructure, which distributes large RAM buffers and allows to offload computational tasks with high throughput and real-time capability.

INTRODUCTION

ELI Beamlines [1] is an emerging high-energy, high-repetition rate laser facility located in Prague, Czech Republic. Four laser sources will supply six experimental halls, which provide various secondary sources to users. Facility commissioning, and installation work of lasers and experiments is progressing, and first enabling experiments are in progress (2018).

The general architecture of the control / DAQ system is introduced in [2]. This paper will specifically discuss the implementation of the DAQ system.

ENVIRONMENT & INFRASTRUCTURE

The core idea of the facility (multiple laser sources distributed to multiple beamlines over a sophisticated beam distribution system) creates certain conditions for the data acquisition system that are similar, but not equal to classical accelerator facilities:

- The data sources produce bunches of data based on the repetition rate of the driving lasers. Most common are single-shot, 10Hz and 1kHz. There can be more than one laser per experiment, and they are not yet synchronized.
- Characteristic data sources are based on very short pulses (femtoseconds) – requiring high sampling rates. Some lasers are operating with high repetition rates and need to be imaged pulse-by-pulse with cameras and other 2D-datactors.
- The data sources are spatially distributed of 100s of meters and located in various, sometimes challenging or hazardous environments (ionized radiation, EMP, high-vacuum, laser, bio / chemical hazards, cleanrooms up to ISO5).
- Lasers and secondary sources are all developed by different stakeholders (ranging from inhouse research groups, industrial partners – from KMEs to major corporations, national labs, partnering academic labs.). This leads to a heterogeneous environment.
- There are hundreds of cameras, tens of digitizers, and tens of specialized detectors and instruments managed by different stakeholders. While there are efforts to standardize interfaces (for example, for cameras [3]), we still see a wide diversity of interfaces.

The DAQ system has the following tasks:

- Acquisition of data and standardization for further processing and storage.
- Buffering and provision of data for further control tasks (feedback loops etc)
- Data processing (online / offline)

The system also has to provide connections for data storage and integrate the electronic timing system. Facility-wide timestamping and synchronization is a challenge because the “heartbeat” of the facility are multiple, currently unsynchronized laser oscillators using different timing systems. Our solution for electronic timing is based on WhiteRabbit [4] and currently being implemented as described in [2].

STRUCTURAL OVERVIEW

The structure of control and data acquisition system is shown on Fig. 1. This figure identifies the basic room layout relevant to the system, namely main control room, dedicated server room, laser halls, experimental halls, and plant / infrastructure rooms (which are not integrated into the DAQ system, but may receive control feedback).

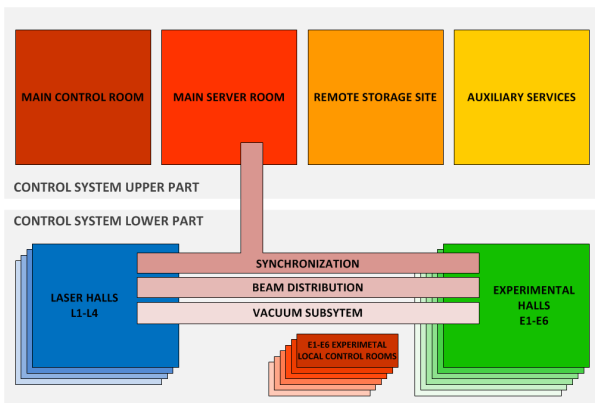

Figure 1: Structural overview.

LOGICAL OVERVIEW

The data acquisition system is divided into two levels, shown in Fig. 2 below:

- **Top level DAQ system**: This system, located in the server room, is responsible for the aggregation and buffering of large amounts of data. It also provides

ISBN: 978-3-95450-200-4 doi:10.18429/JACoW-PCaPAC2018-FRCB1

infrastructure for online / offline processing and interfaces for storage

- **Local level DAQ System**: These systems, located in the field (experimental halls / laser halls) provide an interface for the data sources and some capabilities for local processing / feedback loops. They also serve as a real-time interface for the more powerful computational infrastructure of the Top level DAQ system.

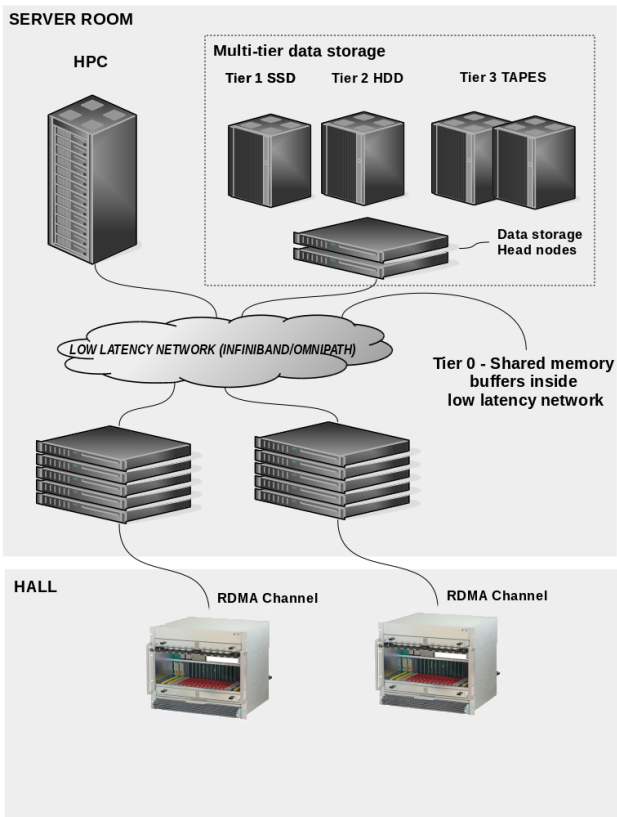

Figure 2: Logical overview of the DAQ system.

TOP LEVEL DAQ SYSTEM

The top level DAQ system is based on 3 data acquisition servers, implemented as a multinode IBM flex system. Each of its seven nodes has the following setup:

- 768GB of RAM (512 GB used for buffer)
- 4 PCIe interfaces for RDMA-enabled NIC cards and FPGA cards
- 24 Cores for further data manipulation and processing
- For each chassis, there is FDR Infiniband infrastructure (2 switches) and Ethernet infrastructure (2 switches, for control and user access).

These servers act as a RAM buffer and provide computational infrastructure.

First Stage: Acquisition and Buffering

Local (Field) data sources are connected to the server point-to-point primarily using the RDMA-enabled NIC cards (10Gb/s).

The data is stored in a pooled RAM buffer (called Tier 0 storage), which is shared using a low-latency network (at the moment: Infiniband 56GB/s, which will be updated to 100GB/s as needed). These low-latency networks cannot work over long distances, but the access via RDMA allows on-demand high-speed distribution of the expensive RAM resources independently of the physical location. This is important because the data load in the facility is highly unbalanced depending on which experiment is running with which lasers.

Second Stage: Acquisition and Processing

In the second stage, data is processed. The CPU cores can be used for this task (trivial standard solution), but we also provide FPGA-based acceleration for online and offline processing.

The field DAQ PCs can access them either using a NIC / RDMA (in this case, a NIC-IP core is implemented on the FPGA of the DAQ server), or using any arbitrary FPGA/FPGA connection. This allows real-time response, and sharing of computational resources (multiple potential inputs per FPGA on the DAQ server, using sx40GB/s interfaces).

Third Stage: Storage

The data then can be saved into data storage which consists of further tiers:

- Tier 0 is the distributed RAM buffer
- Tier1 is set of servers with NVMe in PCIe slots in headnodes of data storage, connected via 100Gb/s Infiniband interface to the DAQ szstem
- Tier 2 is based on conventional harddrives.
- Tier 3 is a tape library for long term data storage.

There are provisions to update this system both horizontally and vertically.

LOW LEVEL DATA ACQUISITION

In the field, we see three types of data sources that are connected to our top-level network:

- Some detectors come with their own inbuilt data aggregation and pre-processing systems
- **PCIe-based DAQ computers** (*Supermicro*) with 128GB of RAM, 24 cores and 10 PICex8 slots. These servers are low-cost, can provide large memory buffers (terrabytes). There is a wide variety of PCIe-cards for different applications.
- **MTCA-based DAQ systems**, which have the advantage of clock support in the backplane and allow card-to-card-connection without involving the CPU.

As already discussed in [2], we want to use the advantages of both Supermicro and MTCA systems (clock support, card-to-card-connections, large memory buffers). Our MTCA-MCH (NAT-MCH-PHYS80) allows to connect its internal PCIe switch through optical cable to PCIe based local DAQ server. The connection setup

ISBN: 978-3-95450-200-4 doi:10.18429/JACoW-PCaPAC2018-FRCB1

(shown in Fig. 3) provides PCIe x16 interface and can be implemented using PCIe on MTCAs' agnostic backplane and its fat pipes.

Figure 3: Connection between MTCA and PC based system.

OUTLOOK

The technology surrounding DAQ systems is evolving: Low latency (which we use for the shared RAM buffer) is approaching throughputs of 200Gb/s. FPGA technology is accelerating as well, and we see transceivers that are reaching speeds of 30Gb/s.

For such data rates, the PCIe gen.3.0 interface used in our servers will become a bottleneck. Therefore, we anticipate that we will progress to PCIe gen. 4.0, which is unfortunately not yet supported by Intel-based processors – but already supported by current FPGA chips and Open Power processors.

Coherent access to memory (CAPI, OpenCAPI) is also a topic we are actively monitoring and preparing. We have installed a POWER8 server with CAPI enabled interfaces, one FPGA with CAPI support and one Infiniband card for development and evaluation.

This setup seems to be promising as both CAPI and PCIe gen 4.0 are supported by the new Power9 processor generation.

REFERENCES

[1] ELI Beamlines, https://www.eli-beams.eu

[2] P. Bastl *et al*, "Hardware architecture of the ELI Beamlines Control and DAQ System", in *Proc. ICALEPCS'17*, Barcelona, Spain, Oct. 2017. doi: 10.18429/JACoW-ICALEPCS2017-FRAPL05

[3] B. Plötzeneder, V. Gaman, O. Janda, P. Pivonka, P. Bastl, "Cameras in ELI Beamlines: A standardized approach", in *Proc. ICALEPCS'2017*, Barcelona, Spain, Oct. 2017. doi: 10.18429/JACoW-ICALEPCS2017-THMPL06

[4] The White Rabbit Project, https://www.ohwr.org/projects/white-rabbit