DAQDB: a Key-Value Store for Data Acquisition Systems.

Danilo Cicalese [1], Grzegorz Jereczek [2], Fabrice Le Goff [1], Giovanna Lehmann Miotto [1], Jeremy Love [3], Maciej Maciejewski [2], Remigius K Mommsen [4], Jakub Radtke [2], Jakub Schmiegel [2], Aleksandra Jereczek[2], Adrian Pielech[2] and Malgorzata Szychowska [2]

[1]CERN, Geneva, Switzerland [2]Intel, Gdansk, Poland [3]Argonne National Laboratory, Argonne, USA [4]Fermilab, Batavia, USA

Background

Data acquisition systems for high energy physics experiments readout terabytes of data per second from a large number of electronic components. They are thus inherently distributed systems and require fast online data selection, otherwise requirements for permanent storage would be enormous. Still, incoming data need to be buffered while waiting for this selection to happen. Each minute of an experiment can produce hundreds of terabytes that cannot be lost before a selection decision is made. In this context, we introduce DAQDB (Data Acquisition Database) --- a distributed key-value store for high-bandwidth, generic data storage in event-driven systems. DAQDB offers not only high-capacity and low-latency buffer for fast data selection, but also opens a new approach in high-bandwidth data acquisition by decoupling the lifetime of the data analysis processes from the changing event rate due to the duty cycle of the data source. This is achievable by the option to extend its capacity even up to hundreds of petabytes to store hours of an experiment's data. Our initial performance evaluation shows that DAQDB is a promising alternative to generic database solutions for the high luminosity upgrades of the LHC at CERN.

Objective

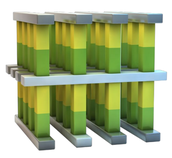
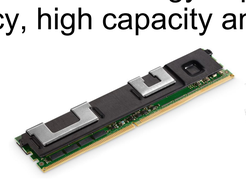
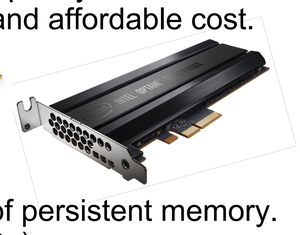
- Create a distributed, O(100) nodes, key-value store that allows organizing data fragments produced by the data acquisition systems of large experiments. Data are queried asynchronously by large software farms for analysis and selection based on physics signatures.

Performance requirements:

- Large, temporary storage of O(10) PB.
- Data insertion at ~ 5 TB/s; O(10k) data producers at 1 MHz each.
- Data retrieval at 3-5 TB/s; O(100k) independent data consumers
- Data storage for several hours.

Emerging Technologies

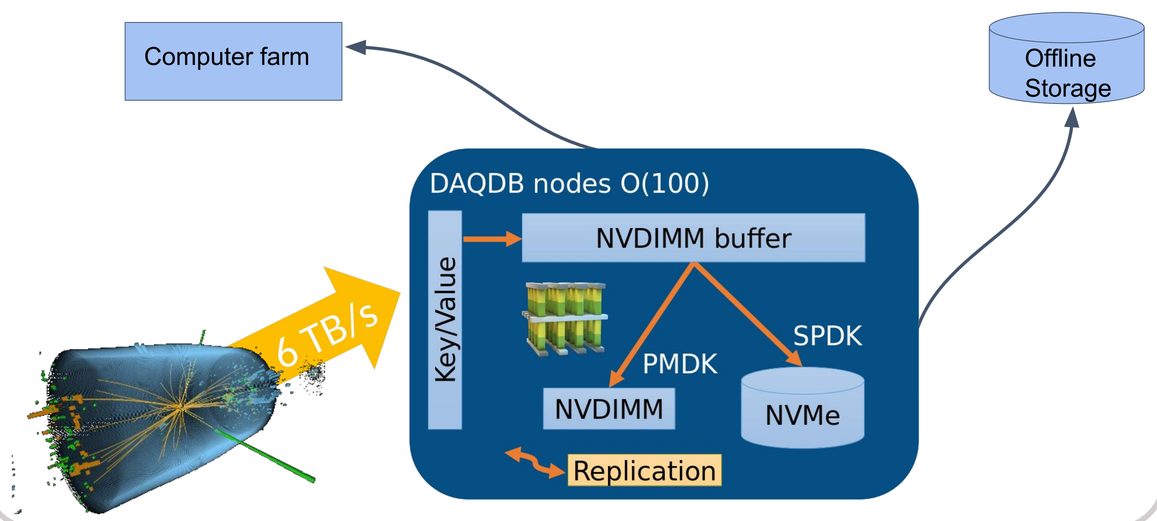
- Memory and Storage: **3D XPoint** a new technology explicitly architected for data center usage with low latency, high capacity and affordable cost.

- Software: **PMDK** [2] to achieve optimal performance of persistent memory. **SPDK**[3] to user-mode access to NVMe devices (SSDs).

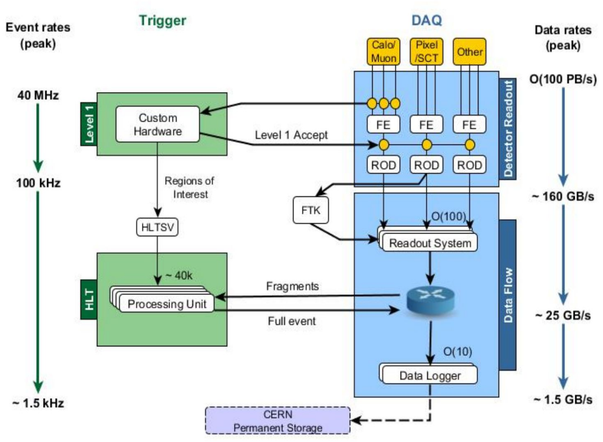
- Connectivity: **eRPC** [1] a general-purpose remote procedure call library that offers performance comparable to specialized systems.

DAQDB for data acquisition system

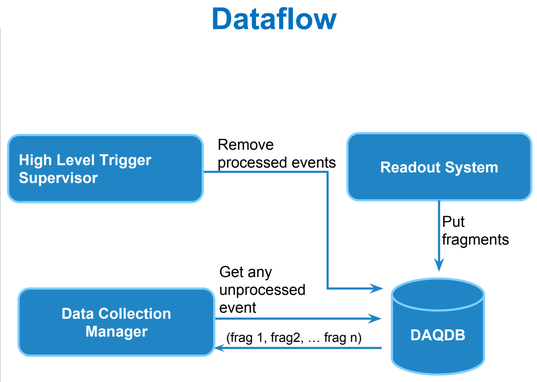


Integration

Atlas TDAQ



Dataflow

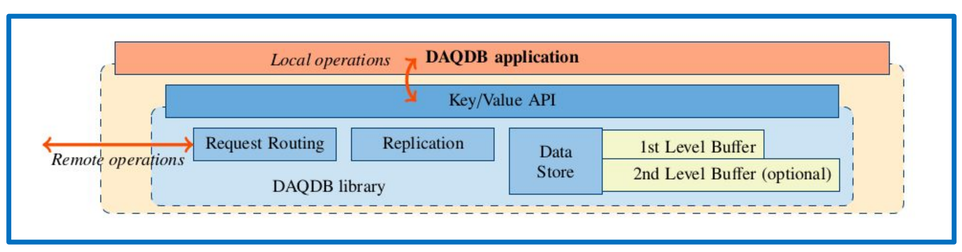


First year progress

✓Detailed requirements capture and understanding of the data flow

✓ Prototype DAQDB library implementation

✓ Initial standalone benchmarking



Outlook

The next steps are:

- Public repo on github
- System deployment on a small scale hardware system
- System testing and performance benchmarking
- Scale-out of the system and testing feasibility assessment

DAQDB @github (coming soon)

Talk @CHEP'18

Talk @CERN openlab TW '18

References

[1] Kalia, Anuj, Michael Kaminsky, and David G. Andersen. "Datacenter RPCs can be General and Fast." *arXiv preprint arXiv:1806.00680* (2018).
[2] Persistent Memory Development Kit. http://pmem.io
[3] Storage Performance Development Kit. http://spdk.io

EP-DT Detector Technologies

https://openlab.cern/

Contact information : danilo.cicalese@cern.ch, giovanna.lehmann@cern.ch