Experience with DAQDB
Key-value store for the ATLAS DAQ system

Adam Abed Abud [12], *Danilo Cicalese* [2], *Grzegorz Jereczek* [3], *Fabrice Le Goff* [2], *Giovanna Lehmann Miotto* [2], *Jeremy Love* [4], *Maciej Maciejewski* [3], *Remigius K Mommsen* [5], *Jakub Radtke* [3], *Jakub Schmiegel* [3], *Malgorzata Szychowska* [3]

[1] University of Liverpool, Liverpool, UK [2] CERN, Geneva, Switzerland [3] Intel, Gdansk, Poland
[4] Argonne National Laboratory, Argonne, USA [5] Fermilab, Batavia, USA

Motivation and objectives

The increase in data rates expected at the HL-LHC will create challenging conditions for the experiments. For the ATLAS experiment, it is anticipated that the readout system will sustain a throughput of 5 TB/s and will be implemented with approximately 500 servers. A storage system of O(10) PB will act as temporary buffer between the readout and filtering systems. Implementing such a large storage system with DRAM is not economically feasible. On the other hand, current SSD devices do not have sufficient throughput performance and endurance. One possible solution is to adopt a **logical event building** approach where data fragments are stored in a distributed storage buffer based on a key-value store (KVS). Event building will be taken care internally by the KVS by only issuing metadata operations instead of physically moving the data fragments.

Data Acquisition Database (**DAQDB**) is an open-source **distributed key-value store** for high-bandwidth data storage architectures. The goal is to have a system capable of providing a O(10) PB buffer and able to sustain an input and output rate of O(5) TB/s at 1 MHz. DAQDB uses the **Intel 3D XPoint memory and storage technology** to achieve the target goals of high-throughput, low-latency and high endurance needed by the LHC experiments (ATLAS, CMS).

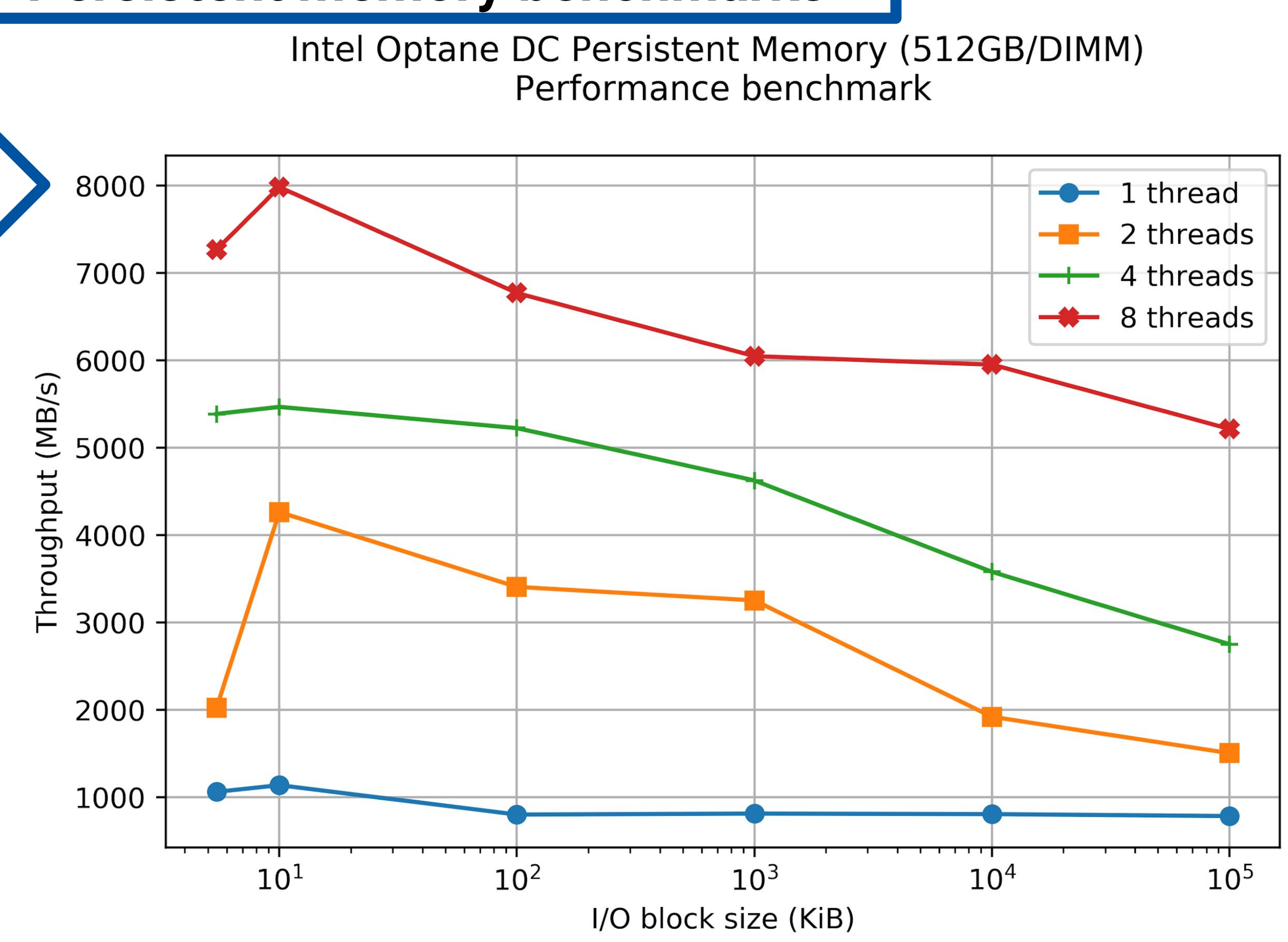
Persistent memory evaluation

The first step before integrating DAQDB into the ATLAS TDAQ framework is to benchmark the performance of the underlying memory technology. This gives insights into the maximum performance that can be achieved by the system. For this reason benchmarks of the 3D XPoint memory devices (**persistent memory**) were executed. The benchmarks show that a single device can achieve a throughput of O(10) GB/s with a limited number of writing threads. In the case of 8 threads, the maximum performance of 8 GB/s is achieved for a block size of 10 KiB.

Testing platform

- C628 chipset + Cascade Lake XeonⓇ Platinum L SKU processor
- 192GB RAM + 6TB IntelⓇ Optane™ DC NVDIMMs

Persistent memory benchmarks



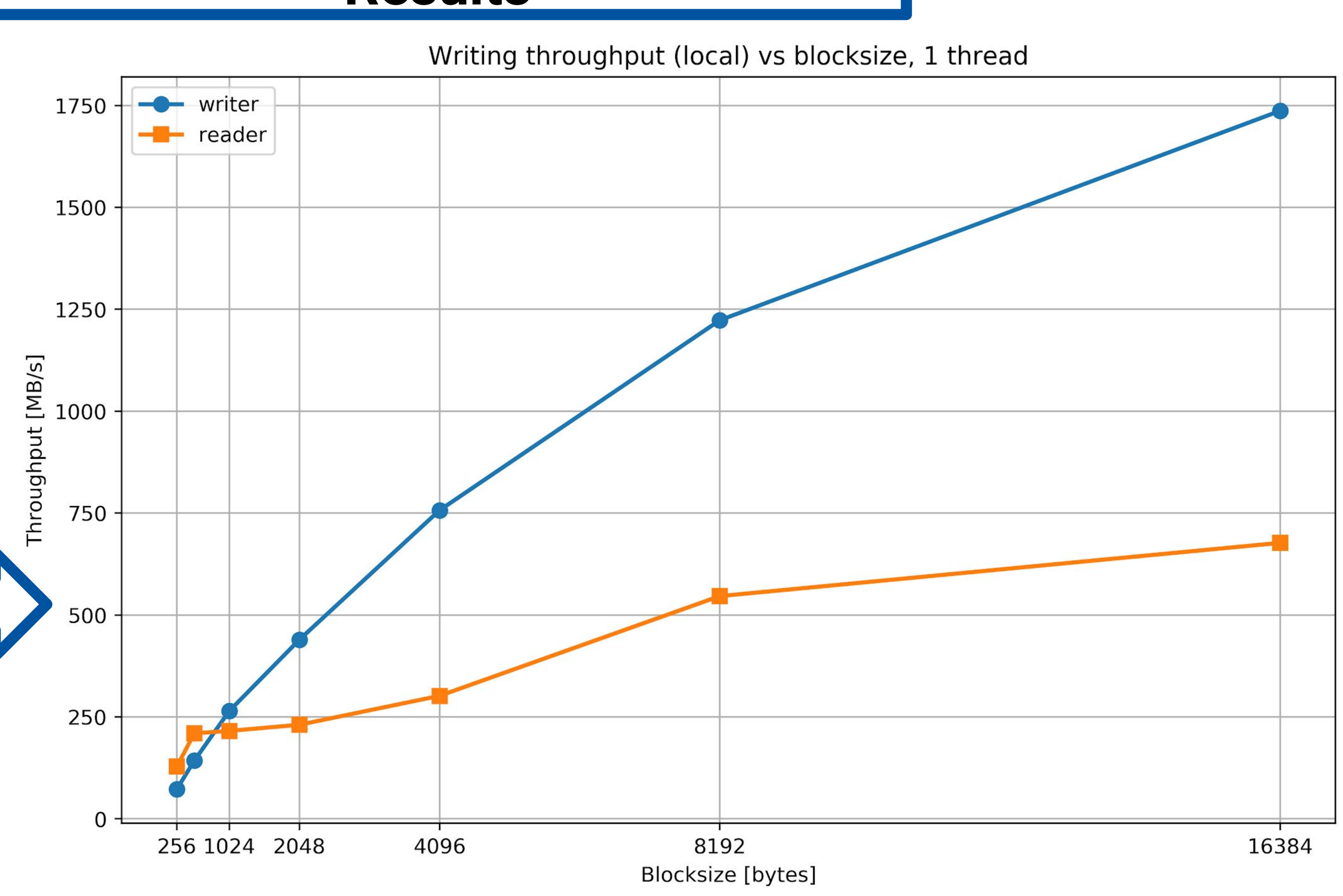
Integration of DAQDB

In order to **integrate** DAQDB into the **ATLAS TDAQ** framework, **standalone producer and consumer applications** have been created. These are used to emulate the traffic patterns of the readout and filtering units.

Preliminary results of the throughput were collected for both writing (readout) and reading (filtering) operations by one single-threaded application.

For a block size of 10 KiB, comparable to the anticipated HL-LHC ATLAS fragment size, the writing throughput is greater than 1.2 GB/s. This is a promising result as with multiple nodes and with less than 10 threads it could be possible, within the maximum rate available in the persistent memories, to reach the target throughputs needed for the ATLAS upgrade. The results for the reader show a low performance (O(500) MB/s). Investigations on how to exploit the capabilities of the hardware are currently being explored.

Results



Achievements and outlook

Benchmarks of persistent memory technologies were carried out. DAQDB has been integrated with the ATLAS TDAQ software framework and first results have been obtained.

Future directions of this work consist in making large scale tests with multiple nodes and testing the system in multiple configurations.

References

Github

Intel Optane Technology: http://cern.ch/go/kv9R

M. Maciejewski, **Persistent Memory based Key-Value Store for Data Acquisition Systems**, IXPUG2019, https://cds.cern.ch/record/2692497

G. Jereczek, **Let's get our hands dirty: a comprehensive evaluation of DAQDB**, key-value store for petascale hot storage., CHEP2019, https://indi.to/wYB2j

J. Radtke, **A Key-Value Store for Data Acquisition Systems**, SPDK, PMDK and VTune(tm) Summit 04'19 Santa Clara

adam.abed.abud@cern.ch