The 2023 Fermilab GPU Cluster “LQ2”

Tim Skirvin (HPC architect),
David Fagan (hardware architect),
Smita Darmora (LQCD site co-manager),
Ken Herner (LQCD site co-manager), and
James Simone (HPC coordinator)

April 2023

FERMILAB-SLIDES-23-031-CSAID

Background

Fermilab will provide GPU computing to USQCD beginning October 2023.

New cadence: buy this year is at FNAL, the first “tick”.
Next year’s buy at BNL will be the “tock” . . . and repeat.

Continue to maintain strong LQCD hardware programs at both BNL and FNAL.

Build a bigger facility each year by sending equipment funds to a single lab.

Ease the procurement effort for both BNL and FNAL.

Cluster specification

Survey of existing USQCD allocated resources highlighted the lack of recent-generation data-center GPU resources.

Panel of domain experts wrote hardware recommendations for a mult-use GPU cluster optimized for LQCD but capable of running representative workloads from CMS, AI, and DUNE.

Panel recommended procurement be open to both AMD and NVIDIA GPUs.

For the first time, an LQCD multi-GPU benchmark would be needed to rate price-to-performance of proposed systems, *i.e.*, AMD *vs* NVIDIA.

The RFP asks for sufficient workers to deliver aggregate sustained four-GPU Dslash performance of at least 40 TFlop/s (base) plus an additional 100 TFlops/s in purchase options.

Challenges

Redirection of extra EQ funds to Fermilab was after RFP was released, however, RFP options were sufficient.

Supply chain disruptions impacted timely availability of systems and components.

Inflation and economic conditions in the tech industry increased risks and costs for system builders to respond to the RFP.

System builders had difficulty getting access to systems they hoped to bid to run the required LQCD benchmark.

Smaller integrators, in particular, were hesitant to risk the time and effort needed to download and run an unfamiliar benchmark.

Fermilab extended the RFP reponse deadline by six weeks.

Timeline

13 Jan 22 – Hardware requirements committee kickoff meeting.

15 Apr 22 – Hardware recommendations submitted to USQCD and Fermilab.

22 Apr 22 – Fermilab CIO creates charge to procurement committee.

08 Aug 22 – RFP approved and purchase requisition created.

05 Oct 22 – RFP posted to vendors with 25 Oct deadline.

16 Dec 22 – The extended deadline for vendor responses.

09 Jan 23 – Winning bid recommendation sent to procurement.

28 Mar 23 – Purchase officially awarded to vendor (Koi).

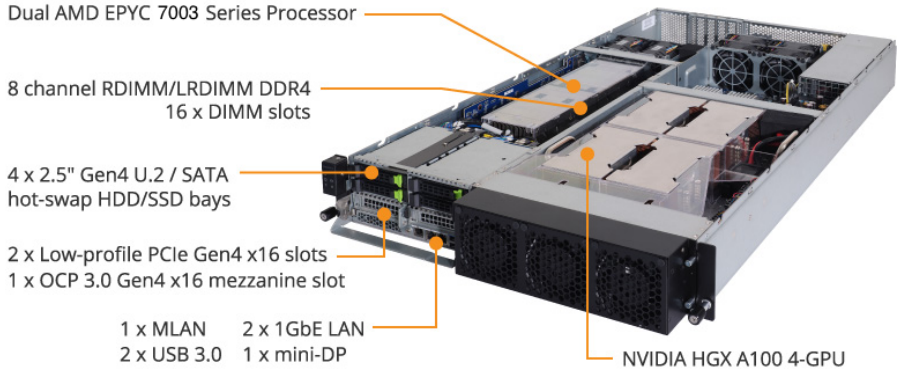
31 Jul 23 – First 13 servers are expected to ship to vendor.

07 Aug 23 – Begin install, burn-in and friendly user testing.

01 Oct 23 – LQ2 cluster becomes an allocated USQCD resource.

GPU worker node

Servers are Gigabyte model G262-ZR0 / HGX 4x A100 ”Redstone”



GPU worker node features

Each of the 18 worker nodes will have

Quad NVIDIA A100 GPUs each with 80 GB of HBM2e device memory

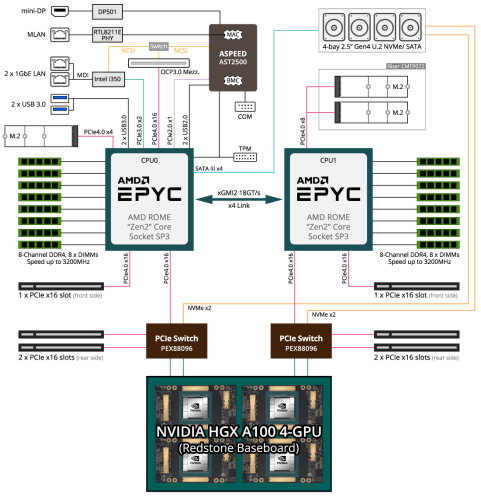
Dual 32-core AMD EPYC 7003 Zen 3 (Milan) 7543 2.8 GHz processors

Dual 1-port ConnectX-7 NDR/200 IB cards, 200 Gbps (400 Gbps per worker)

1 TB RAM total - 2x 8-channel 64 GB DDR4 3200 MHz

960 GB SSD scratch disk

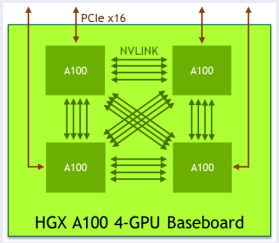
Servers designed for GPU-accelerated HPC



HGX A100 “Redstone” baseboard

four point-to-point NVlink channels between every pair of GPUs

GPU peer-to-peer bandwidth is 200 GB/s (bi-directional)



LQ2 HPC network

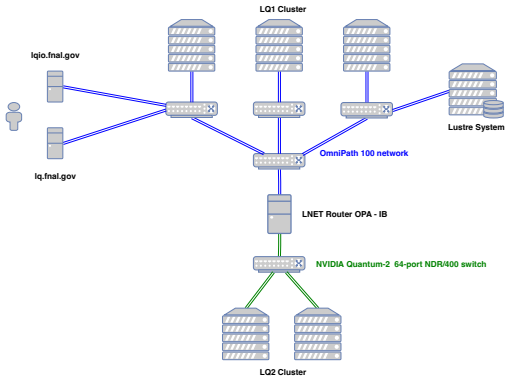
NVIDIA Quantum-2 QM9700 Infiniband switch

Switch supports 64 400 Gb/s ports *or* 128 200 Gb/s ports

Spare ports for potential expansion.

Each LQ2 worker will have dual 1-port 200 Gb/s network interfaces

Bridging the OmniPath and Infiniband networks



Existing LQ1 HPC network is OmniPath 100 Gbps

LQ2 network is faster NDR/400 Infiniband for GPU b/w requirements.

Access to existing LQ Lustre requires a gateway between OmniPath and IB.

Will build an LNET router using an existing server.

What to expect

Fermilab will schedule one or more downtimes to upgrade the LQ complex to EL8 (AlmaLinux 8.x) and to do any reconfiguration needed to integrate the LQ2 cluster with the existing LQ1 services.

Core HPC software frameworks will be upgraraded to OpenHPC verion 2.x

We will strive to open LQ2 to friendly user testing ASAP. Preference will be given to groups that have requested A100 allocations in the 2023-2024 period.

Allocations will begin October 01, 2023.