OAK RIDGE National Laboratory

Frontier

2025-07-28 - Argonne Training Program on Extreme-Scale Computing (ATPESC)

John K. Holmen

HPC Engineer, System Acceptance & User Environment

Oak Ridge Leadership Computing Facility

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

U.S. DEPARTMENT OF ENERGY

What is a Leadership Computing Facility (LCF)?

- Partner with users to enable science and engineering breakthroughs
- Mission: Provide capability computing resources for the most difficult problems
- Resources are unique scientific instruments
  - Closer in use, intent, purpose, and scale to Large Hadron Collider and James Webb telescope than your laptop
- Allow users to investigate otherwise inaccessible systems across scales
  - Galaxy Formation to Nanomaterials

https://www.flickr.com/photos/olcf/52117623798

https://www.chicagomag.com/wp-content/uploads/2023/01/C202302-Aurora-Supercomputer-nodes.jpg

Who Uses Leadership Computing Facilities?

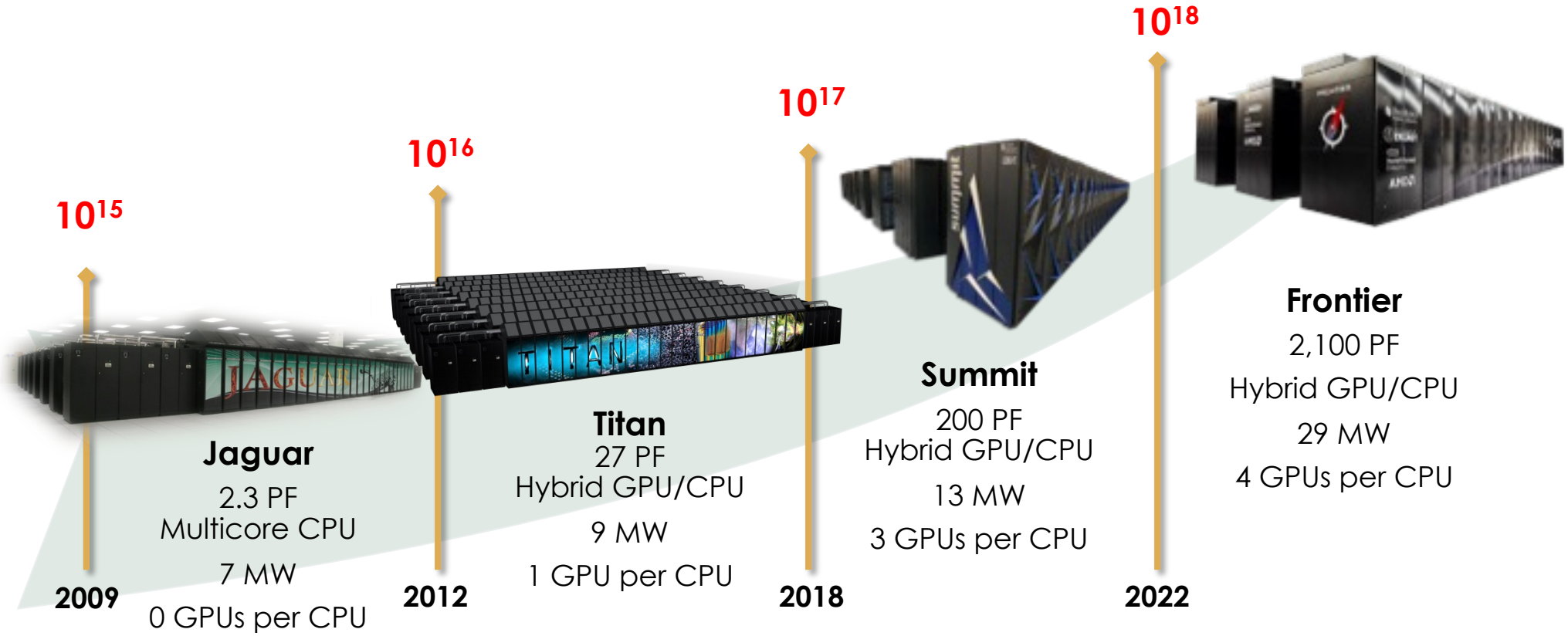
- Users from across the world
  - Academia, industry, national laboratories
- Time awarded through allocation programs
  - INCITE (Large)
    - Innovative and Novel Computational Impact on Theory and Experiment Program
  - ALCC (Small-Medium)
    - ASCR Leadership Computing Challenge
  - DD (Small)
    - Director's Discretionary
  - OLCF Pathways to Supercomputing Initiative (Startup)
    - Time and staff support to prepare for larger allocations

Oak Ridge Leadership Computing Facility (OLCF)

- One of two Department of Energy LCFs
- Based in Oak Ridge, TN at the Oak Ridge National Laboratory (ORNL)
- Department of Energy-funded research
  - Neutron Science, High-Performance Computing, Advanced Materials, Biology and Environmental Science, Nuclear Science and Engineering, Isotopes, and National Security
- Largest, most modern center for unclassified computing in the US

https://www.ornl.gov/sites/default/files/styles/basic_page_hero/public/2008-P01679.jpg

From Petascale to Exascale



A Note on Energy Efficiency

- 145x reduction in energy per FLOPS from Jaguar to Frontier
  - 3,043 MW/EF in 2009 -> 21 MW/EF in 2022
- 1 Frontier cabinet has 10% more performance than all of Titan
  - Using 309 kW compared to Titan's 9 MW

1 Frontier Cabinet (2022)

200 Titan Cabinets (2012)

Frontier System Overview

- HPE Cray EX Supercomputer architecture
  - 9,856 compute nodes across 77 cabinets
  - 2.1 EF peak double-precision performance
    - 1.102 EF HPL performance (June 2022 debut)
    - 1.353 EF HPL performance(June 2025)
  - AMD 64-core Optimized 3rd Gen EPYC CPUs
  - AMD Instinct MI250X GPUs
  - HPE Slingshot interconnect
  - Cray and AMD ROCm prog. environments
  - 679 PB Lustre filesystem, "Orion"
  - NFS storage (/ccs/home, etc.)

https://www.flickr.com/photos/olcf/53567220071/

HPE Cray EX4000 Cabinet

- 8x compute chassis per cabinet
  - Up to 64x compute blades
  - Up to 64x HPE Slingshot switch blades
- Direct liquid cooling using a shared Coolant Distribution Unit (CDU)
- 4x power shelves with a max of 8x rectifiers each
- Supports 400 kW

https://www.flickr.com/photos/olcf/52117623763/in/album-72177

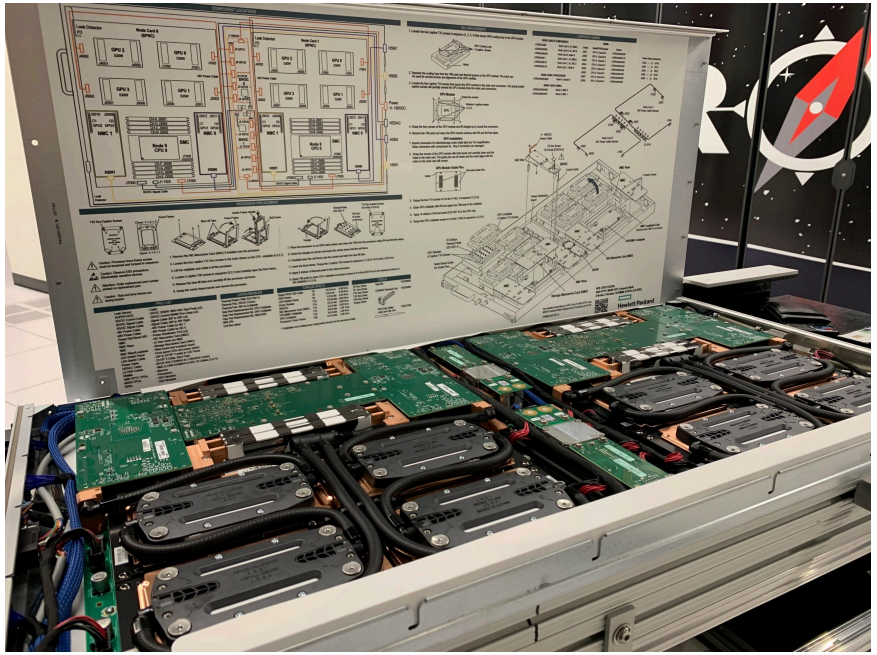
HPE Cray EX4000 Cabinet

- 8x compute chassis per cabinet
  - 8x compute blades per chassis
  - 4x switch blades per chassis
- Frontier cabinets consist of:
  - 64x compute blades
  - 128x compute nodes
  - 512x network interface cards (NIC)
  - 32x HPE slingshot switches
  - 1x Dragonfly group
- 8,000 pounds per cabinet

https://www.flickr.com/photos/olcf/52117623798/in/album-72177720299483343

HPE Cray EX235A Blade

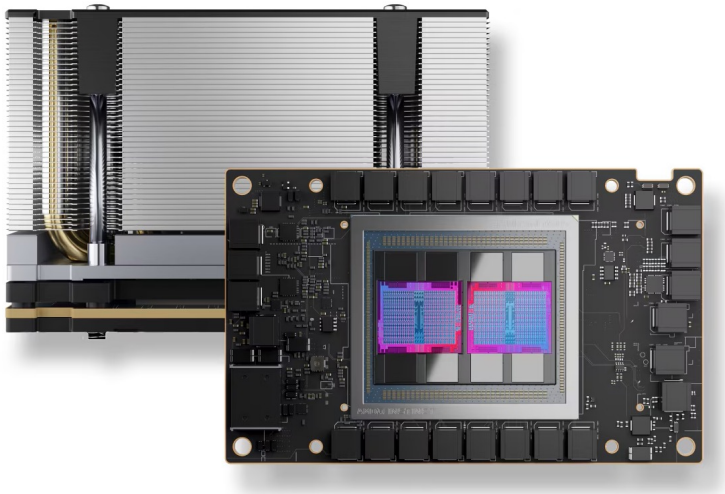
- 2x compute nodes per blade
  - Direct liquid cooled for all components
- Each node has:
  - 1x AMD Optimized 3rd Gen EPYC CPU
  - 4x AMD Instinct MI250X GPUs
  - 512 GB DDR4 on CPU
  - 512 GB HBM2e per node
  - 2x 1.92 TB NVMe, "Burst Buffer"
  - Full CPU & GPU connectivity with AMD Infinity Fabric
  - 4x HPE Slingshot 200 GbE NICs



https://www.hpcwire.com/wp-content/uploads/2022/06/ORNL-Frontier-blade-display-closer-June2022_4000x-scaled.jpg

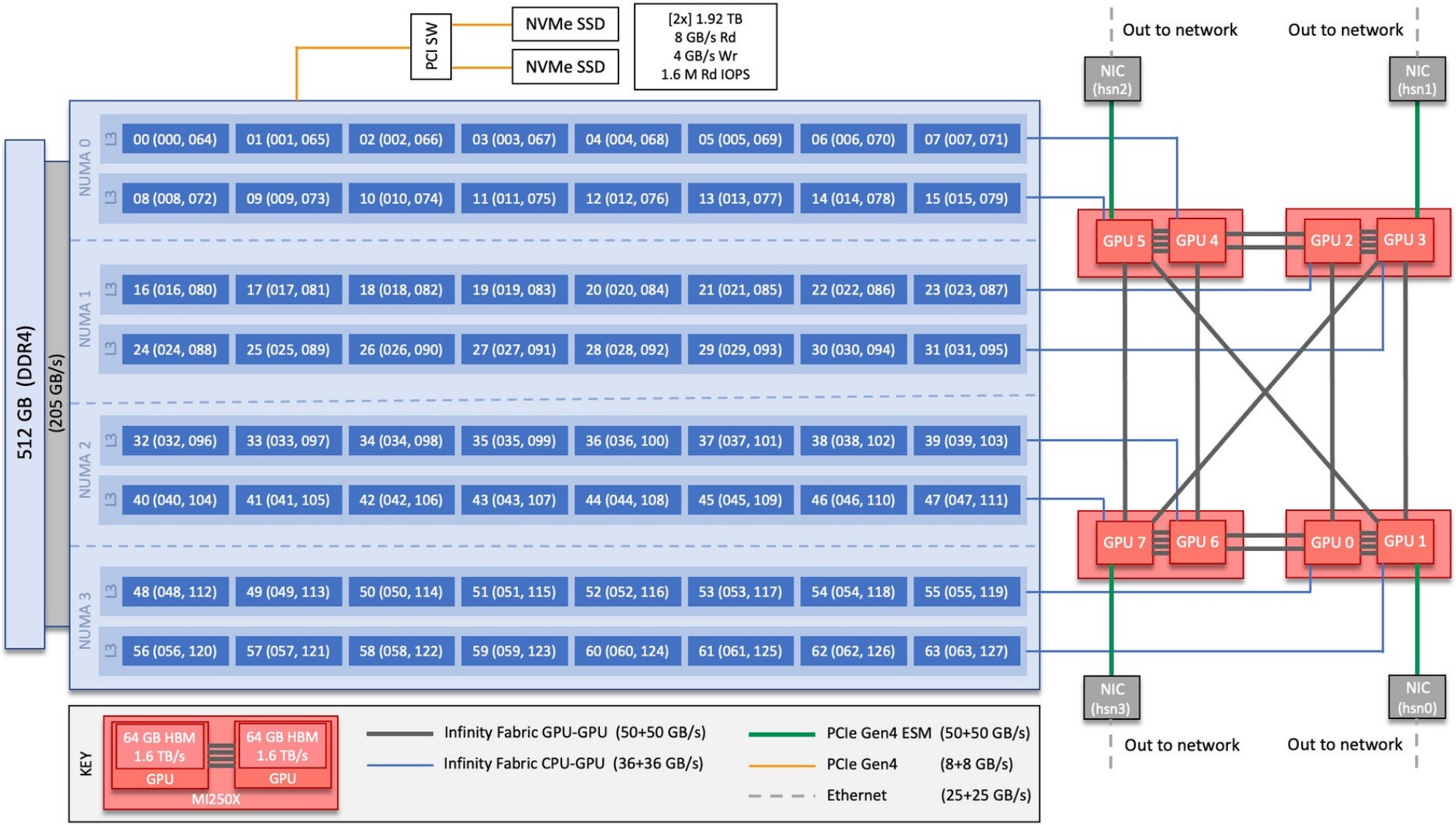
AMD Instinct MI250X GPU

- 2 Graphics Compute Dies (GCDs) per GPU
  - Shown by OS as 2 GPUs
- Effectively 8 GPUs per node, each with:
  - 110 Compute Units
  - 26.5 TFLOPs double-precision peak
  - 64 GB of HBM
  - 1.6 TB/s Memory Bandwidth
- Each associated with a CPU L3 cache region
- 1 NIC connected to each MI250X

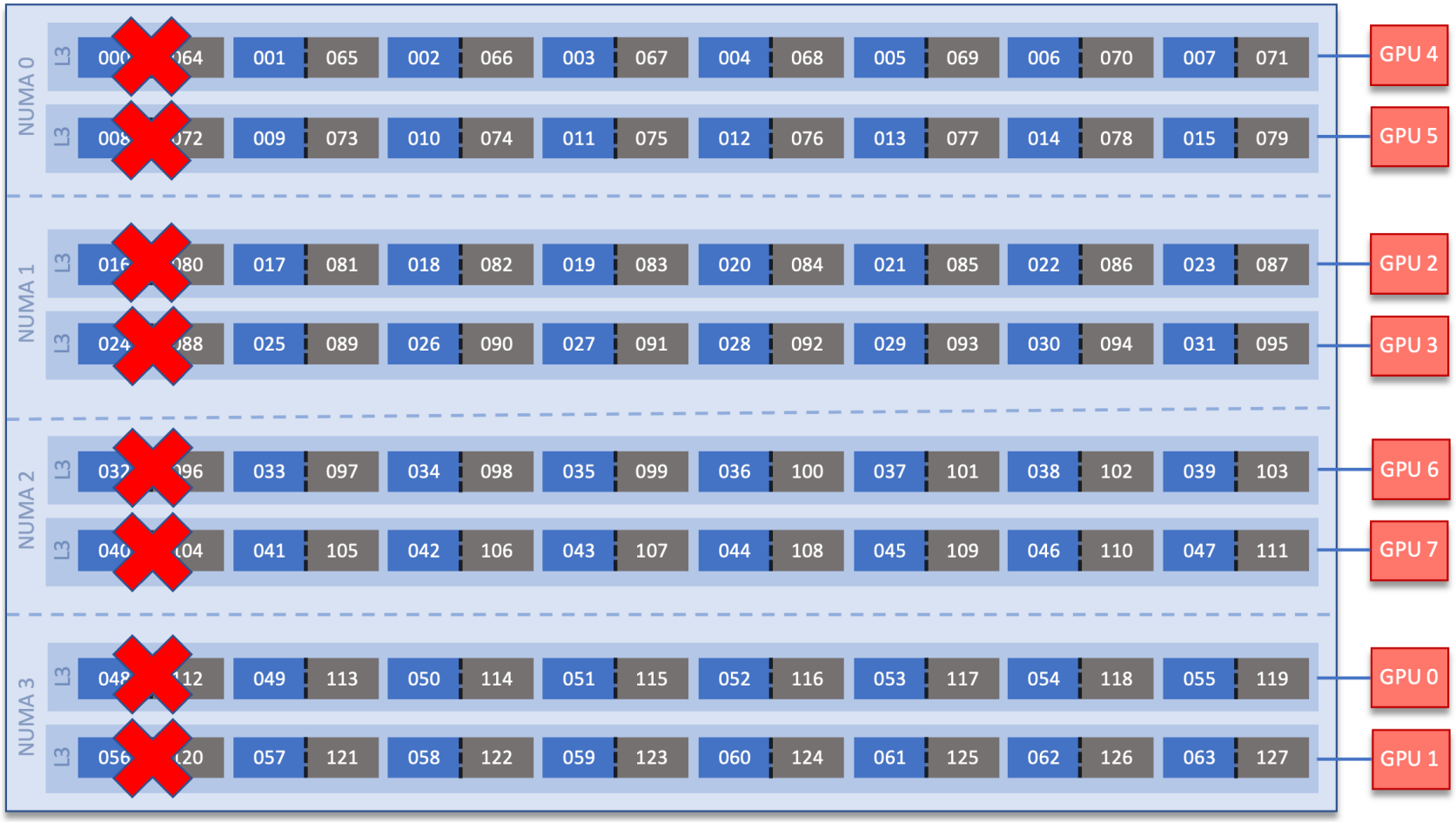


https://www.amd.com/content/dam/amd/en/images/products/data-centers/2325906-amd-instinct-mi250x-product.jpg

Frontier Node Diagram

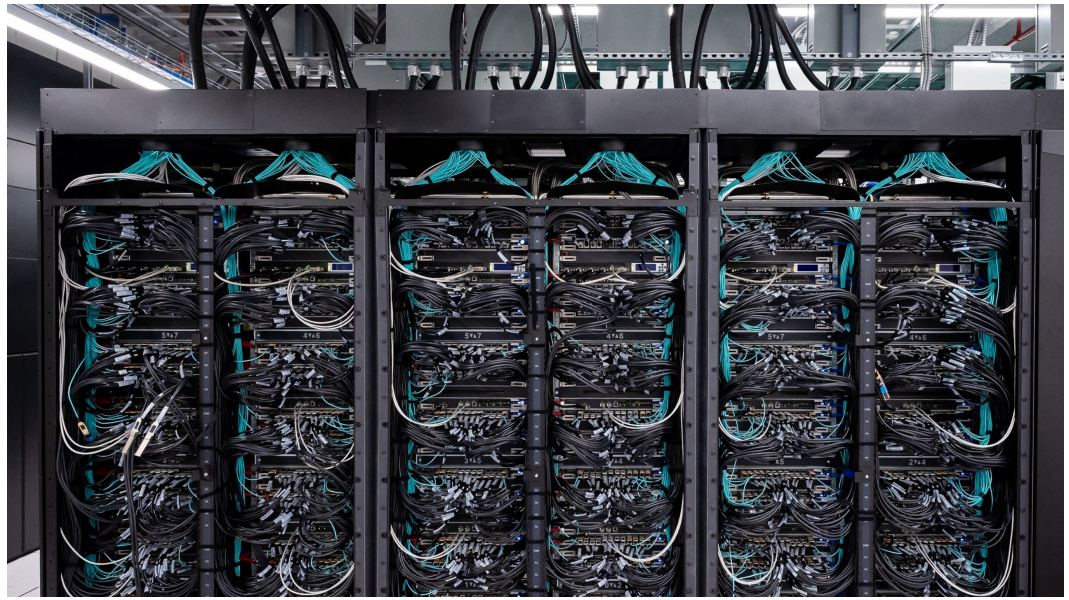


Frontier Node Diagram – Default Configuration



HPE Slingshot Interconnect

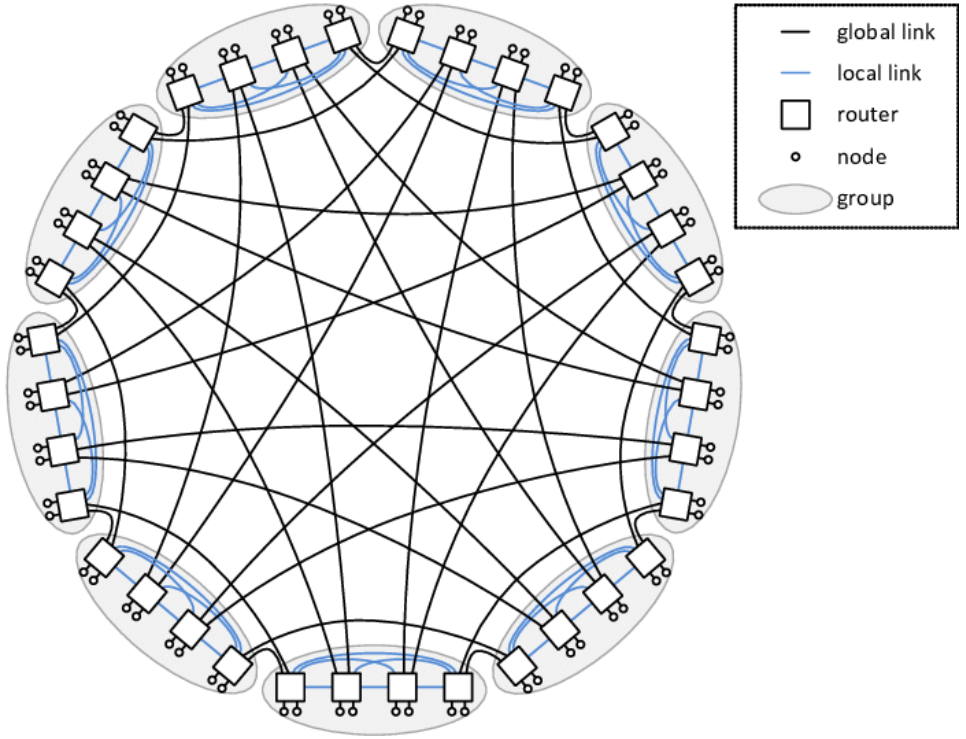
- High-speed, low latency network architecture
- HPE Slingshot switches (64 ports)
  - 25 GB/s bi-directional BW per port
- HPE Slingshot NICs
  - 25 GB/s bi-directional BW per link
- Slingshot is a superset of Ethernet with optimized HPC functionality
- Frontier uses dragonfly topology



https://www.ornl.gov/sites/default/files/2022-05/Side%20view%20Frontier%20cabinets.jpg

Frontier Network Topology

- Dragonfly groups
  - A group of endpoints connected to switches that are connected all-to-all
- Dragonfly topology
  - A set of groups connected all-to-all
    - Each group has >=1 link to every other group
- Frontier has 77 compute groups
  - 128 nodes per compute group
  - 32 switches per computer group
  - 4 NICs per node



https://www.researchgate.net/profile/Enrique-Vallejo-2/publication/261313973/figure/fig2/AS:667782257573894@1536223105142/Sample-Dragonfly-topology-with-h2-p2-a4-36-routers-and-72-compute-nodes.png

Frontier Programming Environment

- Compilers
  - Cray CCE
    - C/C++ LLVM-based
    - Cray Fortran
  - AMD ROCm
    - C/C++ LLVM-based
  - GCC
  - oneAPI DPC++
    - LLVM-based
    - user-managed

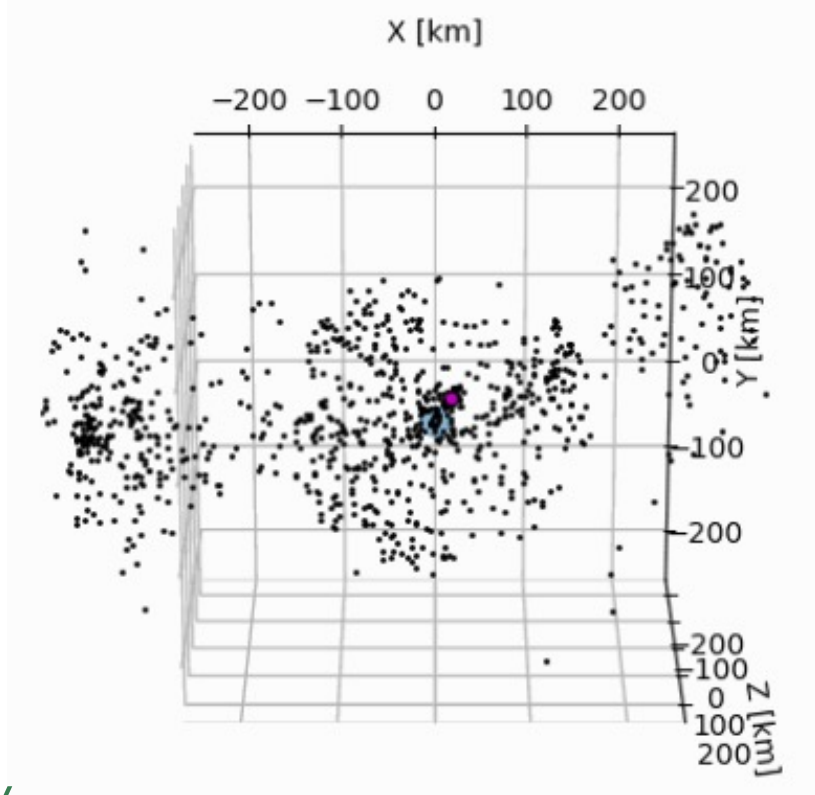
- Programming Models & Abstraction Layers
  - OpenMP
  - HIP
  - Kokkos
  - RAJA
  - SYCL
    - via user-managed DPC++
  - OpenACC
    - C/C++ via user-managed clacc
  - OpenCL
  - Julia
  - UPC++

OLCF System Access

- Odo available through TRN038
  - Odo is a 30-node Frontier training system
- System user guides:
  - https://docs.olcf.ornl.gov/systems/odo_user_guide.html
  - https://docs.olcf.ornl.gov/systems/frontier_user_guide.html
- Details on applying for an allocation of your own:
  - https://docs.olcf.ornl.gov/accounts/index.html
- Startup allocations with support also available:
  - https://www.olcf.ornl.gov/community/pathways-to-supercomputing/

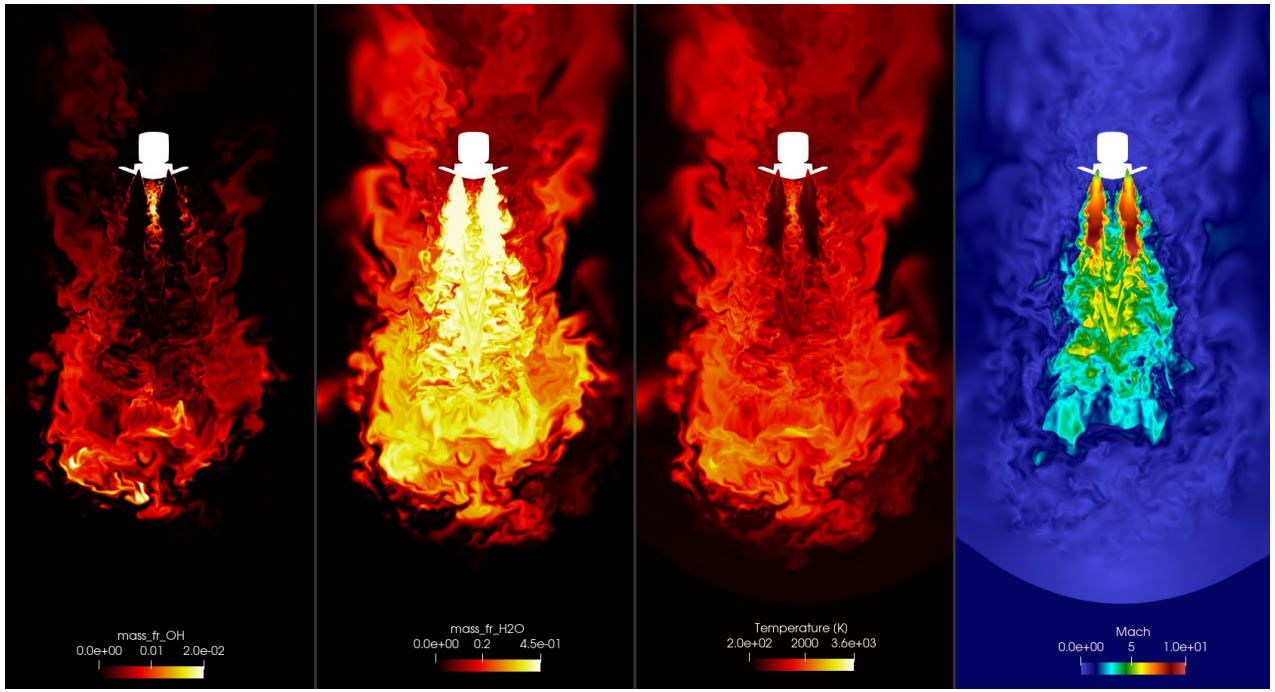
Hand-On Challenges

- Collection of self-guided challenges available to try:
  - https://github.com/olcf/hands-on-with-odo
- Challenges cover UNIX, programming environments, programming model basics, job schedulers, etc.
  - e.g., simulate and visualize galaxies colliding!
- OLCF support available in Slack
  - Additional support through help@olcf.ornl.gov



Example Frontier Use Case

- NASA is exploring ways to safely land a vehicle bringing humans to Mars
- Unable to flight-test in Martian environment
- Frontier enabled first-of-kind autonomous test flights
  - New levels of resolution, physical modeling, and temporal duration

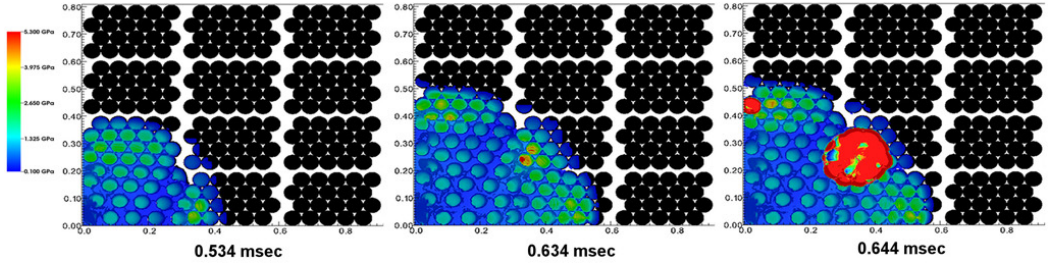


https://www.olcf.ornl.gov/wp-content/uploads/static_fine-1500x811.png

Example Leadership-Class Use Case

- Semi hauling 35,000 pounds of mining explosives crashed in Utah
- Caught fire and caused dramatic explosion leaving a 30'x70' crater
- Debris launched up to 1/4 mile
- Leadership-class systems (e.g., Titan) used to recreate explosion
- Simulations helped identify safer ways to pack mining explosives

https://www.summitdaily.com/news/trailer-full-of-explosives-blows-hole-in-utah/



https://www.jics.tennessee.edu/files/images/accidental-explosion1.jpg

Fun Facts

- 1 Exaflop => 10^{18} Calculations per Second
  - Frontier can do in 1 second what'd take over 4 years if everyone on Earth did 1 calculation/s
- 2.1 Exaflop peak double-precision
  - Compute similar to ~204K PS5s
- 77 cabinets weighing 8,000 pounds each
  - Total weight similar to a Boeing 747
- 700 PB of storage
  - 25 Mt. Everests of DVDs

https://www.flickr.com/photos/olcf/52117839159/

OAK RIDGE National Laboratory

Questions?

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

holmenjk@ornl.gov