# A Common Shorthand for Parts Nomenclature

**Working Group Outbrief**

Scalable Tools Workshop

July 26-30, 2026

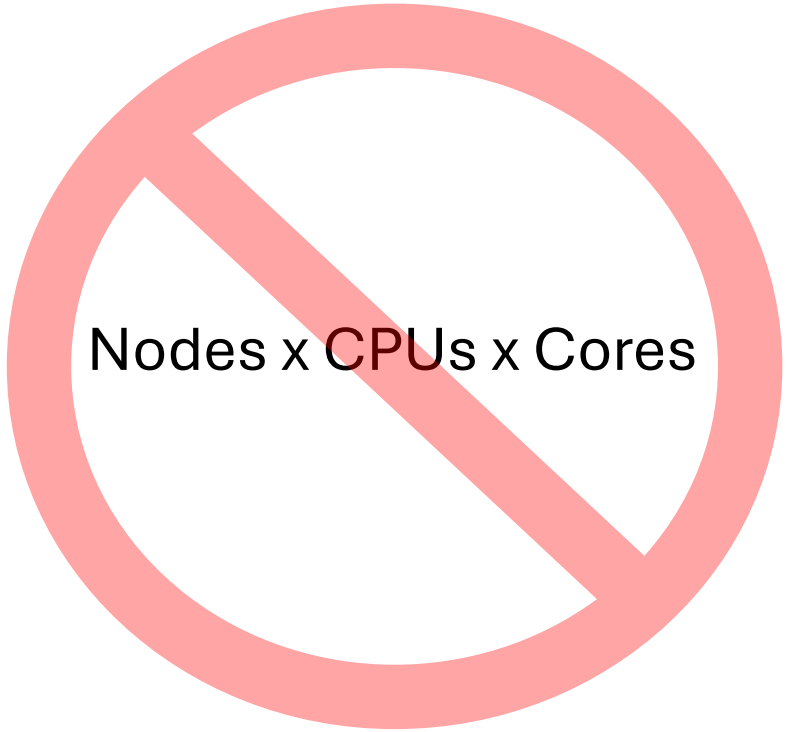
# Suddenly, There's A Wealth Of Architectures

- Common shorthand for GPU and Memory subcomponents
  - Accurately conveys the basic architecture
- Can we communicate the architecture in a way that conveys performance characteristics in an ***abbreviated*** way?
- Back in the day, we describe a machine as follows:

Nodes x CPUs x Cores

- 48 x 128 x 1  (LANL's ASC White machine)
- 18688 x 2 x 8 (ORNL' Jaguar machine)

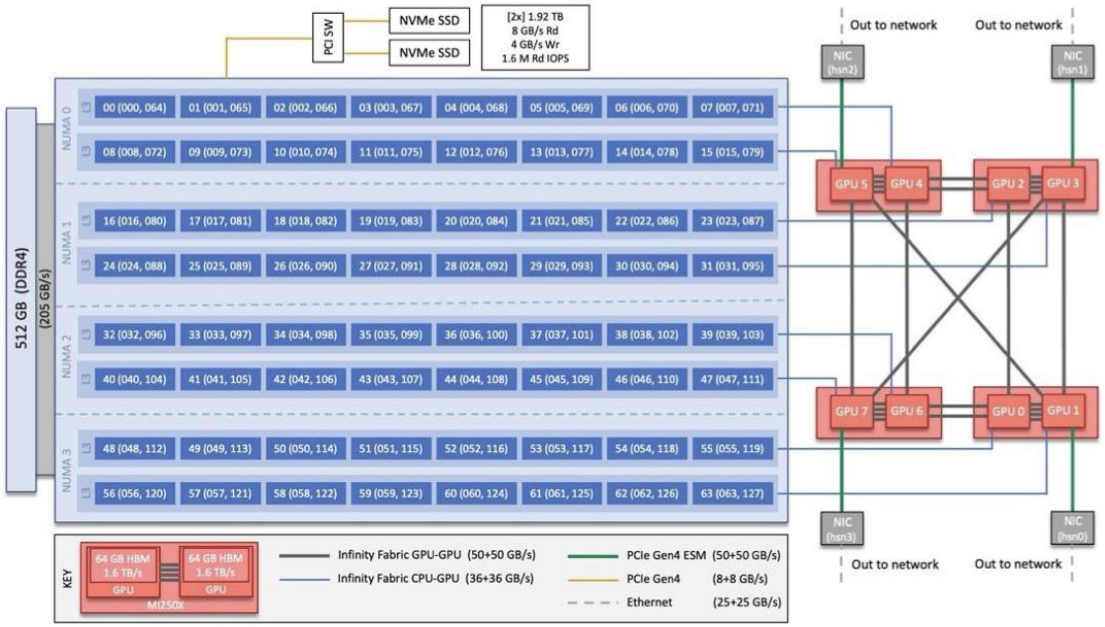
# Now We Have




- Nodes comprised of (Dies and Chiplets)
- Processing
  - Distinct CPU
  - Distinct GPU
  - Distinct AI Accelerator
  - Merged Combinations
- Memory Parts (On CPU, On GPU, On Node, Off node)
  - Cache levels
  - DDRx
  - HBM
  - Fabric Attached
- Memory NUMA Domains
  - Distinct DDRx and HBM
  - Unified combinations

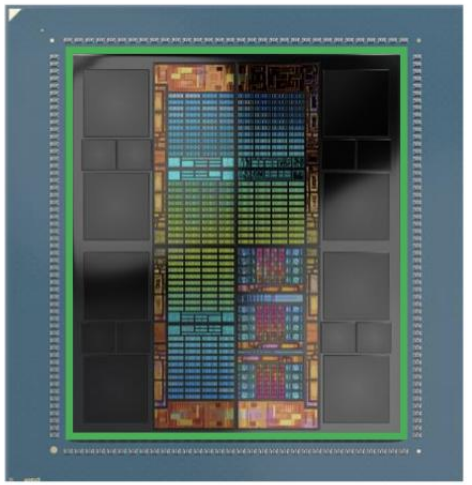
# Some AMD Architectures

AMD 3rd Gen EPYC CPU + AMD Instinct MI250X GPUs



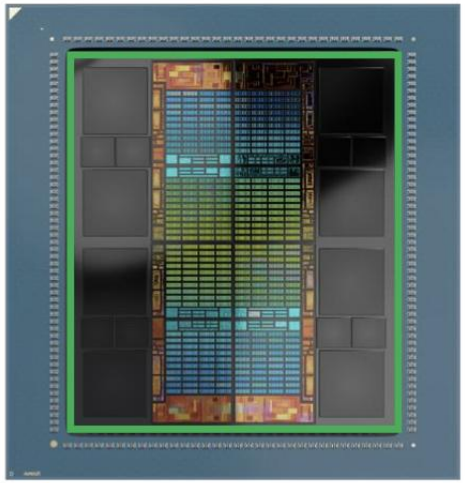

https://docs.olcf.ornl.gov/systems/crusher_quick_start_guide.html

AMD Instinct MI300A



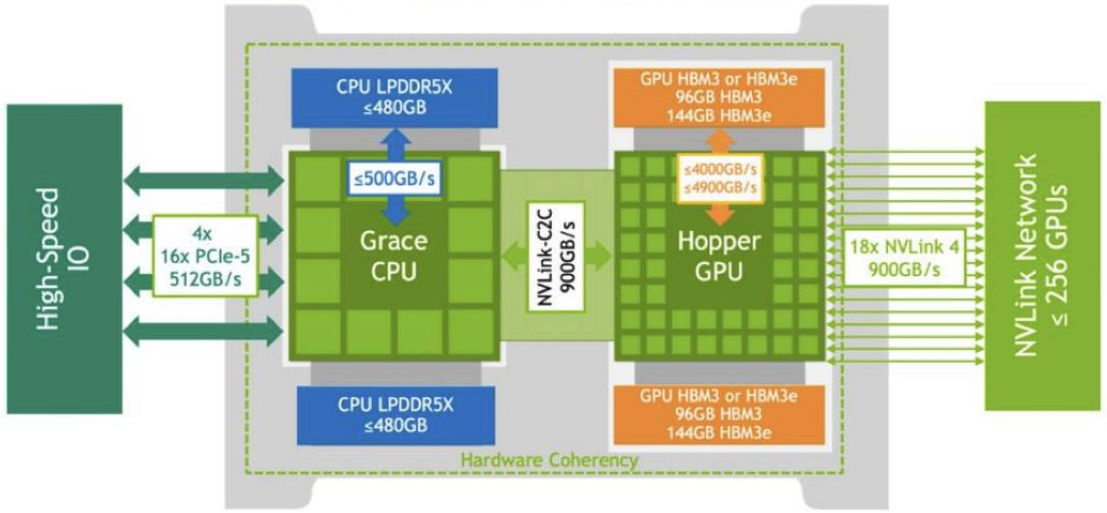
Source: AMD

AMD Instinct MI300X



Source: AMD

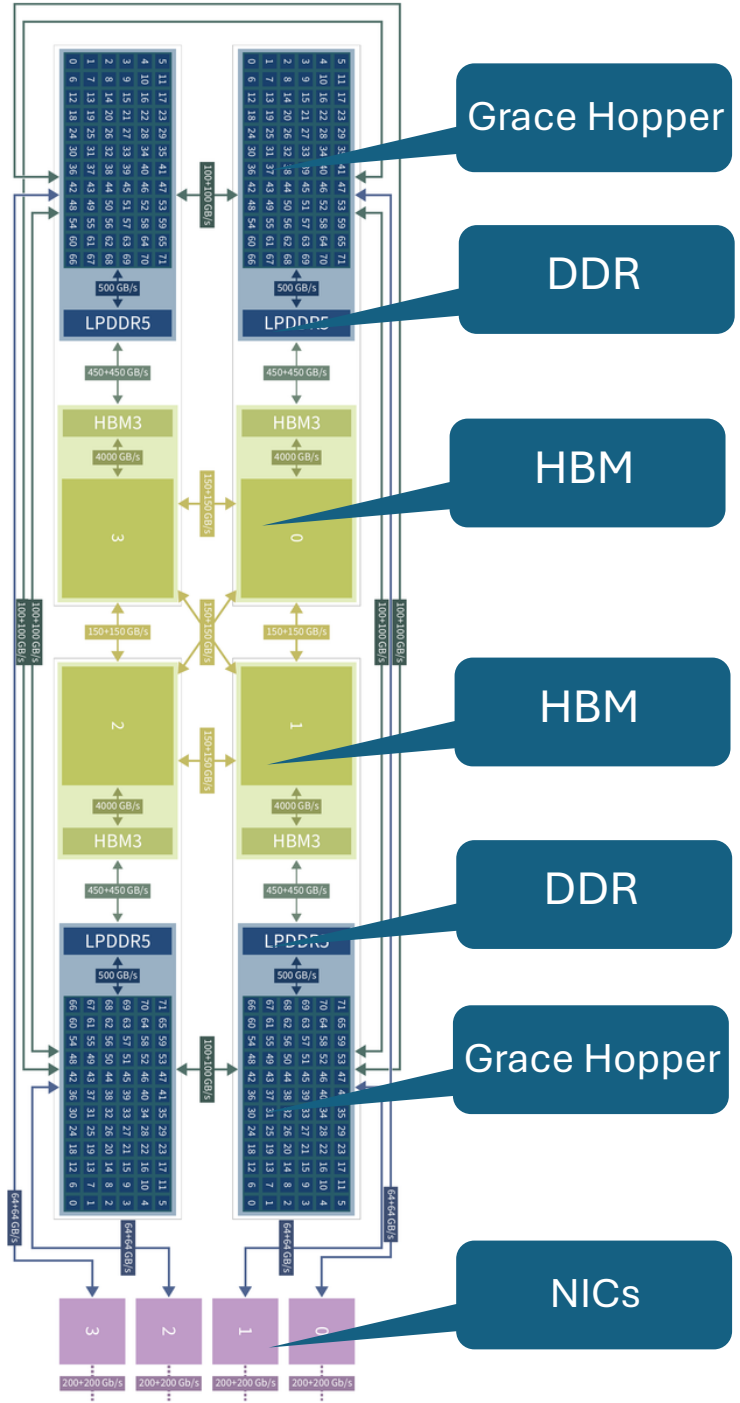
NVIDIA GH200 Grace Hopper Superchip



# Julich’s Jupiter (NVIDIA)

**24,000 NVIDIA GH200 Grace Hopper Superchips** housed across roughly 6,000 compute nodes and 125 racks.

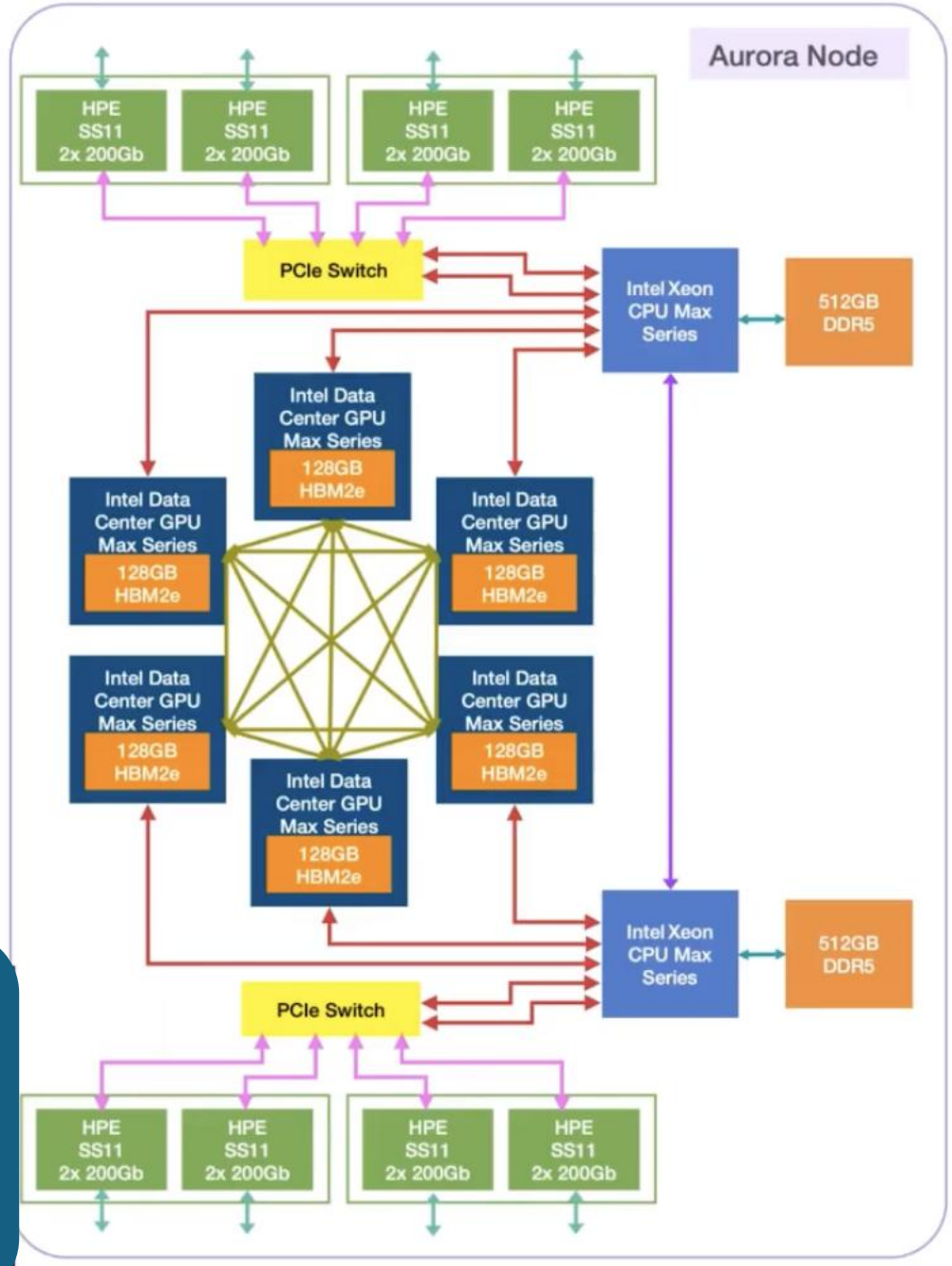
**Node Design:** Each superchip tightly couples an ARM-based Grace CPU (72 arm cores) with a Hopper GPU (132 Streaming Multiprocessors, 16,896 cuda cores) via a high-speed NVLink interconnect.



# Argonne’s Aurora (Intel)

| NODE CHARACTERISTICS |
| --- |
| **6** GPUs - Intel Data Center GPU Max Series |
| **2** CPUs - Intel Xeon CPU Max Series |
| **768 GB** GPU HBM Memory |
| **19.66 TB/s** Peak GPU HBM BW |
| **128 GB** CPU HBM Memory |
| **2.87 TB/s** Peak CPU HBM BW |
| **1024 GB** CPU DDR5 Memory |
| **0.56** TB/s Peak CPU DDR5 BW |
| **≧ 130 TF** Peak Node DP FLOPS |
| **200** GB/s Max Fabric Injection |
| **8** NICs |

Aurora has 10,624 nodes



# Motivation

- Notions of what constitutes a “node” or “CPU” has changed.
- Sometimes “CPU” means “socket”, sometimes it means “core.”
- Ambiguity of each terms semantic meaning makes system utilization and communication confusing.

# Motivation

- Memory hierarchy: HBM, cache levels, DDR, *etc.*
- Local Data Share, Shared Memory, Core Memory Group
- Other hardware: NIC, GPU, APU:
  - Titan: One “GPU” per card/device
  - Frontier: Two “GPUs” per card/device: “GCDs”
  - El Capitan: “XCDs”

# How to get the ball rolling?

- ***hwloc*** would be a good starting point, as it serves as a consistent way of describing available hardware.

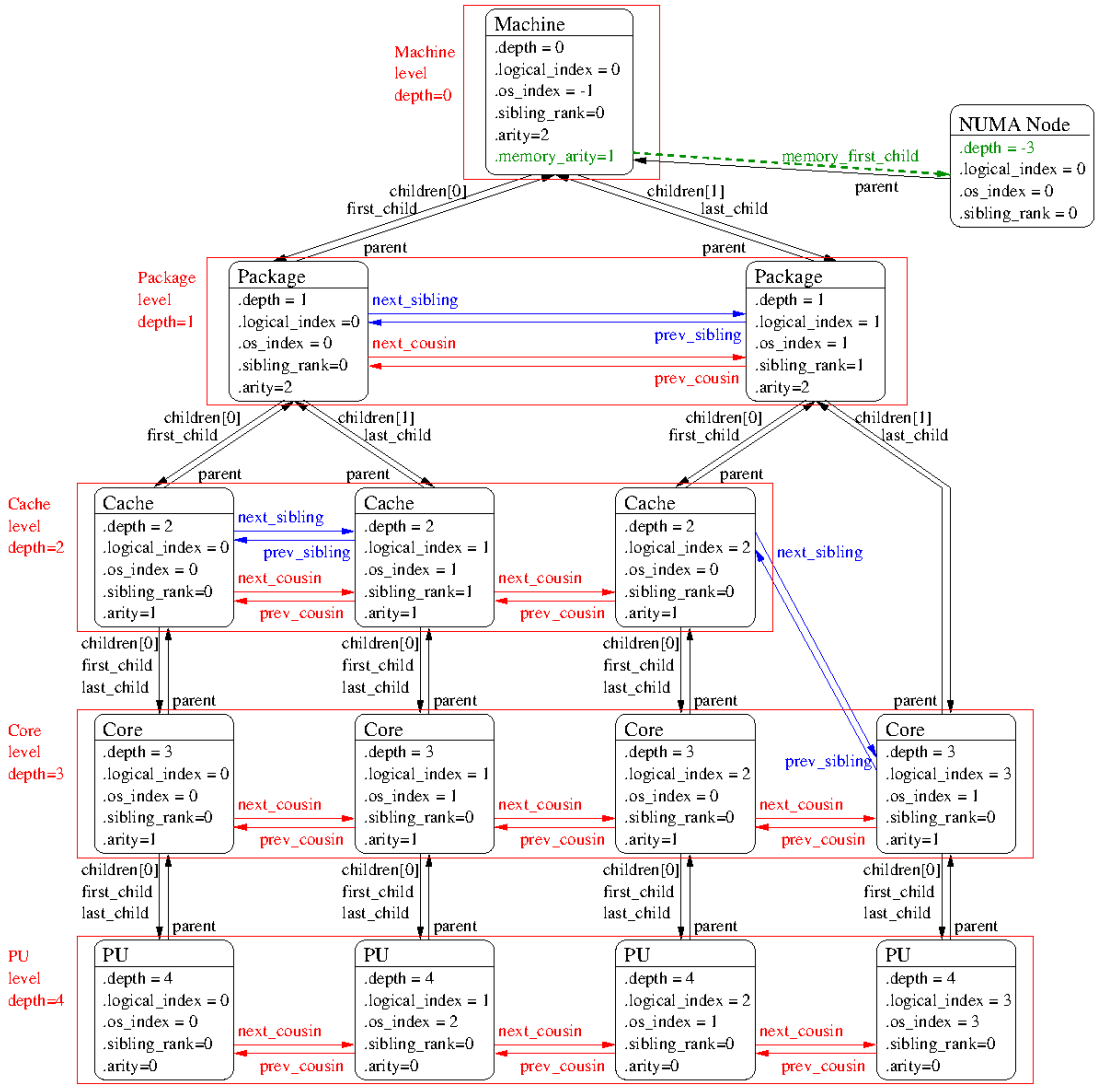
- It already describes the SMT level by the number of “PUs” per “Core” as well as the cores per cache level.

# How to get the ball rolling?

- ***SMT*** serves as a general term for how many hardware threads are allowed on a given core.
  - Example: Summit had modes “SMT[1|2|4]”

- We could have something similar for GPU partitioning.
  - Think XCD mapping on El Capitan.
  - One way to say this is GPU Partitioning: “GP[1|2|6]”

# Considering the Big Picture




- Anyone can design their own “synthetic topologies” and create an XML file for this.
- The hwloc utilities can function using this file.

https://www.open-mpi.org/projects/hwloc/doc/v2.14.0/termsanddefs.html

# Conclusions

- Request computing centers to have a public-facing model showing the topology.
- OLCF created the Job Step Viewer. Something like ***hwloc*** could make this more generalizable to other systems.
- Possibly account for system service tasks on disparate cores: Summit, Frontier, El Capitan.
- Consider a query-based tool to which a domain scientist answers a series of questions, and then it gives a salloc/srun command for the scientist to run their app.

**https://www.olcf.ornl.gov/wp-content/uploads/2020/01/OLCF_March_Con_Call_Job_Step_Viewerpdf.pdf**

# Participants

- Terry Jones
- Edgar Leon
- Zach McMichael
- Daniel Barry