OAK RIDGE National Laboratory

Summit System Overview

Tom Papatheodore – OLCF
Verónica Melesse Vergara – OLCF
Subil Abraham – OLCF
Dan Dietz – OLCF
Bill Renaud – OLCF
Brian Smith – OLCF

SIAM CSE21: Hands-on with the Summit Supercomputer
March 4, 2021

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

U.S. DEPARTMENT OF ENERGY

OLCF Summit Overview

The system includes

- 4,608 nodes
- Dual-port Mellanox EDR InfiniBand network
- 250 PB IBM file system transferring data at 2.5 TB/s

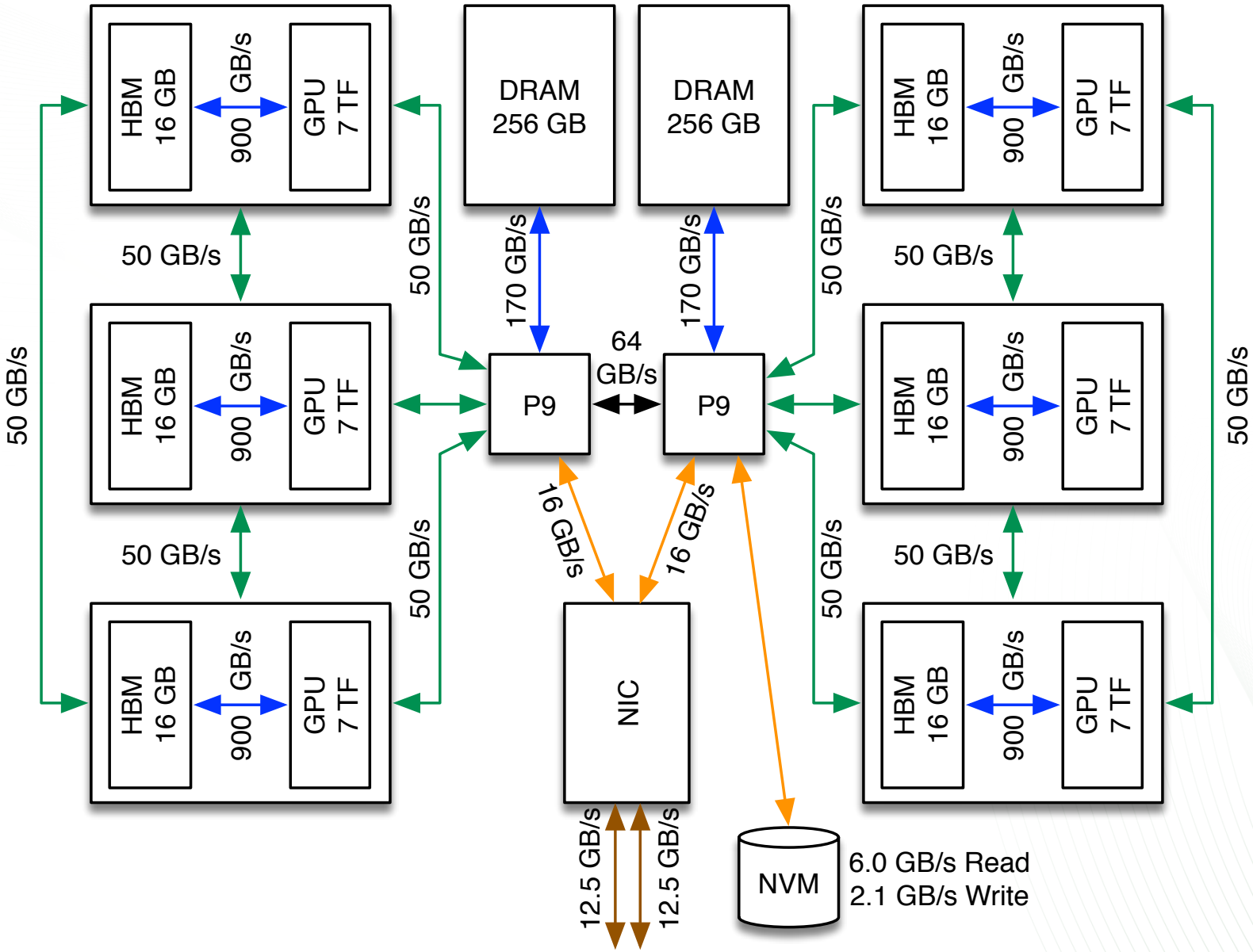
System Performance

- Peak of 200 Petaflops (FP_{64}) for modeling & simulation
- Peak of 3.3 ExaOps (FP_{16}) for data analytics and artificial intelligence

Each node has

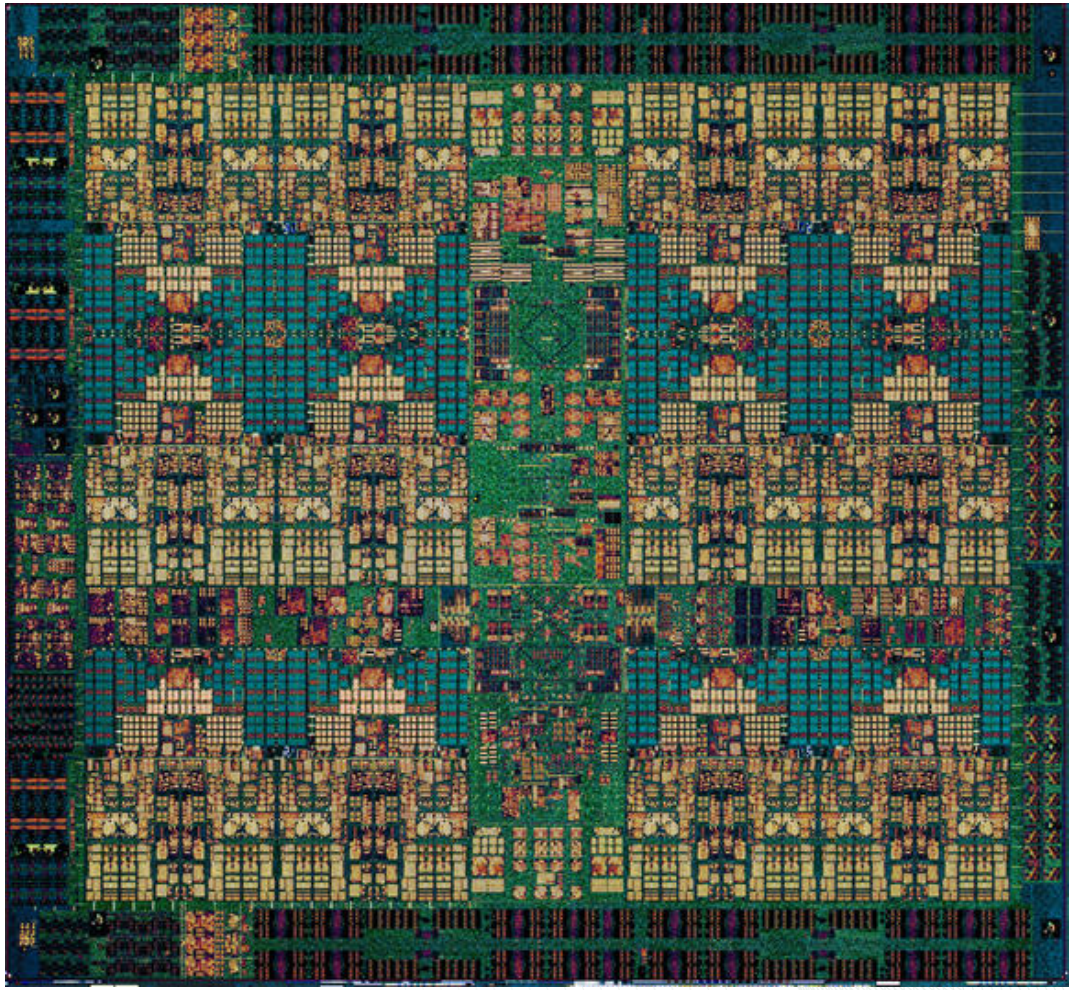
- 2 IBM POWER9 processors
- 6 NVIDIA Tesla V100 GPUs
- 608 GB of fast memory (96 GB HBM2 + 512 GB DDR4)
- 1.6 TB of NV memory

Summit Node Schematic

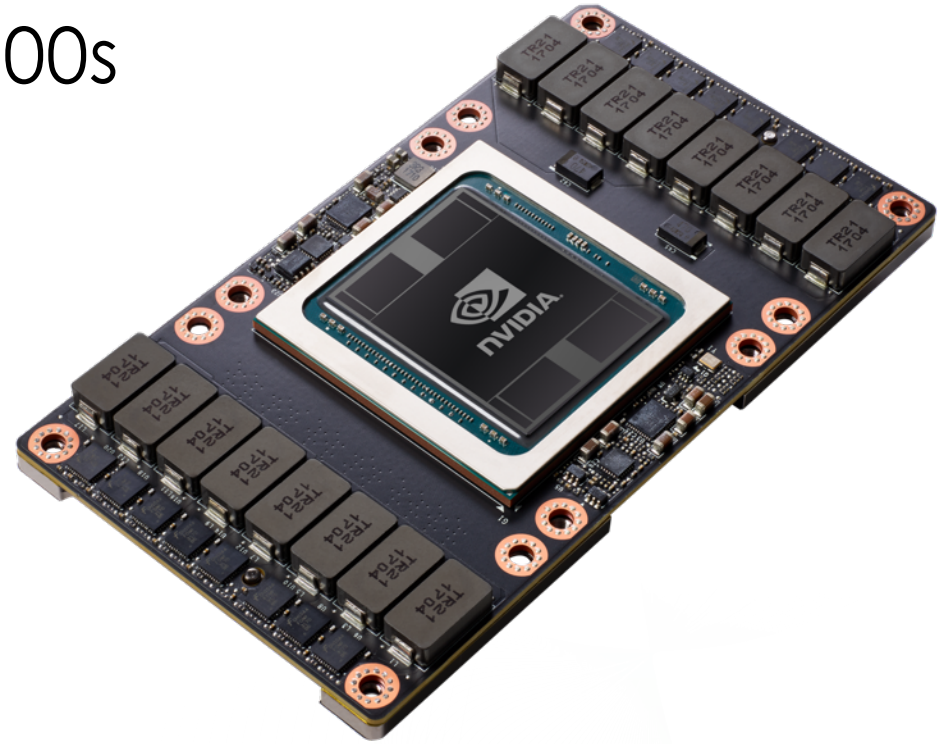


IBM Power9 Processor

- Summit's P9s: 22 cores (4 hwthreads/core)
- PCI-Express 4.0
  - Twice as fast as PCIe 3.0
- NVLink 2.0
  - Coherent, high-bandwidth links to GPUs
- 14nm FinFET SOI technology
  - 8 billion transistors
- Cache
  - L1I: 32 KiB (per core, 8-way set associative)
  - L1D: 32 KiB (per core, 8-way)
  - L2: 512 KiB (per pair of cores)
  - L3: 120 MiB eDRAM, 20-way (shared by all cores)



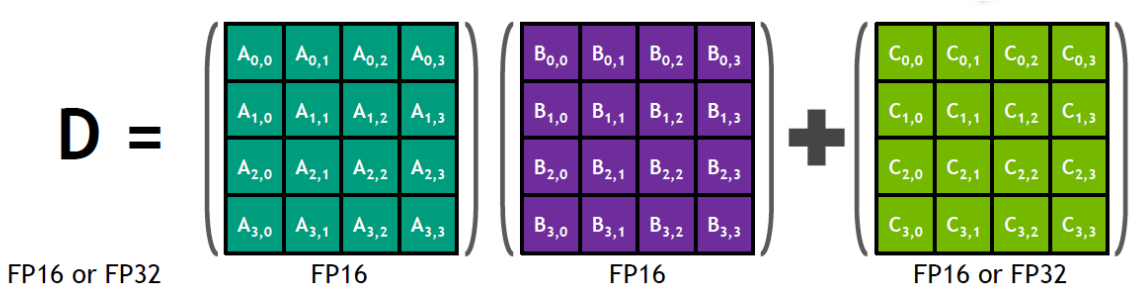
Summit Contains 27,648 NVIDIA Tesla v100s



Each Tesla v100 GPU has:

- 150+150 GB/s total BW (NVLink v2.0)
- 5,120 CUDA cores (64 on each of 80 SMs)
- 640 Tensor cores (8 on each of 80 SMs)
- 20MB Registers | 16MB Cache | 16GB HBM2 @ 900 GB/s
- 7.5 DP TFLOPS | 15 SP TFLOPS | 120 FP_{16} TOPS

- Tensor cores do mixed precision multiply-add of 4x4 matrices



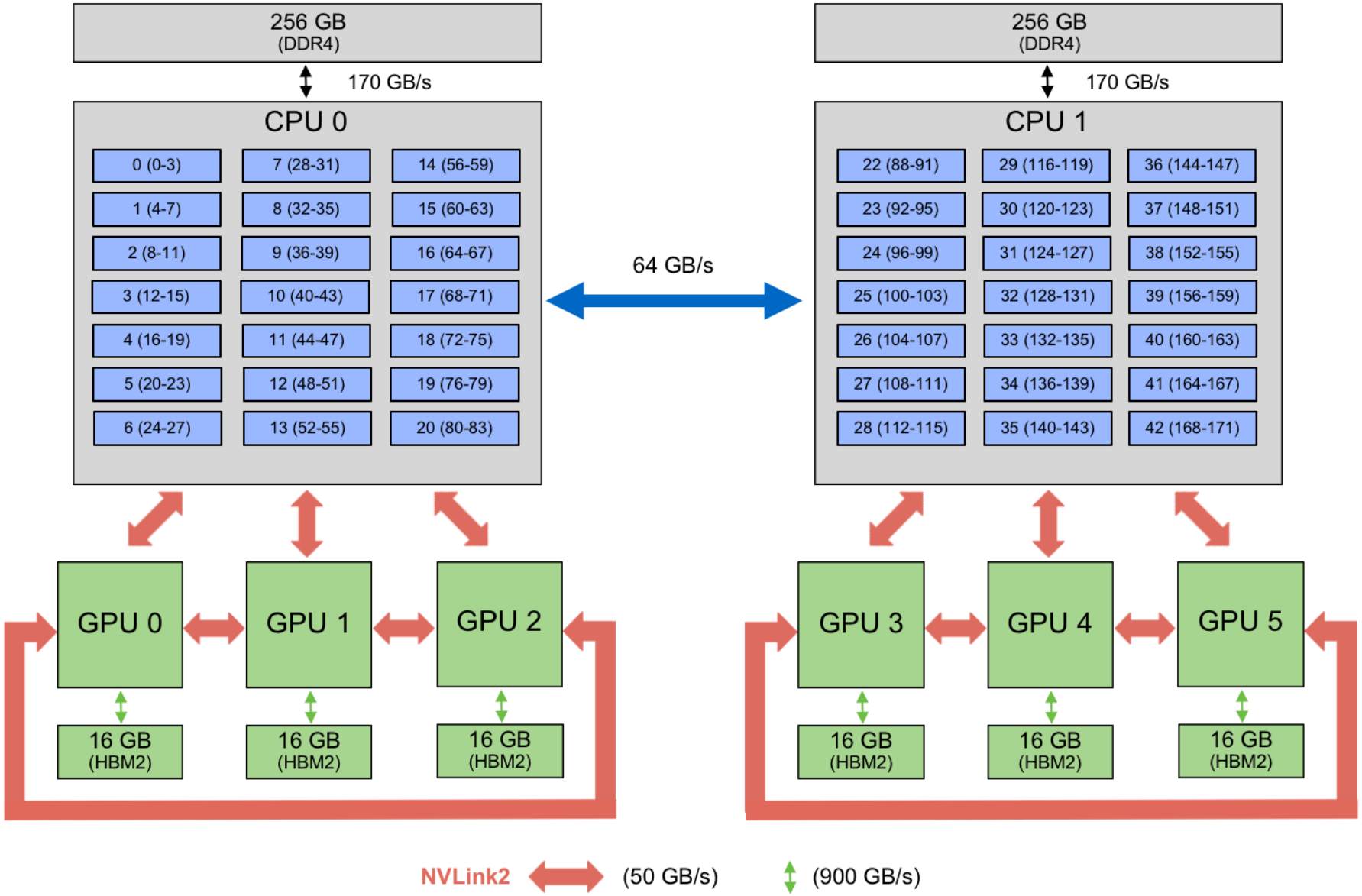
Type	Size	Range	$u = 2^{-t}$
half	16 bits	$10^{\pm 5}$	$2^{-11} \approx 4.9 \times 10^{-4}$
single	32 bits	$10^{\pm 38}$	$2^{-24} \approx 6.0 \times 10^{-8}$
double	64 bits	$10^{\pm 308}$	$2^{-53} \approx 1.1 \times 10^{-16}$
quadruple	128 bits	$10^{\pm 4932}$	$2^{-113} \approx 9.6 \times 10^{-35}$

$$D = AB + C$$

https://docs.olcf.ornl.gov/systems/summit_user_guide.html#tensor-cores

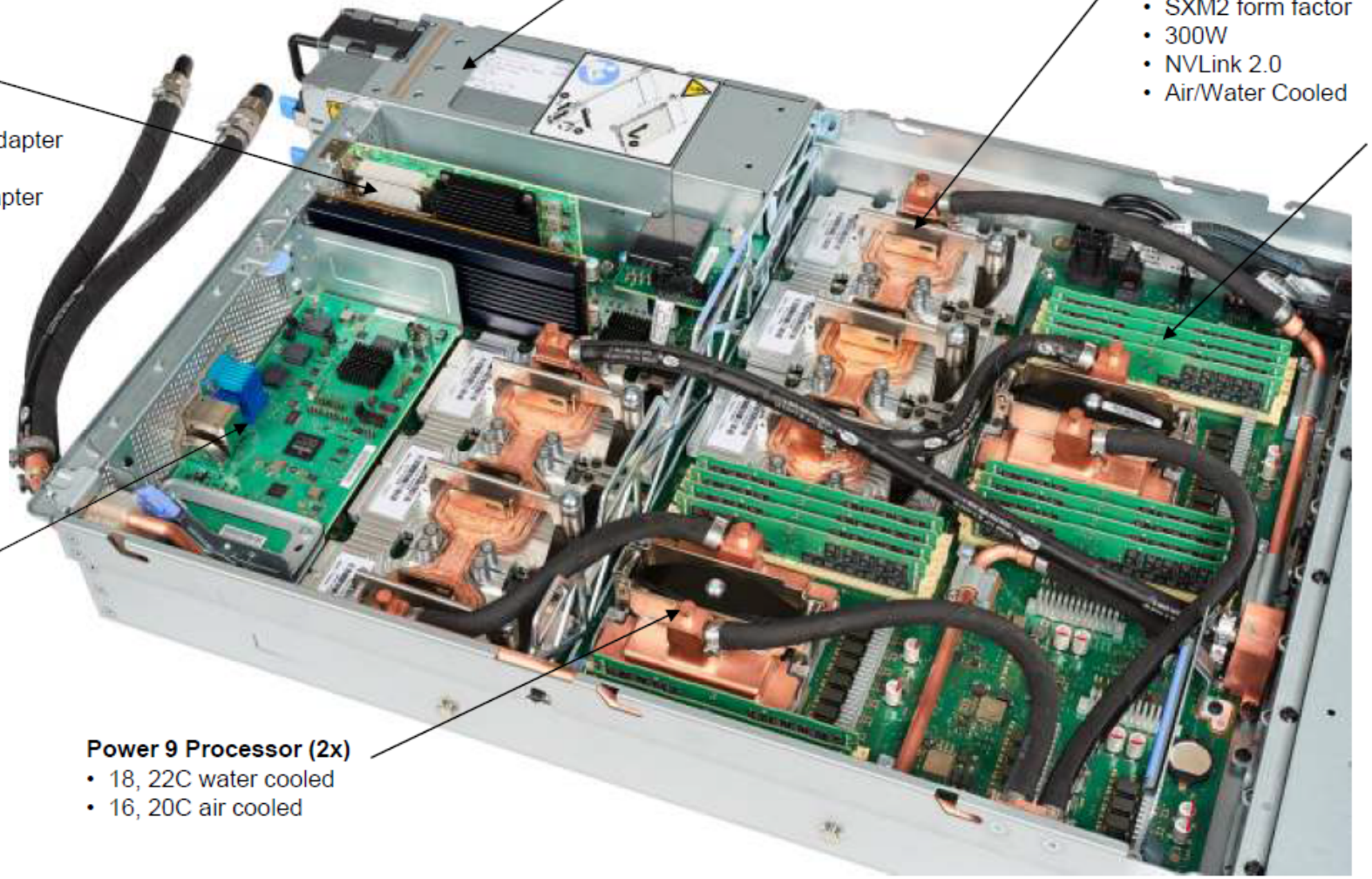
Summit Node

(2) IBM Power9 + (6) NVIDIA Volta V100



OAK RIDGE National Laboratory

Summit Board (1 node) showing the Water Cooling



Power Supplies (2x)
- 2200W
- 200VAC, 277VAC, 400VDC input

NVidia Volta GPU
- 3 per socket
- SXM2 form factor
- 300W
- NVLink 2.0
- Air/Water Cooled

PCIe slot (4x)
- Gen4 PCIe
- 2, x16 HHHL Adapter
- 1, Shared slot
- 1 x8 HHHL Adapter

Memory DIMM's (16x)
- 8 DDR4 IS DIMMs per socket
- 8, 16, 32,64, 128GB DIMMs

BMC Card
- IPMI
- 1 Gb Ethernet
- VGA
- 1 USB 3.0

Power 9 Processor (2x)
- 18, 22C water cooled
- 16, 20C air cooled

Summit Specs

Feature	Summit
Peak FLOPS	200 PF
Max possible Power	13 MW
Number of Nodes	4,608
Node performance	42 TF
Memory per Node	512 GB DDR4 + 96 GB HBM2
NV memory per Node	1.6 TB
Total System Memory	2.8 PB + 7.4 PB NVM
System Interconnect	Dual Port EDR-IB (25 GB/s)
Interconnect Topology	Non-blocking Fat Tree
Bi-Section Bandwidth	115.2 TB/s
Processors on node	2 IBM POWER9™ 6 NVIDIA Volta™
File System	250 PB, 2.5 TB/s, GPFS™

Available File Systems / Storage Areas on Ascent

NFS Directories – This is where you might want to keep source code and build your application.

NOTE: These directories are read-only from the compute nodes!

`/ccsopen/home/userid`

- Your personal home directory

`/ccsopen/proj/gen153`

- Can be accessed by all participants of this event

GPFS Directories (parallel file system) – This is where you should write data when running on Ascent's compute nodes.

`/gpfs/wolf/gen153/scratch/userid`

- Your personal GPFS scratch directory

`/gpfs/wolf/gen153/proj-shared`

- Can be accessed by all participants of the event

Questions?