

# Summit Updates

Presented to  
2021 OLCF User Meeting

Don Maxwell  
HPC Scalable Systems Group Leader

June 22, 2021

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

# Summit IBM AC922

## Specifications and Features

- Processor: IBM Power9™ (2/node)
- GPUs: ~~27,648~~ **28,080** NVIDIA Volta V100s (6/node)
- Nodes: ~~4,608~~ **4,680**
- Node Performance: 42TF
- Memory/node: 512GB DDR4 + 96GB HBM2
- NV Memory/node: 1.6TB
- Total System Memory: >10PB DDR4 + HBM + Non-volatile
- Interconnect Topology: Mellanox EDR 100G InfiniBand, Non-blocking Fat Tree
- Peak Power Consumption: 13MW



TOP 500 The List.

HOME LISTS STATISTICS RESOURCES ABOUT MEDIA KIT

Home / Lists / TOP500 / June 2018

### JUNE 2018

The TOP500 celebrates its 25<sup>th</sup> anniversary with a major shakeup at the top of the list, the first time since November 2012, the U.S. claims the most powerful supercomputer in the world.

#### TOP 10 Sites for November 2019

For more information about the sites and systems in the list, click on the links or view the complete list.

Summit, an IBM Power CPU + NVIDIA Volta GV100 GPU System at DOE/SC/Oak Ridge National Laboratory

Rank	System	Cores	Rmax (TFlop/s)	HPCG (TFlop/s)
1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory	2,414,592	148,600.0	2925.75

HPCG List for November 2019

Rank	System	Cores	Rmax (TFlop/s)	HPCG (TFlop/s)
1	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory, United States	2,414,592	148,600.0	2925.75

The GREEN 500 CERTIFICATE

Summit, an IBM Power CPU + NVIDIA Volta GV100 GPU System at DOE/SC/Oak Ridge National Laboratory

TOP 500 CERTIFICATE

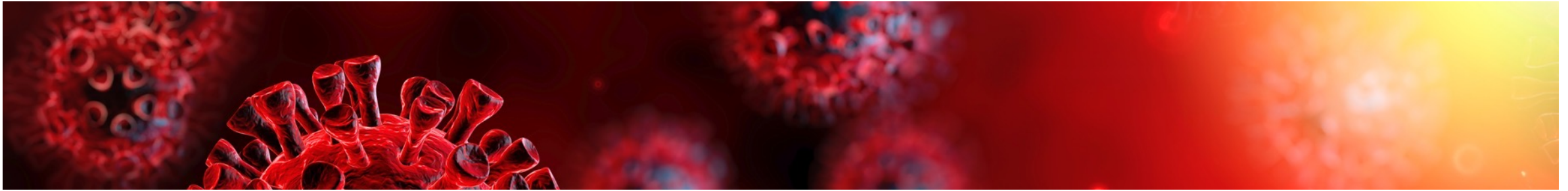
Summit, an IBM Power System AC922 at the U.S. Department of Energy / SC / Oak Ridge National Laboratory, TN, USA

is ranked **No. 1** among the World's TOP500 Supercomputers with 122.3 PFlop/s Linpack Performance on the TOP500 List published at ISC High Performance, June 25, 2018

Congratulations from the TOP500 Editors

Eric Steinhauser NERSC/Berkeley Lab  
Dhananjay Choudhary University of Tennessee  
Harish Shirohetti NERSC/Berkeley Lab  
Markus Muehle SC Group

# Summit COVID Hardware



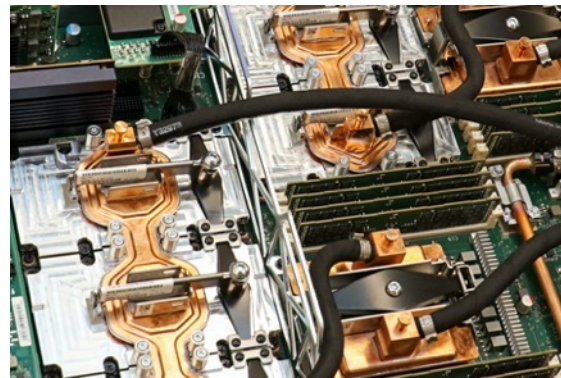
- Add 54 nodes, each with double the HBM and 4x the DDR4 and NVMe
- Allows jobs that need larger on-node memory to run on up to 54 nodes
- Applications that benefit
  - Computational Chemistry
  - AI/DL for medical image analysis

Feature	Current Summit Nodes	Large Memory Nodes
Peak FLOPS <sub>64</sub>	200 PF	203 PF
Number of Nodes	4,680	54
Node performance	43 TF	43 TF
Memory per Node	512 GB DDR4 + 96 GB HBM2	2048 GB DDR4 + 192 GB HBM2
NV memory per Node	1.6 TB	6.4 TB
Total System Memory	2.8 PB + 7.4 PB NVM	2.9 PB + 7.7 PB NVM
System Interconnect	Dual Rail EDR-IB (25 GB/s)	Dual Rail EDR-IB (25 GB/s)
Interconnect Topology	Non-blocking Fat Tree	Non-blocking Fat Tree
Bi-Section Bandwidth	115.2 TB/s	115.2 TB/s
Processors on node	2 IBM POWER9™ 6 NVIDIA Volta™	2 IBM POWER9™ 6 NVIDIA Volta™
File System	250 PB, 2.5 TB/s, GPFS™	250 PB, 2.5 TB/s, GPFS™

# Summit Contingency Hardware

3 New  
Summit  
Cabinets

54 additional nodes  
added in March 2021



Summit  
Specs

Effectively same  
specifications as  
original Summit  
hardware

6<sup>th</sup> Year  
Spares

Would be used as  
spare parts if Summit  
extended into a 6<sup>th</sup>  
year



# How do I access the new nodes?

- COVID nodes are a separate partition due to heterogeneity
  - Batch queue name has a trailing -hm (high memory)
    - batch-hm
  - All users have access
- Contingency nodes are included in regular queues
  - Simply increased the normal machine size providing the ability to submit jobs up to 4662 nodes
- Queues for longer walltimes for small jobs
  - COVID nodes (**batch-hm**) provides longer walltime flexibility for small jobs (i.e., a single-node job can run for 24 hours)
  - **killable** queue provides this functionality to regular Summit partition
    - Caveats
      - Preemptable with the following guarantees

Bin	Min Nodes	Max Nodes	Max Walltime (Hours)	Guaranteed Walltime
4	46	91	24.0	6.0 (hours)
5	1	45	24.0	2.0 (hours)

[https://docs.olcf.ornl.gov/systems/summit\\_user\\_guide.html#killable-queue-policy](https://docs.olcf.ornl.gov/systems/summit_user_guide.html#killable-queue-policy)

# Summit Software

- Currently
  - RHEL 7.6
  - CUDA 10.1
  - MOFED 4.7
  - LSF 10.1.0.9
  - xCAT 2.15
- Summit IBM HPC Software Stack Upgrade
  - Software Updates
    - RHEL 8.2
    - CUDA 11
    - MOFED 4.9
    - LSF 10.1.0.11
    - xCAT 2.16
    - Updates to LAPACK, LAPACKE, and LAPACK GPU support in ESSL
    - GDRCopy support in Spectrum MPI
    - JSM
      - Improved GPU Binding options
      - Support for multiple SMPI installs with a single JSM





Questions?