Master Informatics Eng.

2021/22

A.J.Proença

Top HPC systems in TOP500 lists

(most slides & images are borrowed)

WIKIPEDIA
The Free Encyclopedia

What is TOP500?

TOP500

From Wikipedia, the free encyclopedia

The **TOP500** project ranks and details the 500 most powerful non-distributed computer systems in the world. The project was started in 1993 and publishes an updated list of the supercomputers twice a year. The first of these updates always coincides with the International Supercomputing Conference in June, and the second is presented at the ACM/IEEE Supercomputing Conference in November. The project aims to provide a reliable basis for tracking and detecting trends in high-performance computing and bases rankings on HPL,[1] a portable implementation of the high-performance LINPACK benchmark written in Fortran for distributed-memory computers.

Currently the latest TOP500 list is the ~~57th~~ 58th, published in ~~June~~ November 2021. Since June 2020, the Japanese Fugaku is the world's most powerful supercomputer, reaching initially 415.53 petaFLOPS and 442.01 petaFlops after an update in November 2020 on the LINPACK benchmarks. China currently dominates the list with ~~188~~ 173 supercomputers, leading the second place [150] (United States).

TOP500	
Key people	Erich Strohmaier Jack Dongarra Horst Simon Martin Meuer
Established	24 June 1993
Website	www.top500.org

The TOP500 list is compiled by Jack Dongarra of the University of Tennessee, Knoxville, Erich Strohmaier and Horst Simon of the National Energy Research Scientific Computing Center (NERSC) and Lawrence Berkeley National Laboratory (LBNL), and, until his death in 2014, Hans Meuer of the University of Mannheim, Germany.

The TOP500 project lists also Green500 and HPCG benchmark list.

Outline

1. TOP500 *(LINPACK)*
   a) TOP10 lists from Nov'17 to Nov'21
   b) Country distribution over the past 25 years
   c) PU chip technology evolution in the past 25 years and since last year
   d) Evolution of the accelerators since they were available
   e) Analysis of some relevant systems and architectures

2. GREEN500
   a) TOP10 lists from Nov'17 to Nov'21
   b) Analysis of some relevant systems and architectures

3. HPCG500
   a) HPCG *vs.* HPL: an overview
   b) TOP10 lists from Nov'17 to Nov'20
   c) Analysis of some relevant systems

4. HPL-AI
   a) High-performance Linpack (HPL) and Artificial Intelligence (AI) workloads

LINPACK benchmarks (HPL)

> HPL measures the sustained floating-point rate (GFLOPs/s) to solve a dense system of linear equations using double-precision floating-point arithmetic

LINPACK benchmarks

From Wikipedia, the free encyclopedia

For the software library, see LINPACK.

The **LINPACK Benchmarks** are a measure of a system's floating point computing power. Introduced by Jack Dongarra, they measure how fast a computer solves a dense *n* by *n* system of linear equations $Ax = b$, which is a common task in engineering.

The latest version of these benchmarks is used to build the TOP500 list, ranking the world's most powerful supercomputers.[1]

The aim is to approximate how fast a computer will perform when solving real problems. It is a simplification, since no single computational task can reflect the overall performance of a computer system. Nevertheless, the LINPACK benchmark performance can provide a good correction over the peak performance provided by the manufacturer. The peak performance is the maximal theoretical performance a computer can achieve, calculated as the machine's frequency, in cycles per second, times the number of operations per cycle it can perform. The actual performance will always be lower than the peak performance.[2] The performance of a computer is a complex

LINPACK benchmarks

Original author(s)	Jack Dongarra, Jim Bunch, Cleve Moler, and Gilbert Stewart
Initial release	1979
Website	www.netlib.org/benchmark/hpl/

Top 10 HPC systems
Nov'17 TOP500

3x systems in Top10
w/ Xeon Phi KNL

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	17,808
3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland				
4	**Gyoukou** - ZettaScaler-2.2 HPC system, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz , ExaScaler Japan Agency for Marine-Earth Science and Technology Japan				
5	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States				
6	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890
7	**Trinity** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	979,968	14,137.3	43,902.6	3,844
8	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	3,939
9	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	2,719
10	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	12,660

Top 10 HPC systems
Nov'18 TOP500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
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	2,397,824	143,500.0	200,794.9	9,783
2	**Sierra** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
3	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
4	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
5	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	387,872	21,230.0	27,154.3	2,384
6	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	979,			
7	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,			
8	**SuperMUC-NG** - ThinkSystem SD530, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path , Lenovo Leibniz Rechenzentrum Germany	305,856	19,476.6	26,873.9	
9	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	8,209
10	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	7,890

Top 10 HPC systems
Nov'19 TOP500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
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	200,794.9	10,096
2	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
3	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
4	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 , NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
5	**Frontera** - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox InfiniBand HDR , Dell EMC Texas Advanced Computing Center/Univ. of Texas United States	448,448	23,516.4	38,745.9	
6	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray/HPE Swiss National Supercomputing Centre (CSCS) Switzerland	387,872			
7	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray/HPE DOE/NNSA/LANL/SNL United States	979,072			
8	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	32,576.6	1,649
9	**SuperMUC-NG** - ThinkSystem SD650, Xeon Platinum 8174 24C 3.1GHz, Intel Omni-Path , Lenovo Leibniz Rechenzentrum Germany	305,856	19,476.6	26,873.9	
10	**Lassen** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Tesla V100 , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	288,288	18,200.0	23,047.2	

Frontera (TACC): successor of Stampede2

Top 10 HPC systems
Nov'20 TOP500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	**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	200,794.9	10,096
3	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	79,215.0	2,646
6	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,[illegible]			
7	**JUWELS Booster Module** - Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos Forschungszentrum Juelich (FZJ) Germany	449,[illegible]			
8	**HPC5** - PowerEdge C4140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, Dell EMC Eni S.p.A. Italy	669,760	35,450.0	51,720.8	2,252
9	**Frontera** - Dell C6420, Xeon Platinum 8280 28C 2.7GHz, Mellanox InfiniBand HDR, Dell EMC Texas Advanced Computing Center/Univ. of Texas United States	448,448	23,516.4	38,745.9	
10	**Dammam-7** - Cray CS-Storm, Xeon Gold 6248 20C 2.5GHz, NVIDIA Tesla V100 SXM2, InfiniBand HDR 100, HPE Saudi Aramco Saudi Arabia	672,520	22,400.0	55,423.6	

3x

Top 10 HPC systems
Nov'21 TOP500

Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	537,212.0	29,899
2	**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	200,794.9	10,096
3	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014[illegible]	125,435.9	15,371
5	**Perlmutter** - HPE Cray EX235n, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 SXM4 40 GB, Slingshot-10, HPE DOE/SC/LBNL/NERSC United States	761,856	70,870.0	93,750.0	2,589
6	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	55[cut off]			
7	**Tianhe-2A** - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,98[cut off]			
8	**JUWELS Booster Module** - Bull Sequana XH200[illegible], AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellano[illegible] HDR InfiniBand/ParTec ParaStation ClusterSuite, [illegible]tos Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	70,980.0	1,764
9	**HPC5** - PowerEdge C4140, Xeon Gold [illegible]52 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR In[illegible]iband, DELL EMC Eni S.p.A. Italy	669,760	35,450.0	51,720.8	2,252
10	**Voyager-EUS2** - ND96amsr_A100_v4, AMD EPYC 7V12 48C 2.45GHz, NVIDIA A100 80GB, Mellanox HDR Infiniband, Microsoft Azure Azure East US 2 United States	253,440	30,050.0	39,531.2	

New in the TOP10 list since June'21, with Zen3

Microsoft Azure, pushed Frontera to 13th…

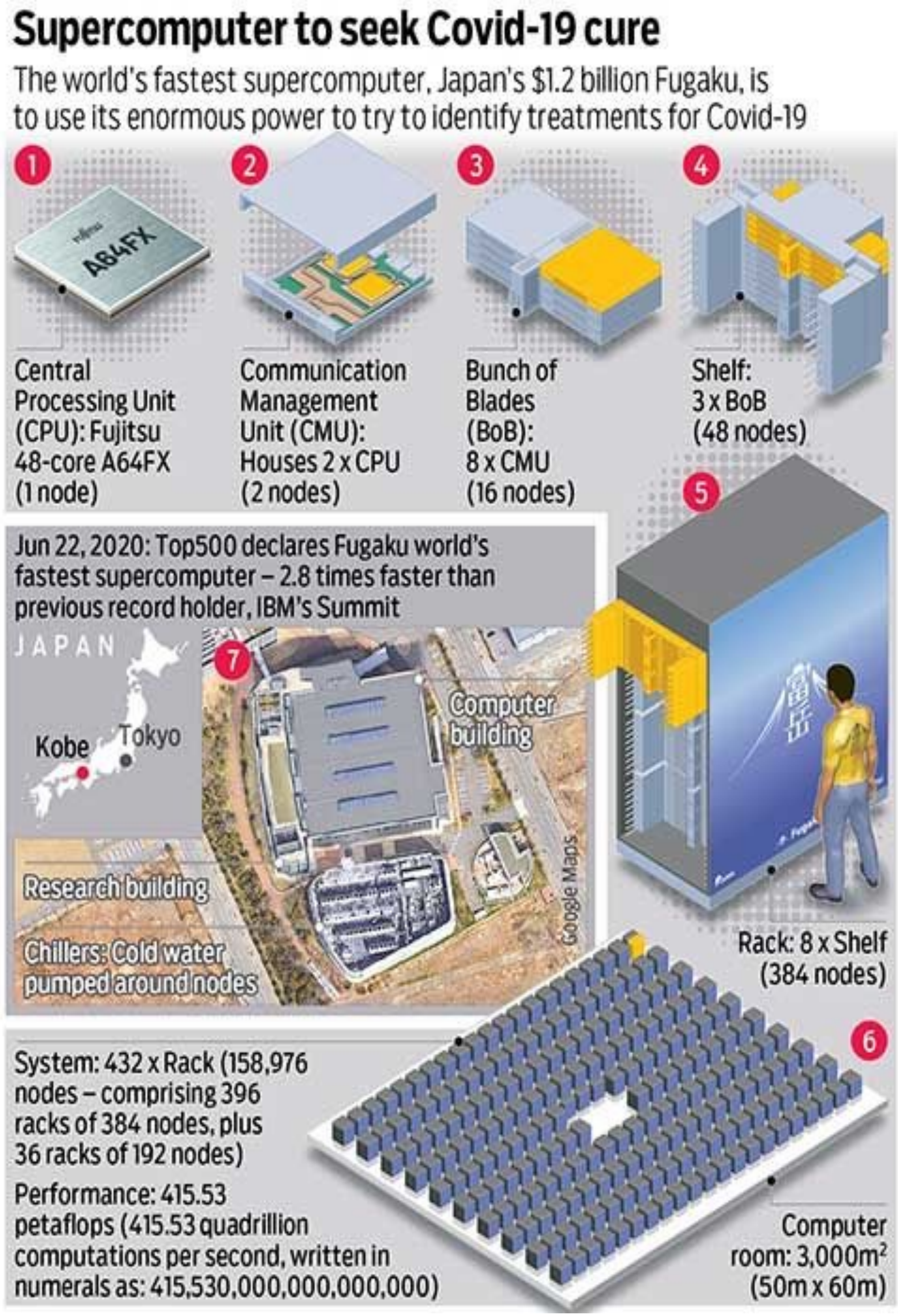
Analysis of the key systems in 2021

1. **#1** in Nov'21 (#1 in Jun'20): **Fugaku** *(Fujitsu A64FX, 48 cores)*, follow-up of
   #1 in Jun'11, **K-Computer** *(SPARC64 VIIIfx, 8 cores)*
2. **#2** in Nov'21 (#1 in Nov'18): **Summit** *(IBM POWER9, 22 cores + NVidia Volta GV100)* +**Sierra,** follow-up of
   #1 in Jun'12, **Sequoia** *(IBM POWER BGQ, 16 cores)*
3. **#4** in Nov'21 (#1 in Nov'17): **TaihuLight** *(Sunway SW26010, 260 c)*
4. **#6** in Nov'21: **Selene** *(AMD Epyc Rome 64 c + NVidia A100)*
5. **#7** in Nov'21: **Tianhe-2A** (MilkyWay-2A) *(Xeon, 12c + Matrix-2000)*, follow-up of
   #1 in Jun'13, **Tianhe-2** (MilkyWay-2) *(Xeon, 6 c + Xeon Phi 31S1P)*
   #1 in Nov'10, **Tianhe-1A** (MilkyWay-1A) *(Xeon, 6 c + NVidia Fermi)*

① 富岳 *Fugaku*

1 **Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu
RIKEN Center for Computational Science
Japan

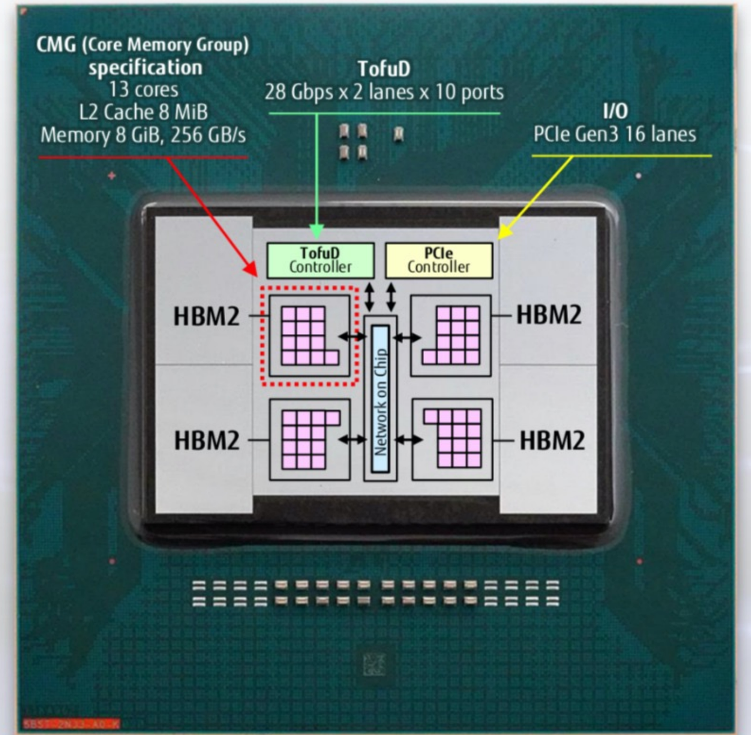
since Jun'20



Fujitsu A64FX in Fugaku

1. High-Performance Arm CPU A64FX in HPC and AI Areas

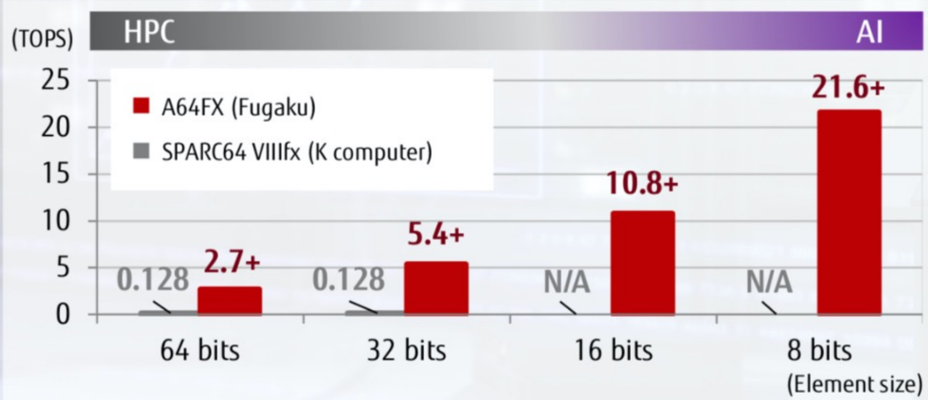
FUJITSU



Architecture features

ISA	Armv8.2-A (AArch64 only) SVE (Scalable Vector Extension) arm
SIMD width	512-bit
Precision	FP64/32/16, INT64/32/16/8
Cores	48 computing cores + 4 assistant cores (4 CMGs)
Memory	HBM2: Peak B/W 1,024 GB/s
Interconnect	TofuD: 28 Gbps x 2 lanes x 10 ports

Peak performance (Chip level)



4

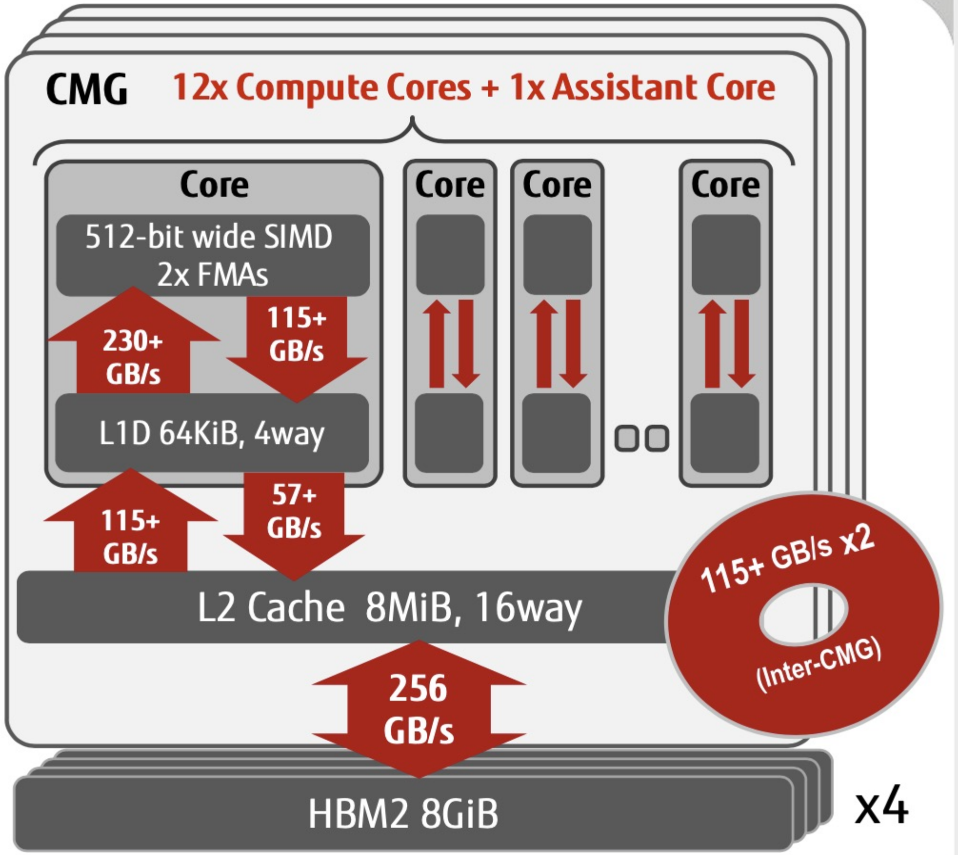
Fujitsu A64FX in Fugaku

Fujitsu-designed CPU Core w/ High Memory Bandwidth

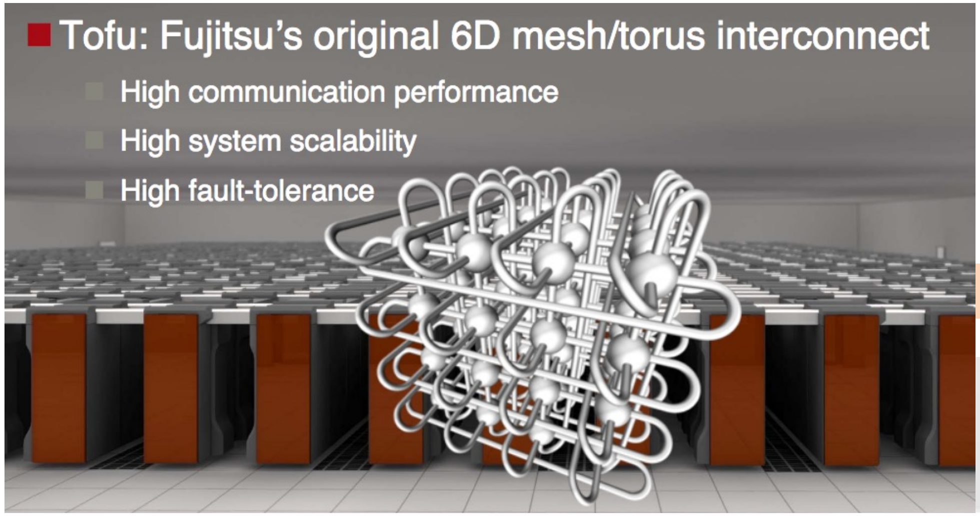
FUJITSU

- A64FX out-of-order controls in cores, caches, and memories achieve superior throughput

BW and calc. perf.	A64FX	B/F
DP floating perf. (TFlops)	2.7+	-
L1 data cache (TB/s)	11+	4
L2 cache (TB/s)	3.6+	1.3
Memory BW (GB/s)	1024	0.37

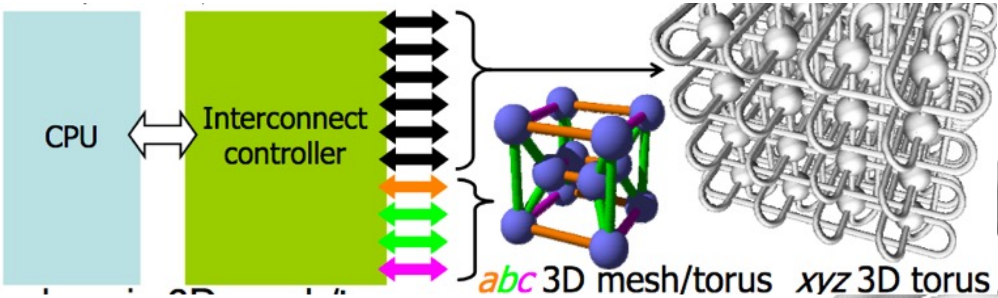


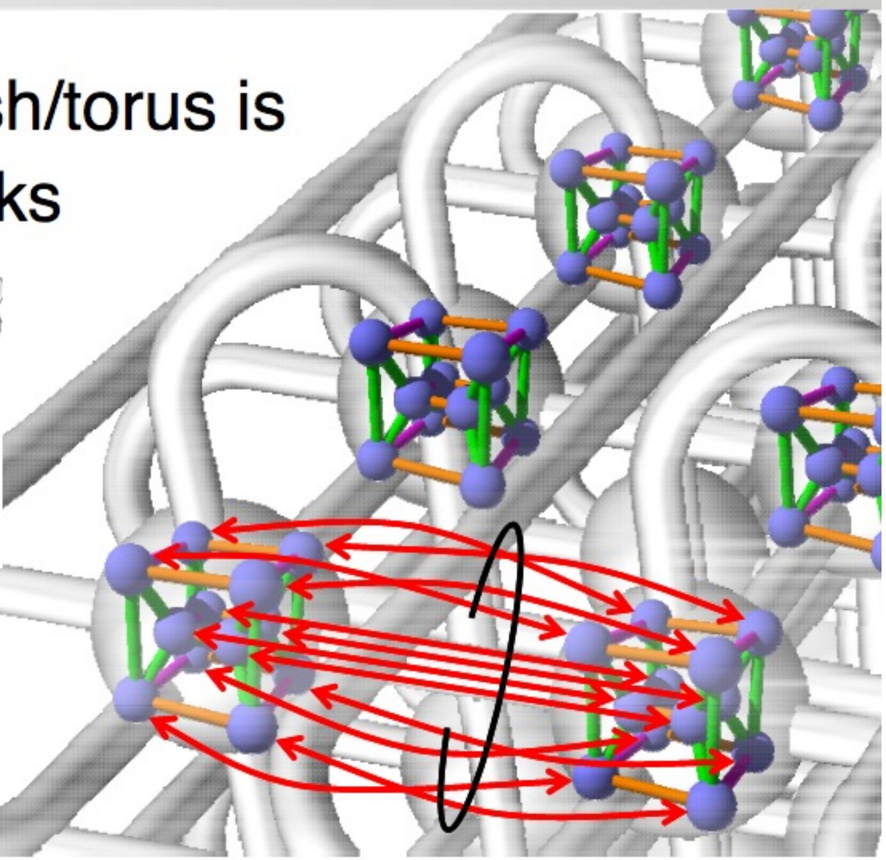
Post-K Activities, ISC19, Frankfurt 3



Tofu3: 6D mesh/torus interconnect

- Each pair of adjacent *abc* mesh/torus is interconnected with twelve links





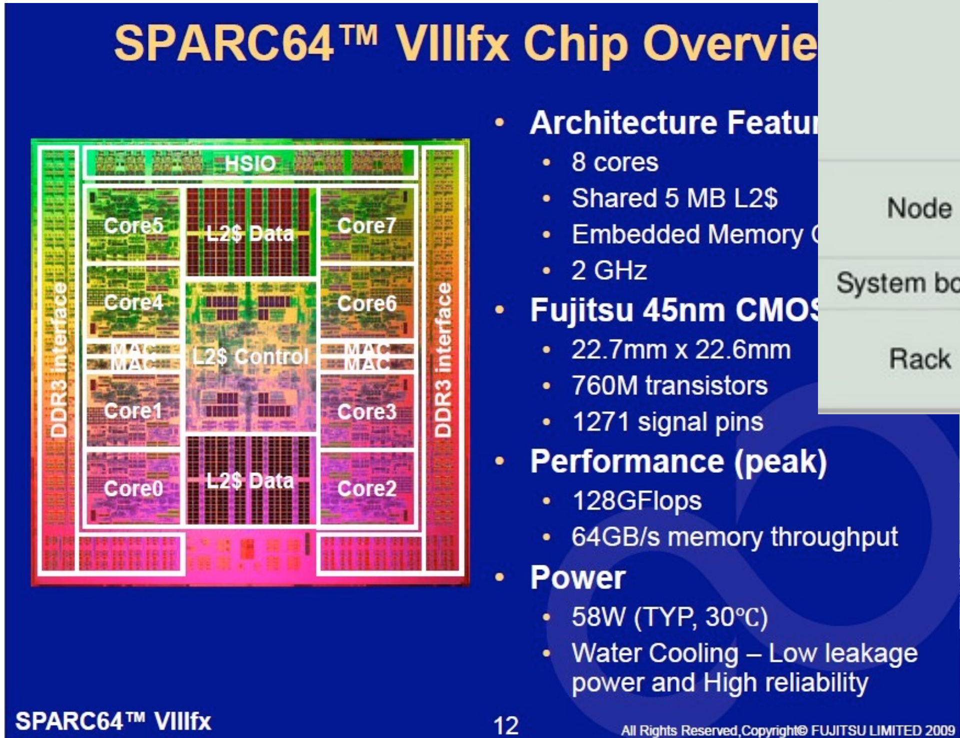
Fujitsu K computer

(the Japanese word "kei" (京) means 10 quadrillion, 10^{16})

		K computer
CPU	Name	SPARC64™ VIIIfx
	Performance	128GFlops@2GHz
	Architecture	SPARC V9 + HPC-ACE extension
	Cache configuration	L1(I) Cache:32KB/core, L1(D) Cache:32KB/core
		L2 Cache: 6MB(shared)
	No. of cores/socket	8
	Memory band width	64 GB/s.
Node	Configuration	1 CPU / Node
	Memory capacity	16 GB
System board	Node/system board	4 Nodes
Rack	System board/rack	24 System boards
	Performance/rack	12.3 TFlops

TOP 500 The List.

Jun'11: #1
Nov'11: #1
Jun'12: #2
Nov'12: #3
Jun'13: #4
Nov'13: #4
Jun'14: #4
Nov'14: #4
Jun'15: #4
Nov'15: #4
Jun'16: #5
Nov'16: #7
Jun'17: #8
Nov'17: #10
Jun'18: #16
Nov'18: #18
Jun'19: #20



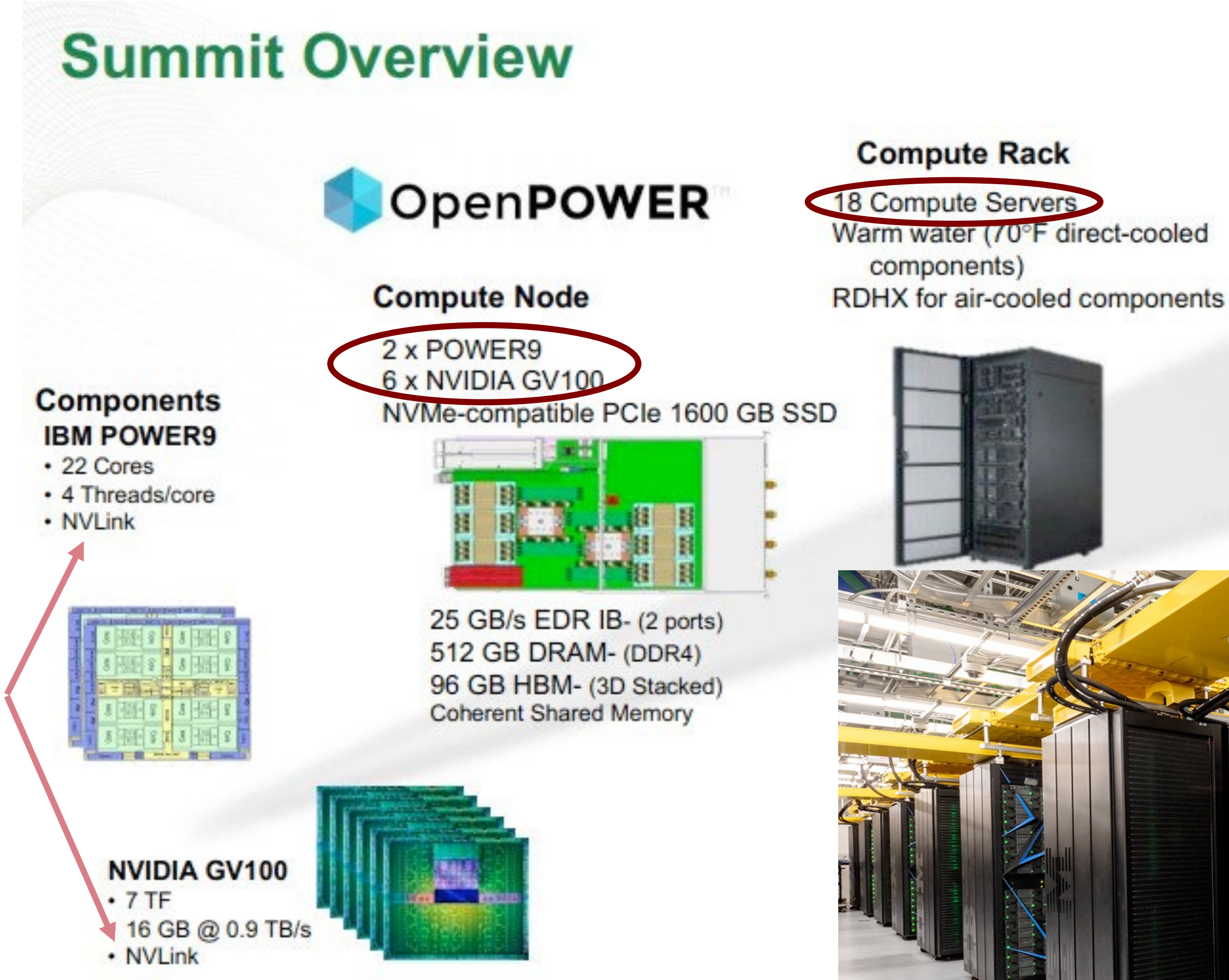
②

IBM POWER9 Summit

(Nov'19 #1 TOP500)

2 **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

since Jun'20



22-core IBM POWER9

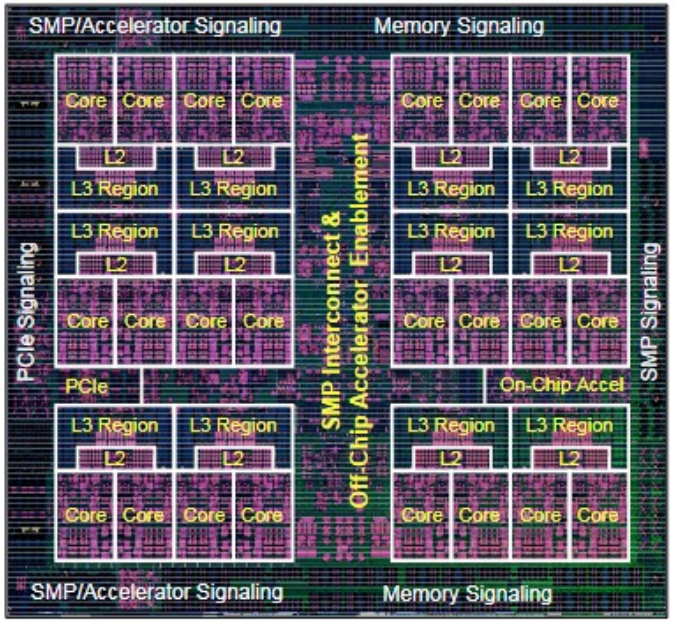
POWER9 Processor – Common Features

Power Systems IBM

New Core Microarchitecture
- Stronger thread performance
- Efficient agile pipeline
- POWER ISA v3.0

Enhanced Cache Hierarchy
- 120MB NUCA L3 architecture
- 12 x 20-way associative regions
- Advanced replacement policies
- Fed by 7 TB/s on-chip bandwidth

Cloud + Virtualization Innovation
- Quality of service assists
- New interrupt architecture
- Workload optimized frequency
- Hardware enforced trusted execution



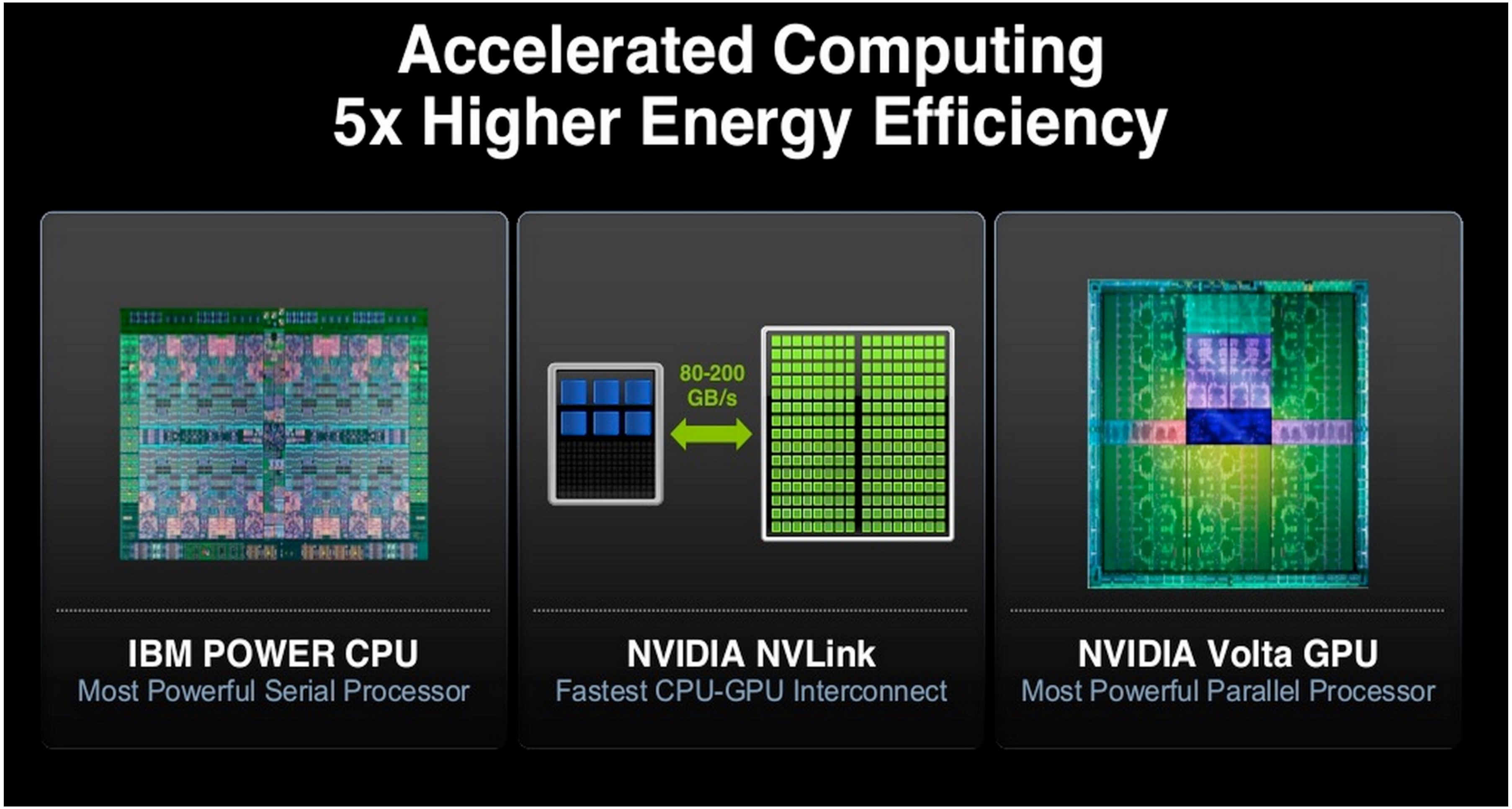
14nm finFET Semiconductor Process
- Improved device performance and reduced energy
- 17 layer metal stack and eDRAM
- 8.0 billion transistors

Leadership Hardware Acceleration Platform
- Enhanced on-chip acceleration
- Nvidia NVLink 2.0: High bandwidth and advanced new features (25G)
- CAPI 2.0: Coherent accelerator and storage attach (PCIe G4)
- New CAPI: Improved latency and bandwidth, open interface (25G)

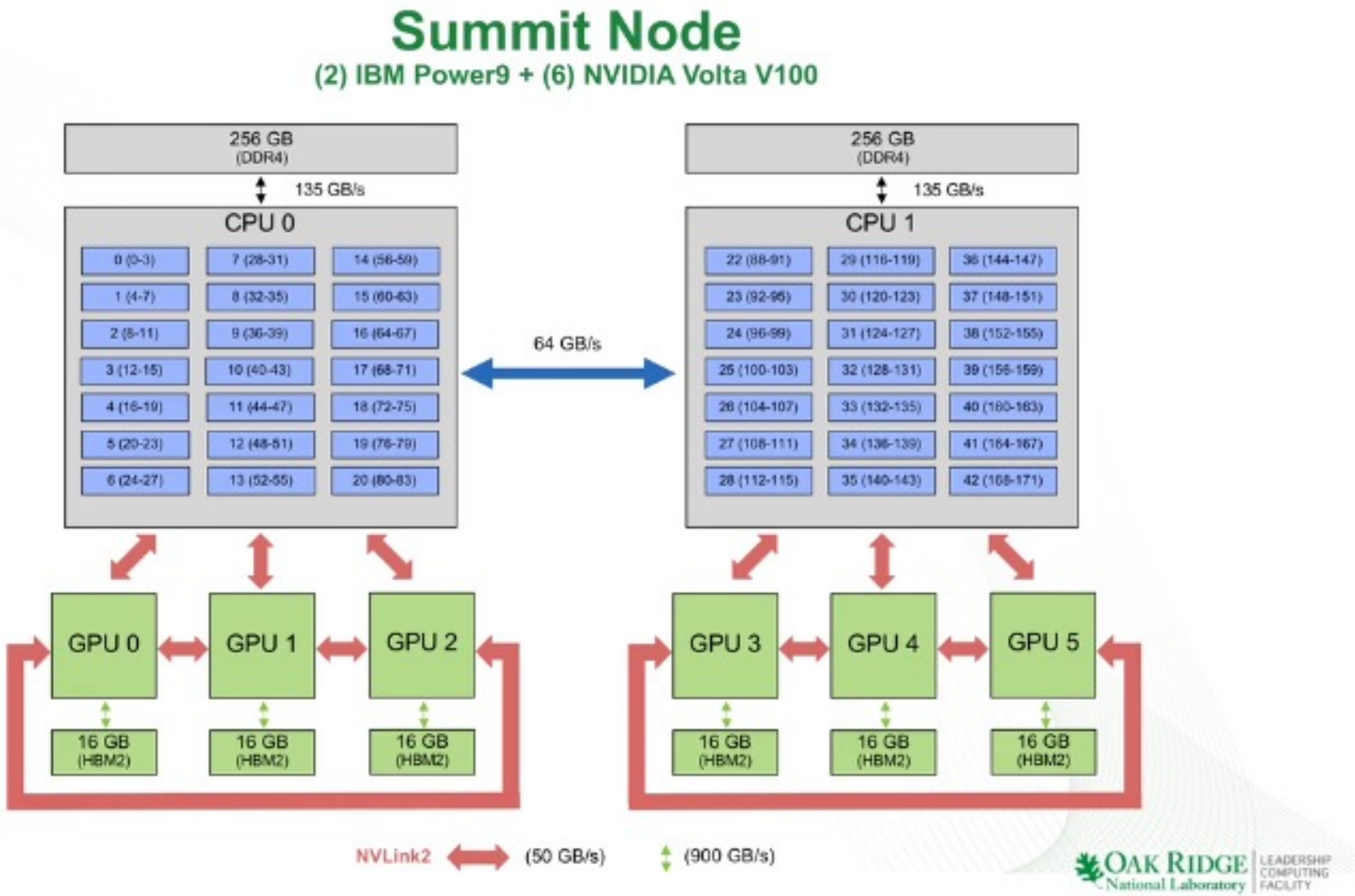
State of the Art I/O Subsystem
- PCIe Gen4 – 48 lanes

High Bandwidth Signaling Technology
- 16 Gb/s interface
  - Local SMP
- 25 Gb/s Common Link interface
  - Accelerator, remote SMP

IBM POWER9 + NVidia V100

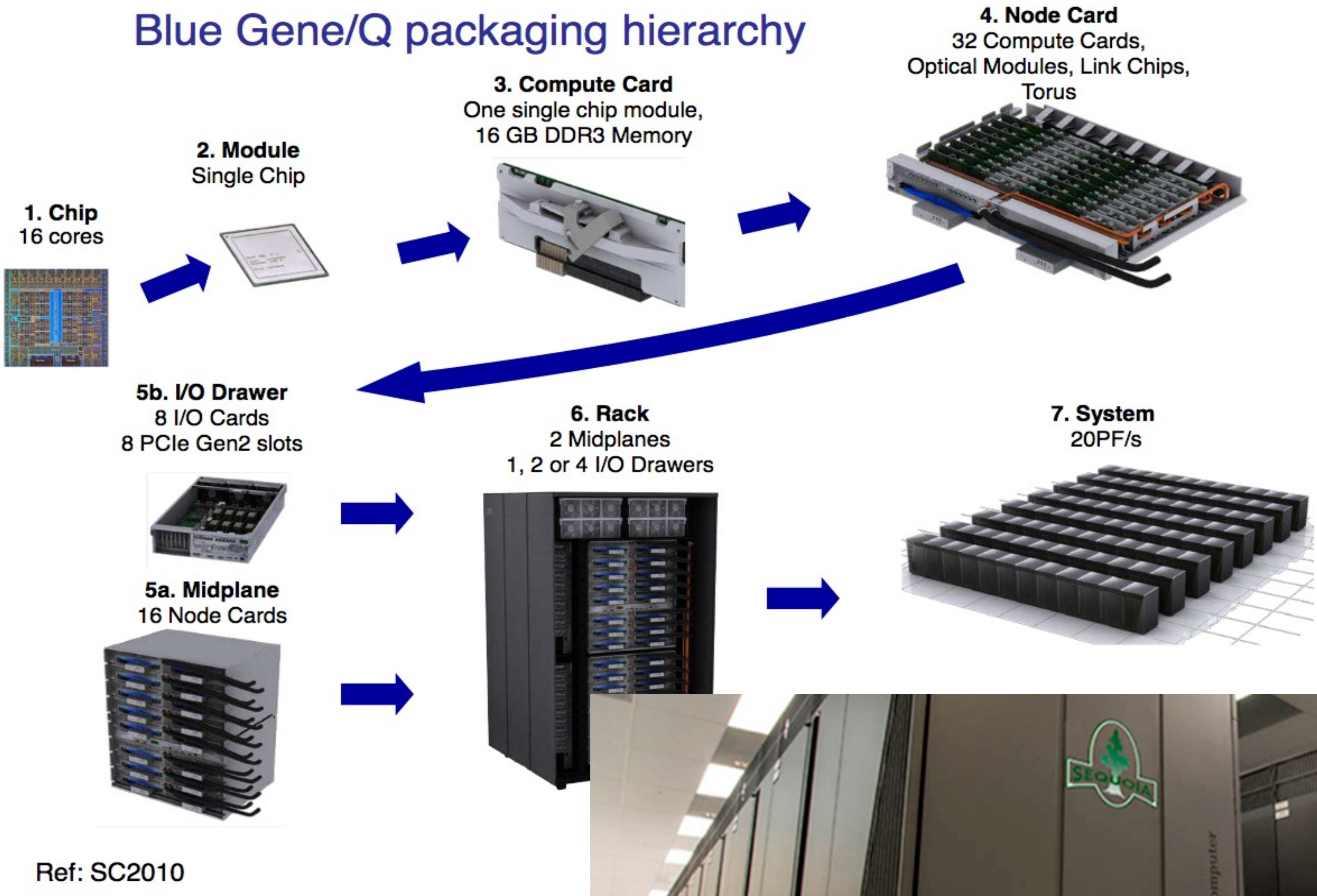


Summit/Sierra node architecture

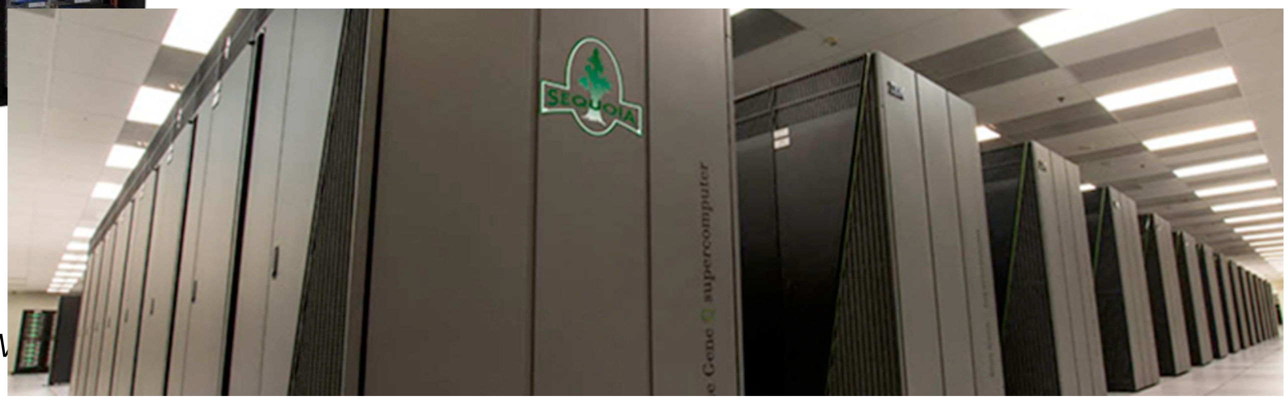


https://en.wikichip.org/wiki/supercomputers/summit

IBM Power BlueGene/Q Compute (Sequoia)



Jun’12: #1
Nov’12: #2
Jun’13: #3
Nov’13: #3
Jun’14: #3
Nov’14: #3
Jun’15: #3
Nov’15: #3
Jun’16: #4
Nov’16: #4
Jun’17: #5
Nov’17: #6
Jun’18: #8
Nov’18: #10
Nov’19: #12



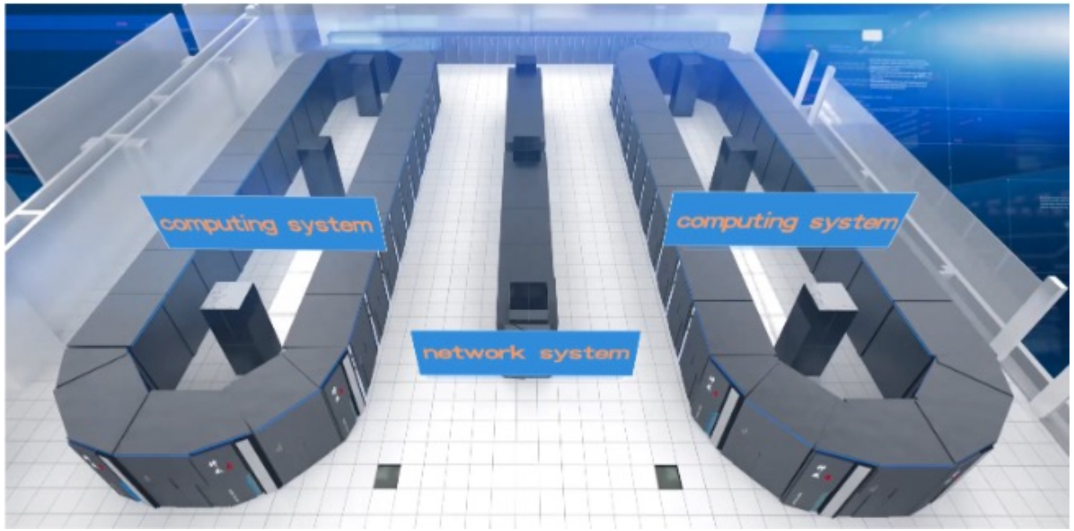
3

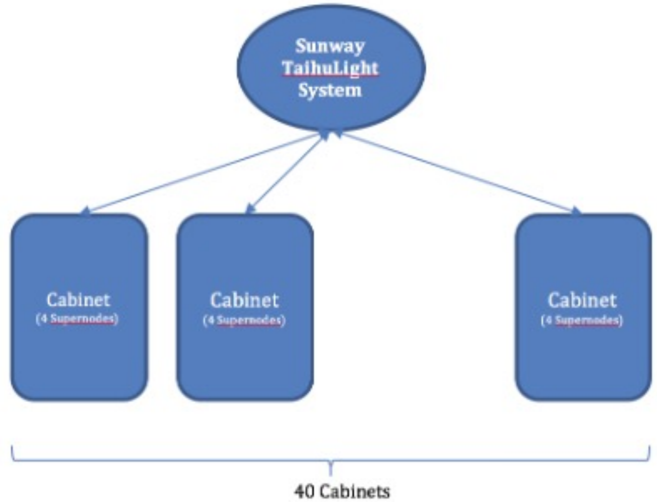
Sunway TaihuLight
(#1 in June'16 TOP500)

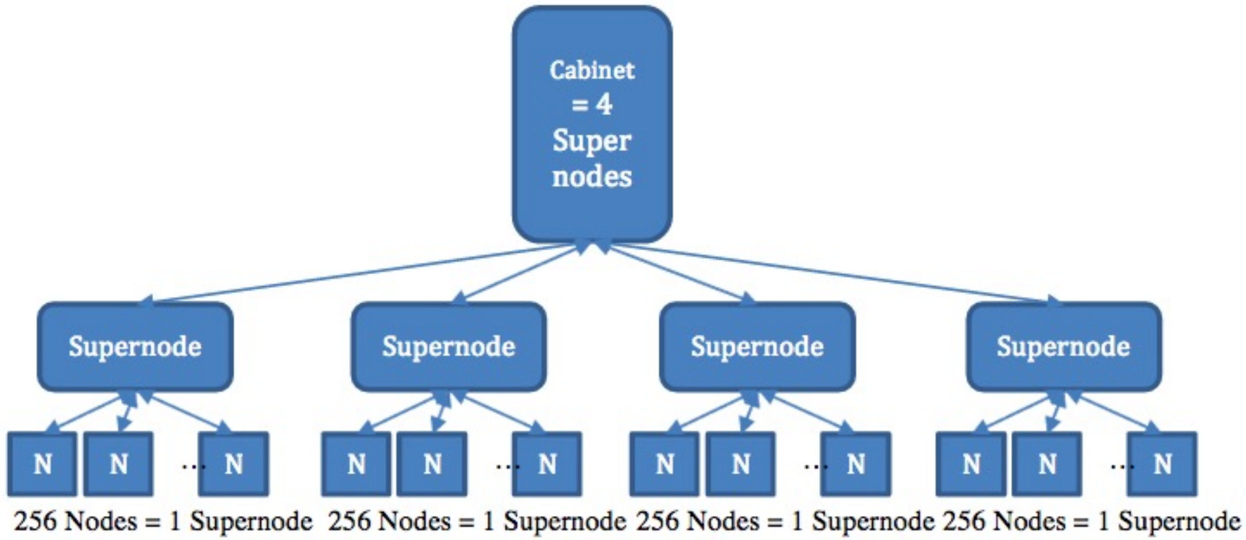
4 **Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC
National Supercomputing Center in Wuxi
China

since Jun'20

Overview of the Sunway TaihuLight System



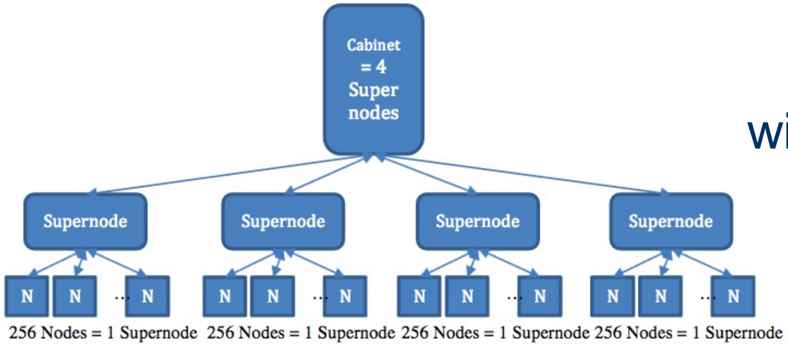




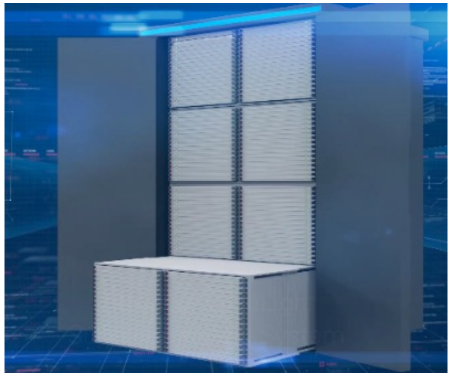
http://www.netlib.org/utk/people/JackDongarra/PAPERS/sunway-report-2016.pdf

Sunway TaihuLight

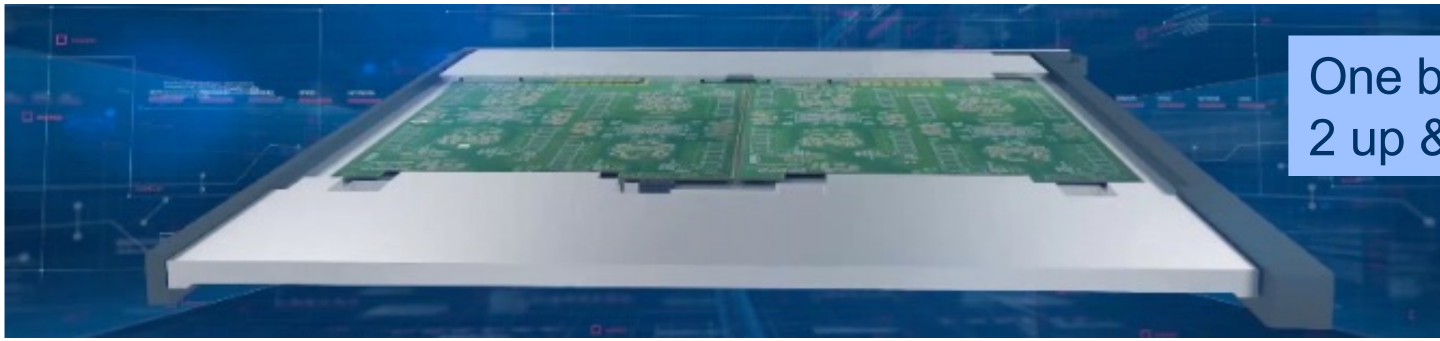
(#1 in June'16 TOP500)



One cabinet
with 4 Supernodes



One Supernode
with 32 boards



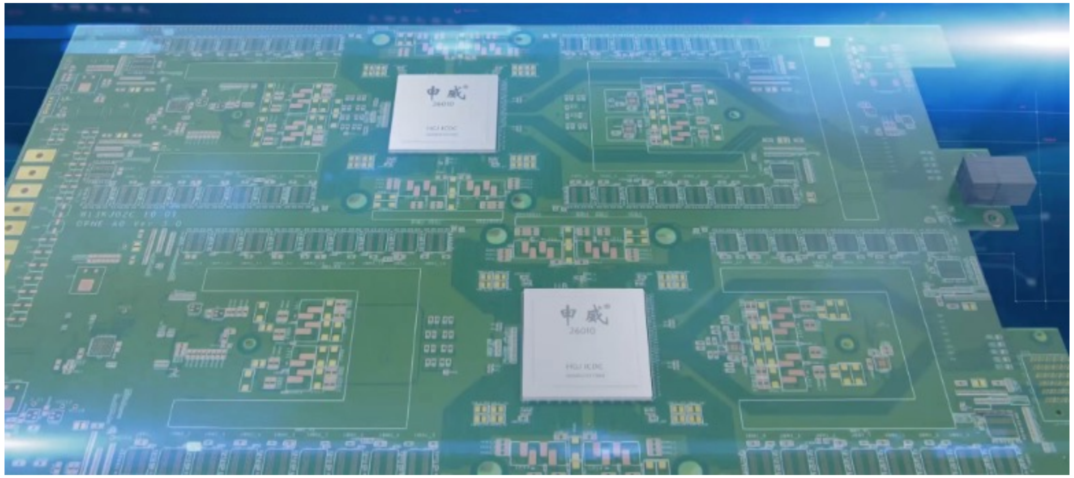
One board with 4 cards,
2 up & 2 down

http://www.netlib.org/utk/people/JackDongarra/PAPERS/sunway-report-2016.pdf

Sunway TaihuLight

(#1 in June'16 TOP500)

One card w/ two PU devices *(two SW26010 chips)*

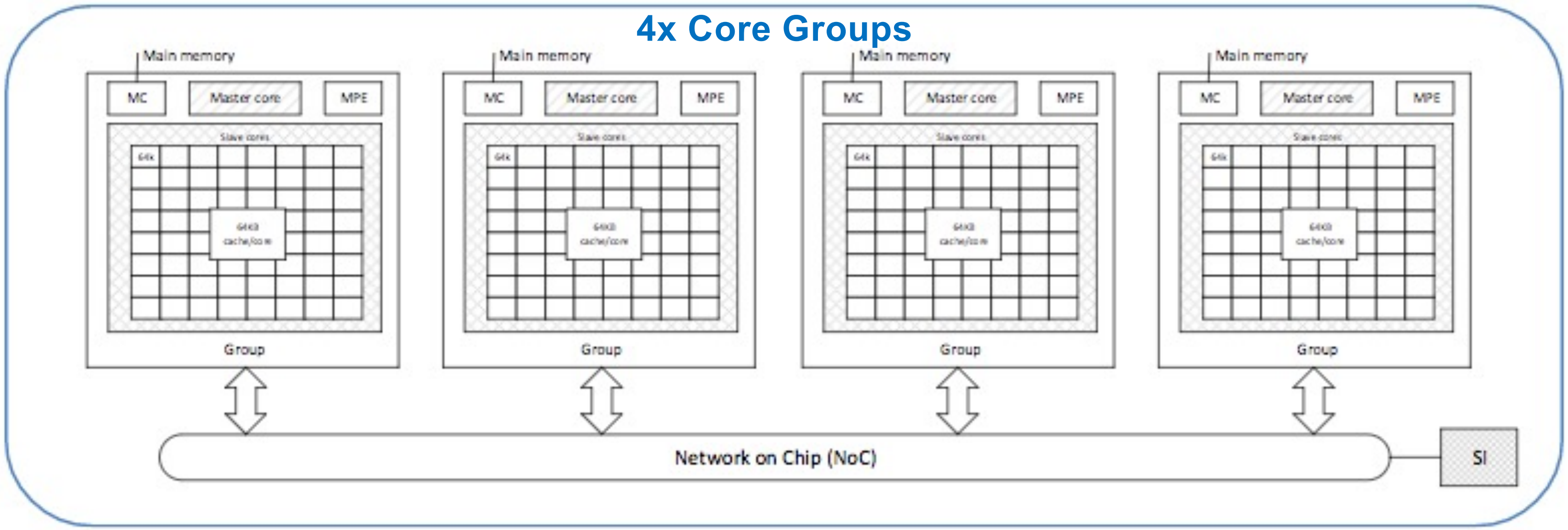


SW26010 chip with
4x NUMA Core Groups (CG).

Each CG follows a hybrid approach:

- 1 fat-core (MPE) for serial work, OoO execution, superscalar, L1 & L2
- 8x8 grid of skinny-cores (CPE), L1 private & L2 shared by the grid
- all cores are 64-bit RISC PU and all support 256-bit vector instructions

SW26010 chip



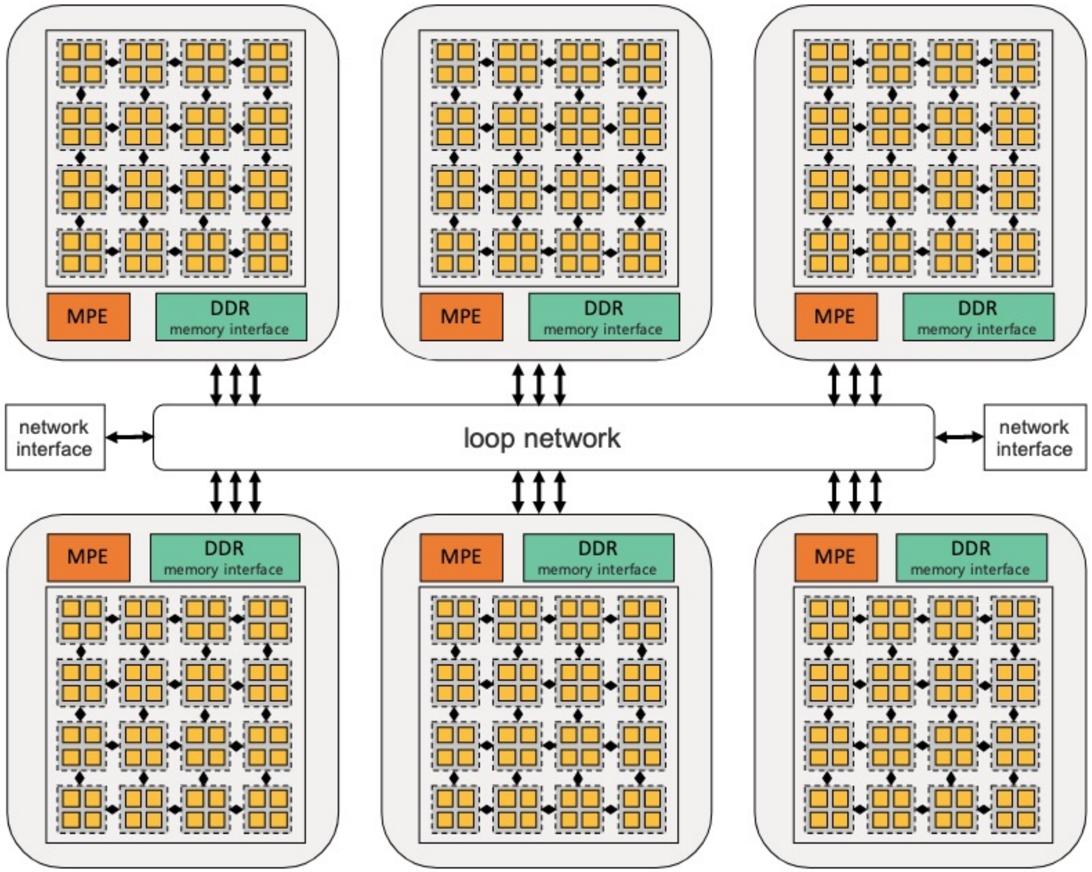
http://www.netlib.org/utk/people/JackDongarra/PAPERS/sunway-report-2016.pdf

The new Sunway OceanLight supercomputer
(April'21, not submitted to TOP500)

- Based on the chip **SW26010Pro**
  - 6x CG, each 1x fat-core & 4x4 mesh of 4 CPE
  - overall **390 cores**
    (= 6 x (1 + 64))
  - mem controller at each CG accesses 16 GiB DDR4 with 51.2 GiB/s bandwidth
  - each chip:
    accesses 96 GiB DDR4
    with 307.2 GiB/s bandwidth
- Single-socket nodes
- More nodes, more cores:
  from 7.6 millions cores to
  over **41.9 millions cores**!

In the original design: **8x** leading to 520 cores

6x core-groups (CG)



https://dl.acm.org/doi/pdf/10.1145/3458817.3487399

4

Overview of Tianhe-2A

(#1 in June'13 TOP500)

7 TOP 500 The List. Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China **Nov'21**

Overview of Tianhe-2A

Comparison

Items	Milkyway-2	Milkyway-2A
Nodes & Performance	16000 nodes with Intel CPU + KNC	**17792 nodes with Intel CPU + Matrix-2000**
	54.9Pflops	**94.97Pflops**
Interconnection	10Gbps, 1.57us	**14Gbps, 1us**
Memory	1.4PB	**3.4PB**
Storage	12.4PB, 512GB/s	**20PB, 1TB/s**
Energy Efficiency	17.8MW, 1.9Gflops/W	**About 18MW, >5Gflops/W**
Heterogeneous software	MPSS for Intel KNC	**OpenMP/OpenCL for Matrix-2000**

https://www.nextplatform.com/2018/02/14/look-whats-store-chinas-tianhe-2a-supercomputer/

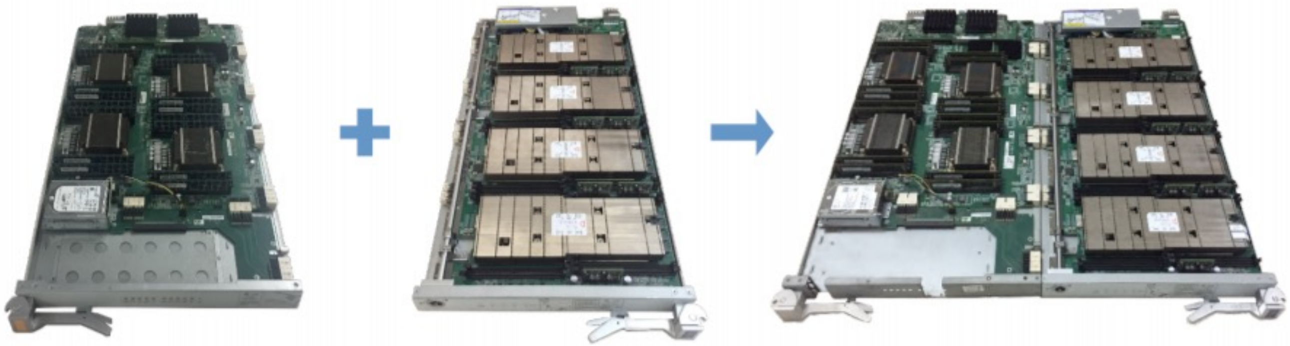
Overview of Tianhe-2A

NSCC GZ Compute nodes

Heterogeneous Compute Blades

– Compute blade = Xeon part + Matrix-2000 part

4 Intel Xeon CPUs + 4 FT Matrix-2000 → 2 Compute Nodes

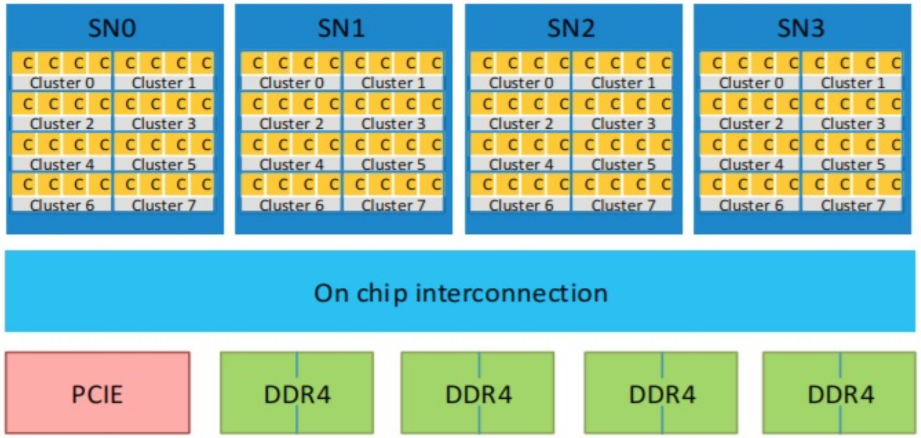


– Use the Matrix-2000 part to replace the KNC part

https://www.nextplatform.com/2018/02/14/look-whats-store-chinas-tianhe-2a-supercomputer/

Replacing the KNC in Tianhe-2A: the Matrix-2000 accelerator

Matrix-2000 accelerator



- **Chip specification**
  - 128cores
    - 4 super-nodes (SN)
    - 8 clusters per SN
    - 4 cores per cluster
    - Core
      - Self-defined 256-bit vector ISA
      - 16 DP flops/cycle per core
  - Peak performance: 2.4576Tflops@1.2GHz

    4 SNs x 8 clusters x 4cores x 16 flops x 1.2 GHz = 2.4576 Tflops
  - Peak power dissipation: ~240w
  - Interface
    - 8 DDR4-2400 channels
    - X16 PCIE 3.0 EP Port

15

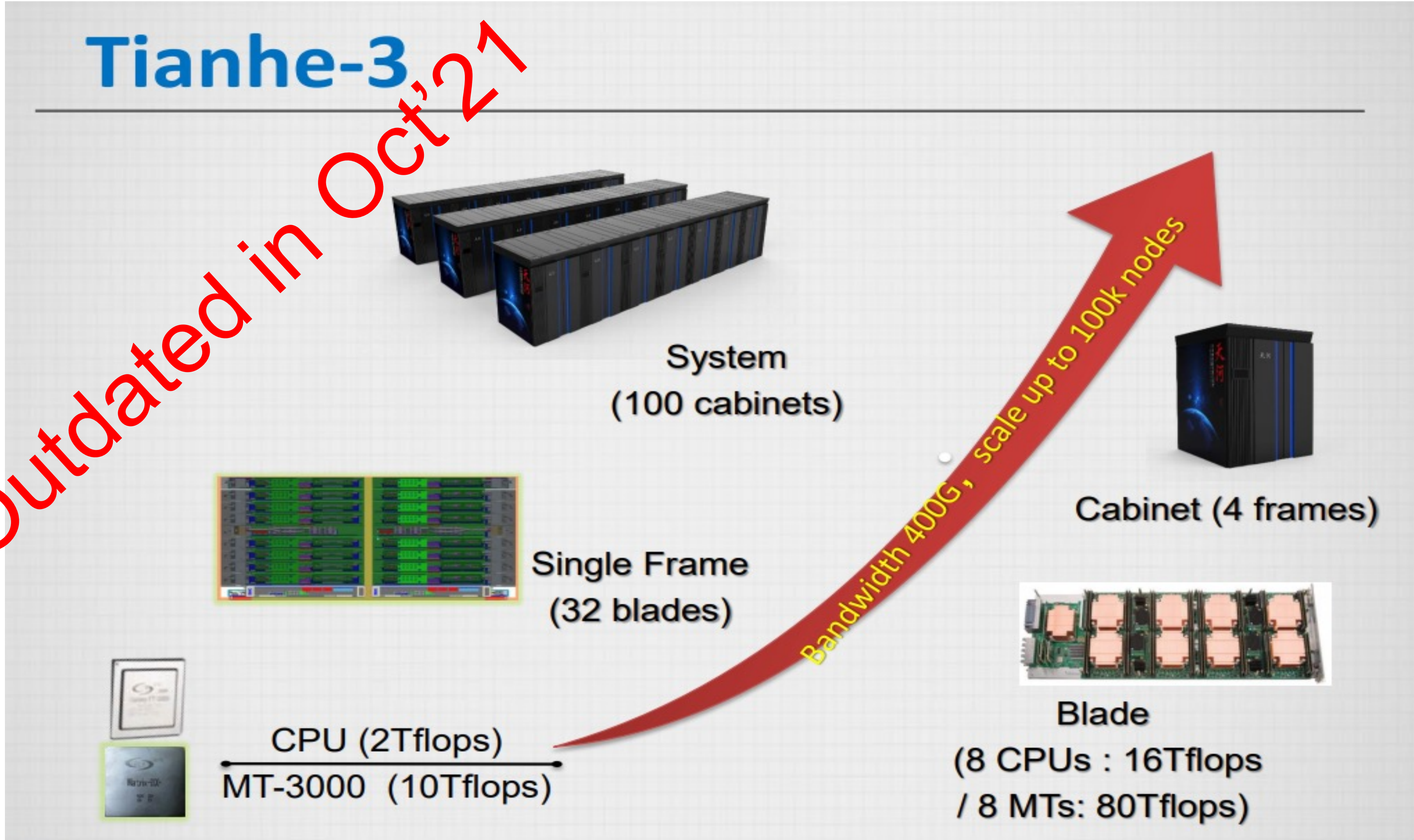
https://www.nextplatform.com/2018/02/14/look-whats-store-chinas-tianhe-2a-supercomputer/

Phytium-2000+ Next: Tianhe-3

with ~~Fujitsu A64FX~~ ARM~~-SVE~~ + Matrix-~~3000~~ accelerators
2000+

https://www.hpcwire.com/2021/11/24/three-chinese-exascale-systems-detailed-at-sc21-two-operational-and-one-delayed/

Outdated in Oct'21



https://www.nextplatform.com/2019/05/02/china-fleshes-out-exascale-design-for-tianhe-3/

5

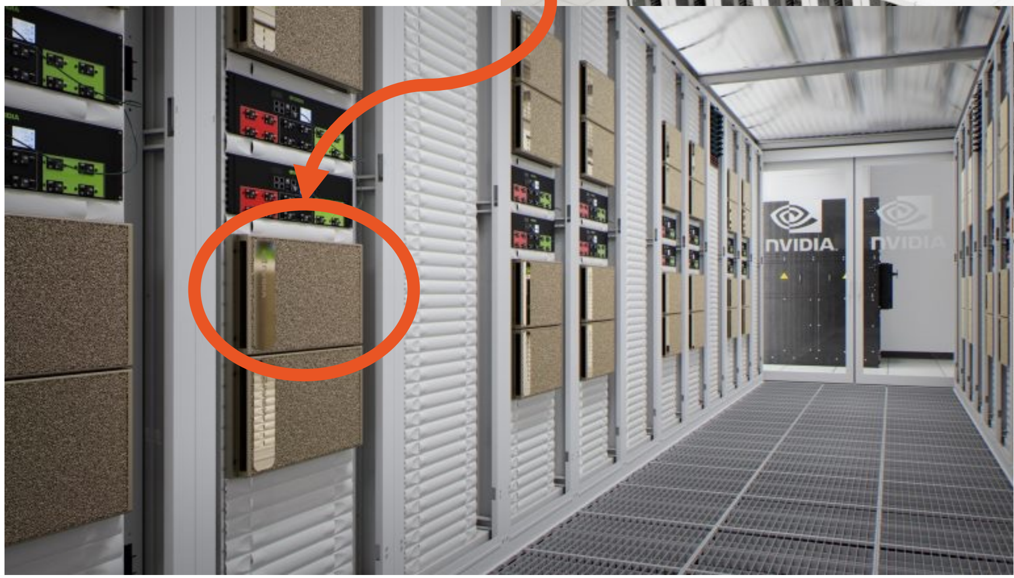
NVidia Selene: 280 DGX A100 nodes

6 | Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia
NVIDIA Corporation
United States

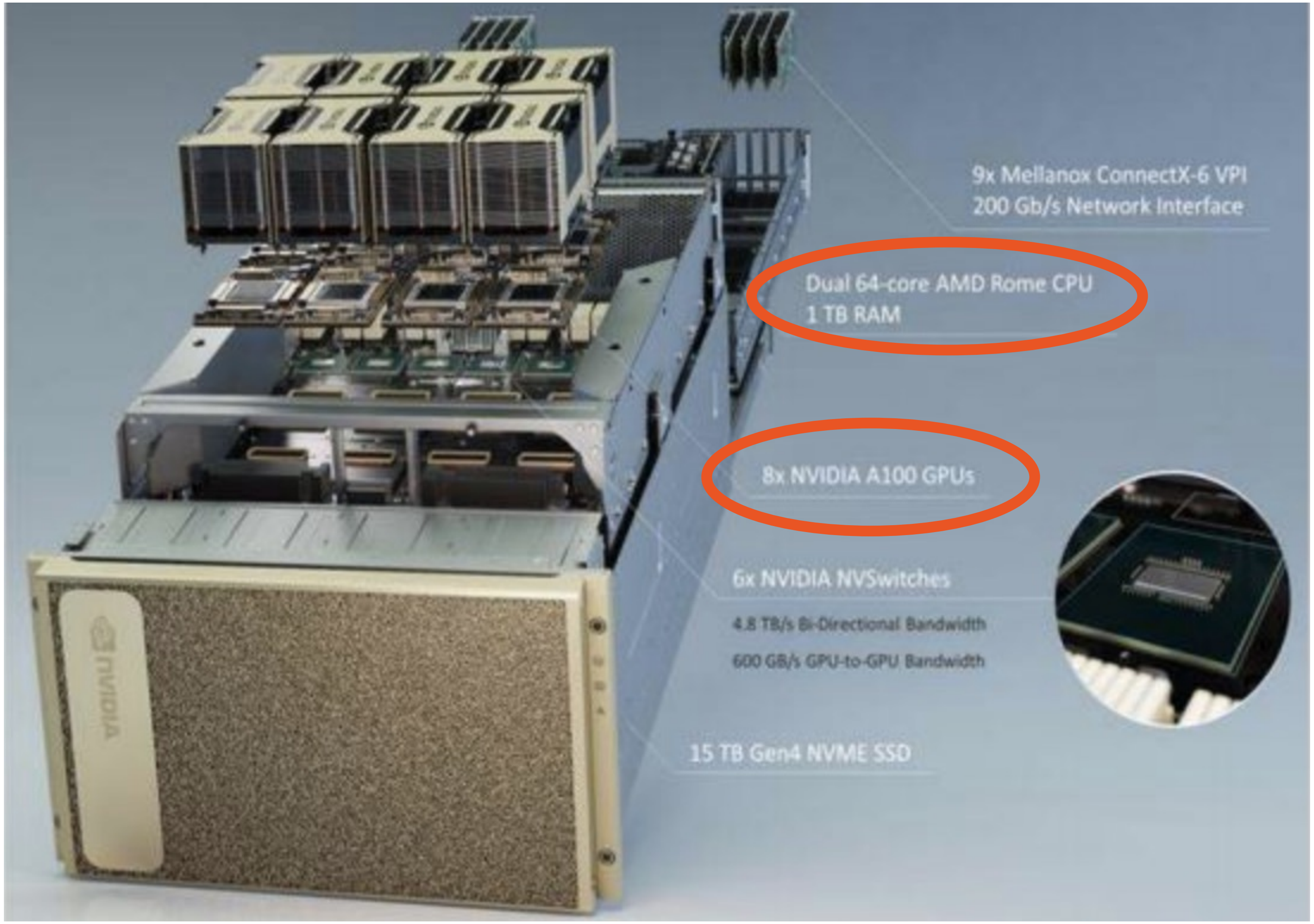
TOP 500 The List.

Nov'21

DGX A100 node



NVidia DGX A100 node



Suggestion of homework for discussion in next session

1. Go to the TOP500 website and analyse & comment:
   i. The country distribution over the past 25 years, in #systems and aggregate performance in the TOP500 list
   ii. The evolution of the key PU chip technologies and the accelerator families in the past 25 years
   iii. The overall impact of each processor technology and accelerator family in the past 3 years

2. EuroHPC is funding 8 supercomputing centres selected in June 2019: 3 pre-exascale & 5 petascale *(peak HPL performance)*
   i. Find & identify these 8 supercomputing centres
   ii. Characterize the architecture of Deucalion in MACC

Outline

1. TOP500

a) TOP10 lists from Nov'17 to Nov'21
b) Country distribution over the past 25 years
c) PU chip technology evolution in the past 25 years and since last year
d) Evolution of the accelerators since they were available
e) Analysis of some relevant systems and architectures

2. GREEN500

a) TOP10 lists from Nov'17 to Nov'21
b) Analysis of some relevant systems and architectures

3. HPCG500

a) HPCG *vs.* HPL: an overview
b) TOP10 lists from Nov'17 to Nov'20
c) Analysis of some relevant systems

4. HPL-AI

a) High-performance Linpack (HPL) and Artificial Intelligence (AI) workloads

The GREEN 500

The Green500 list

The list ranks computers in terms of energy efficiency, typically measured as LINPACK FLOPS per watt.

About the Green500 List

The Green500 list ranks the top 500 supercomputers in the world by energy efficiency. The focus of performance-at-any-cost computer operations has led to the emergence of supercomputers that consume vast amounts of electrical power and produce so much heat that large cooling facilities must be constructed to ensure proper performance. To address this trend, the Green500 list puts a premium on energy-efficient performance for sustainable supercomputing.

The inaugural Green500 list was announced on November 15, 2007 at SC|07. As a complement to the TOP500, the unveiling of the Green500 ushered in a new era where supercomputers can be compared by performance-per-watt.

While the selection of any power-performance metric will be controversial, we currently opt for "FLOPS-per-Watt" given that it has already become a widely used metric in the community and for

Top Green500 systems Nov'19

The GREEN 500

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	159	**A64FX prototype** - Fujitsu A64FX, Fujitsu A64FX 48C 2GHz, Tofu interconnect D , Fujitsu Fujitsu Numazu Plant Japan	36,864	1,999.5	118	16.876
2	420	**NA-1** - ZettaScaler-2.2, Xeon D-1571 16C 1.3GHz, Infiniband EDR, PEZY-SC2 700Mhz , PEZY Computing / Exascaler Inc. PEZY Computing K.K. Japan	1,271,040	1,303.2	80	16.256
3	24	**AiMOS** - IBM Power System AC922, IBM POWER9 20C 3.45GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Volta GV100 , IBM Rensselaer Polytechnic Institute Center for Computational Innovations (CCI) United States	130,000	8,045.0	510	15.771
4	373	**Satori** - IBM Power System AC922, IBM POWER9 20C 2.4GHz, Infiniband EDR, NVIDIA Tesla V100 SXM2 , IBM MIT/MGHPCC Holyoke, MA United States	23,040	1,464.0	94	15.574
5	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	10,096	14.719
6	8	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	1,649	14.423
7	494	**MareNostrum P9 CTE** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Tesla V100 , IBM Barcelona Supercomputing Center Spain	18,360			
8	23	**TSUBAME3.0** - SGI ICE XA, IP139-SXM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 SXM2 , HPE GSIC Center, Tokyo Institute of Technology Japan	135,828	8,125.0	792	13.704
9	11	**PANGEA III** - IBM Power System AC922, IBM POWER9 18C 3.45GHz, Dual-rail Mellanox EDR Infiniband, NVIDIA Volta GV100 , IBM Total Exploration Production France	291,024	17,860.0	1,367	13.065
10	2	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	7,438	12.723
11	48	**Advanced Computing System(PreE)** - Sugon TC8600, Hygon Dhyana 32C 2GHz, Deep Computing Processor, 200Gb 6D-Torus , Sugon Sugon China	163,840	4,325.0	380	11.382

~Ghyoukou

Top Green500 systems Nov'20

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	170	**NVIDIA DGX SuperPOD** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	19,840	2,356.0	90	26.195
2	330	**MN-3** - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan	1,664	1,652.9	65	26.039
3	7	**JUWELS Booster Module** - Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	1,764	25.008
4	146	**Spartan2** - Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR Infiniband, Atos Atos France	23,040	2,566.0	106	24.262
5	5	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	2,646	23.983
6	239	**A64FX prototype** - Fujitsu A64FX, Fujitsu A64FX 48C 2GHz, Tofu interconnect D, Fujitsu	36,864	1,999.5	118	16.876

~Selene

#1 in Jun'20

~Fugaku

The GREEN 500

Top Green500 systems Nov’21

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Power (kW)	Power Efficiency (GFlops/watts)
1	301	**MN-3** - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan	1,664	2,181.2	55	39.379
2	291	**SSC-21 Scalable Module** - Apollo 6500 Gen10 plus, AMD EPYC 7543 32C 2.8GHz, NVIDIA A100 80GB, Infiniband HDR200, HPE Samsung Electronics South Korea	16,704	2,274.1	103	33.983
3	295	**Tethys** - NVIDIA DGX A100 Liquid Cooled Prototype, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 80GB, Infiniband HDR, Nvidia NVIDIA Corporation United States	19,840	2,255.0	72	31.538
4	280	**Wilkes-3** - PowerEdge XE8545, AMD EPYC 7763 64C 2.45GHz, NVIDIA A100 80GB, Infiniband HDR200 dual rail, DELL EMC University of Cambridge United Kingdom	26,880	2,287.0	74	30.797
5	30	**HiPerGator AI** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Infiniband HDR, Nvidia University of Florida United States	138,880	17,200.0	583	29.521
6	403	**Snellius Phase 1 GPU** - ThinkSystem SD650-N V2, Xeon Platinum 8360Y 36C 2.4GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR, Lenovo SURF Netherlands	6,480	1,818.0	63	29.046

#1 in Jun’20

All based on EPYC + A100

Xeon + A100

Additional comments:

- from #7 to #11, #13 to# 21, … all based on EPYC/Xeon + A100
- #12: NA-IT1, follow-up of NA-1 with PEZY-SC3

12	435	**NA-IT1** - ZettaScaler3.0, AMD EPYC 7702P 64C 1.5GHz, PEZY-SC3, Infiniband EDR, PEZY Computing / Exascaler Inc. NA Simulation Japan

Preferred Networks

The MN-3 system

#1 in Jun'20 Green500

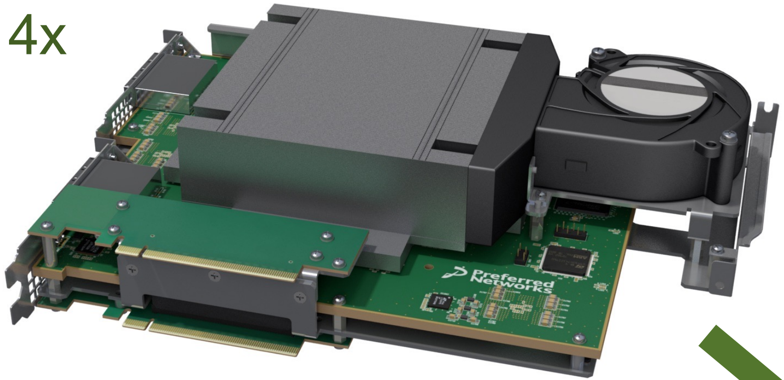
2 330 The GREEN 500 MN-3 - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan

Nov'20

2x

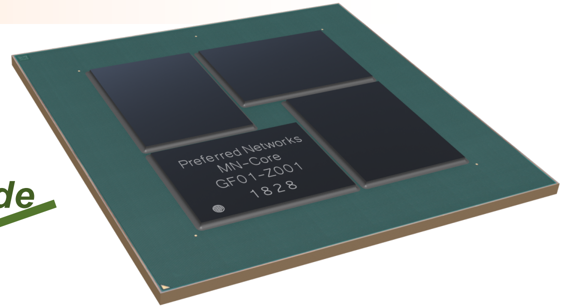
Xeon 8260M 24c

\+ 4x



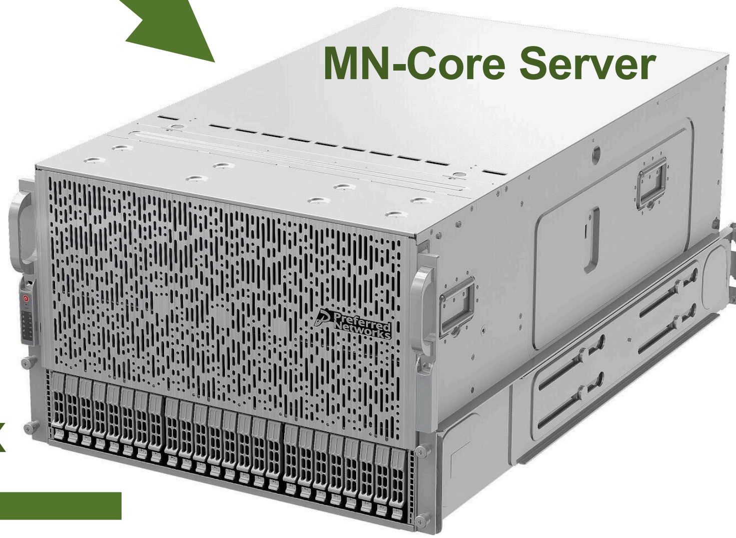
MN-Core Board

inside

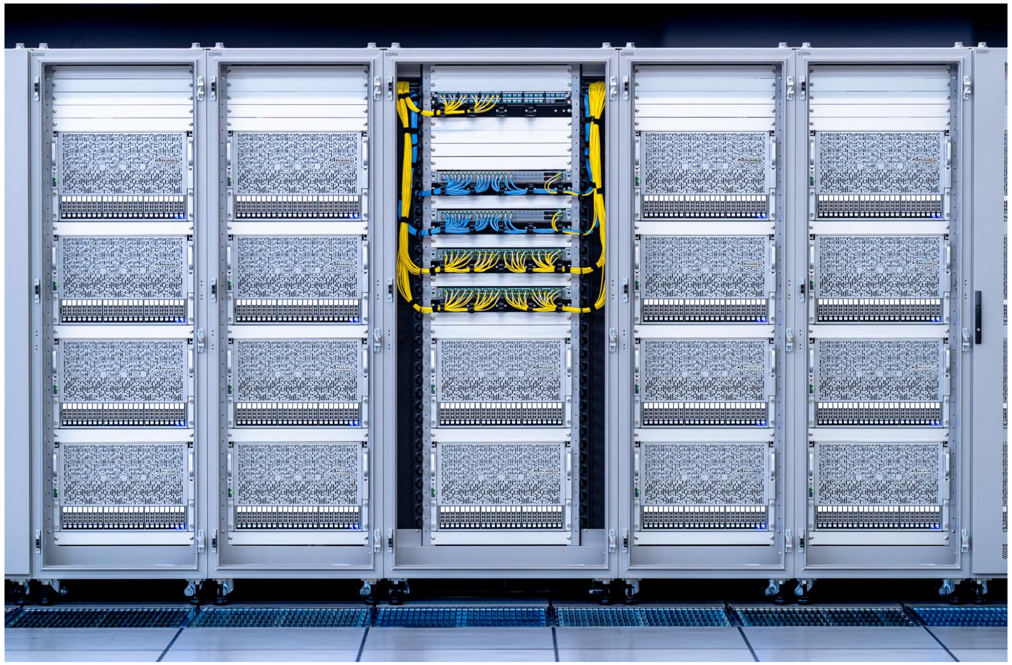


MN-Core accelerator

MN-Core Server



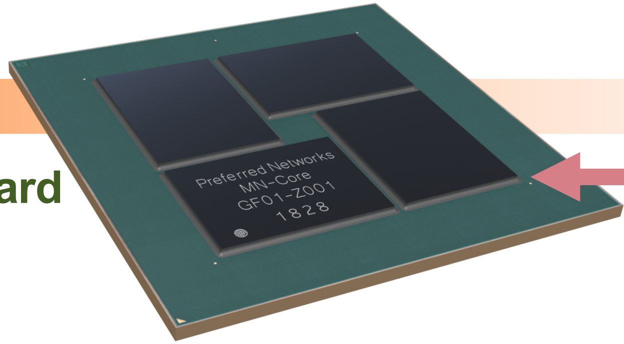
40x



MN-3 *(partial view)*

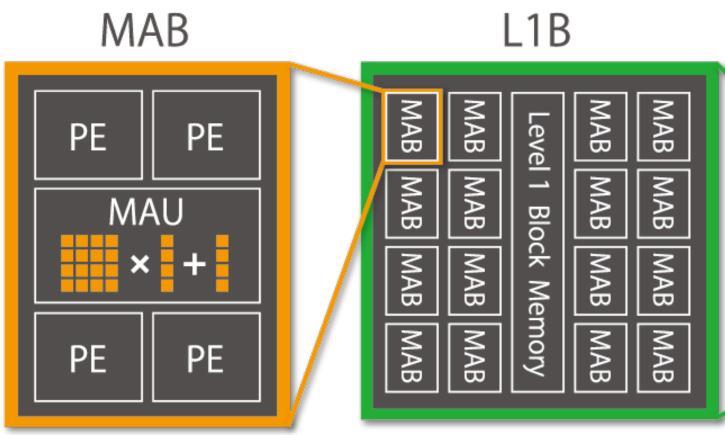
https://projects.preferred.jp/mn-core/en/

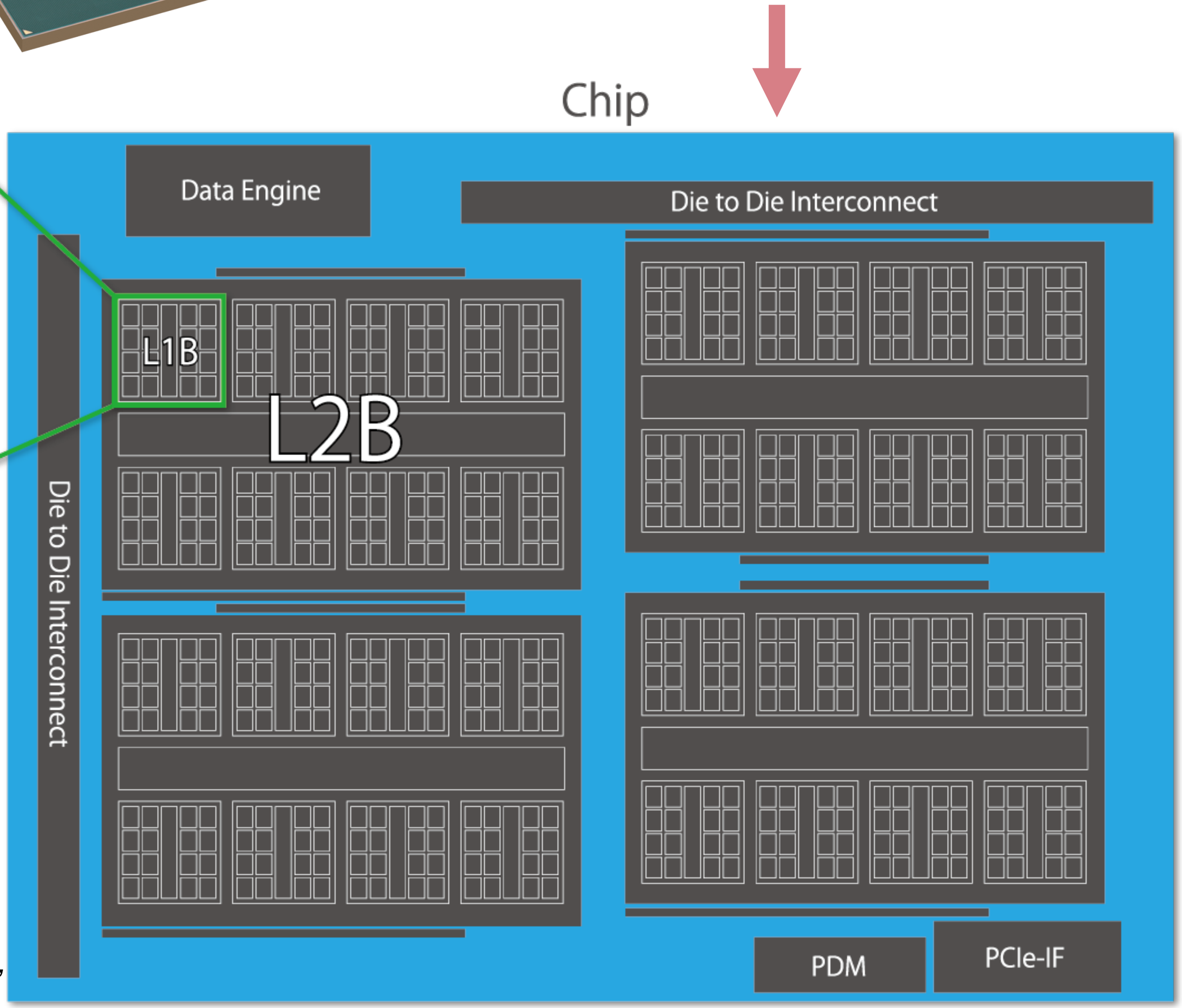
MN-Core architecture



MN-Core Board

4 dies
each die

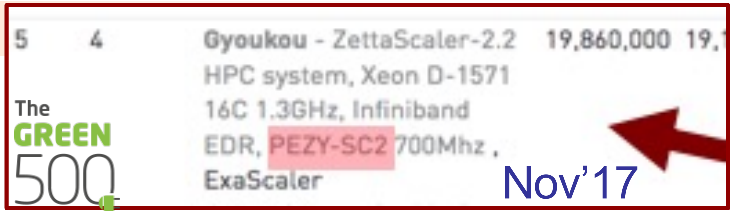




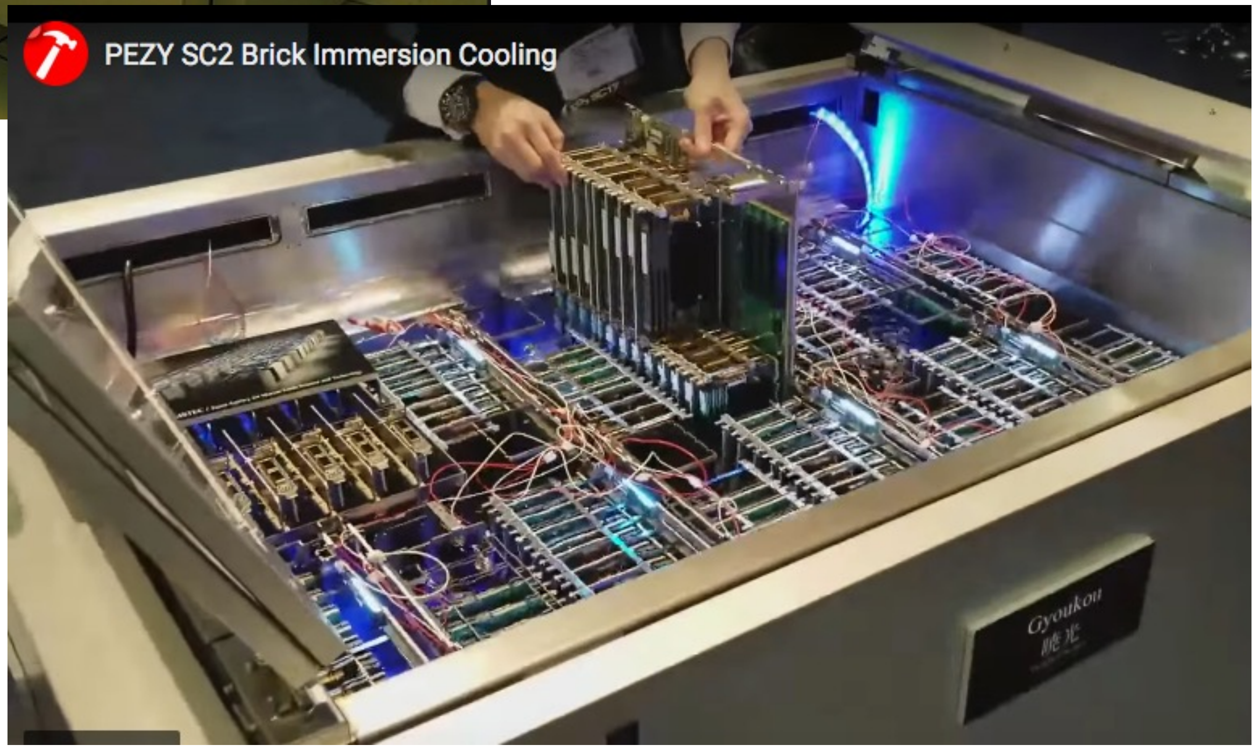
MAU: Matrix Arithmetic Unit
PE: Processing Element
MAB: Matrix Arithmetic Block
L1B: Level 1 Block w/ 16 MABs
L2B: Level 2 Block w/ 128 MABs
1 chip/die with 512 MABs
1 MN-Core package w/ 2048 MABs

Gyoukou ZettaScaler-2.2

10,000 PEZY-SC2 + 1,250 16-cores Xeon =
19.84 M PEZY cores + 20 K Xeon cores

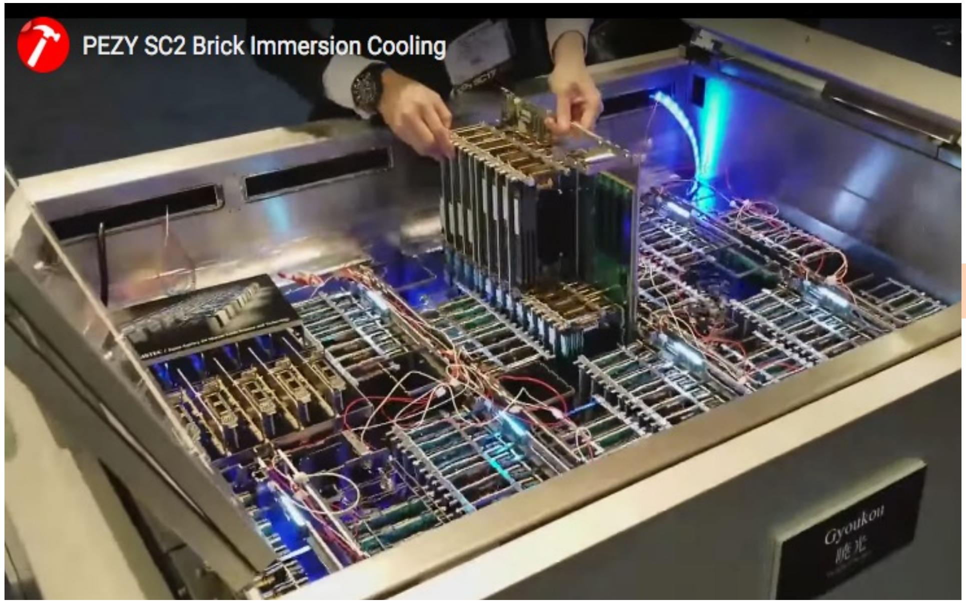


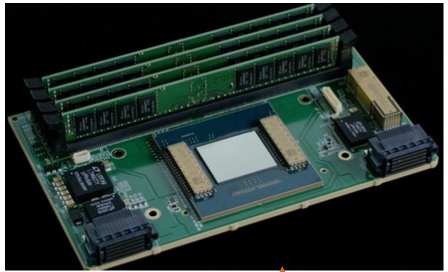
NA-1, #3 in Jun’20 is similar to Gyoukou, but w/less cores

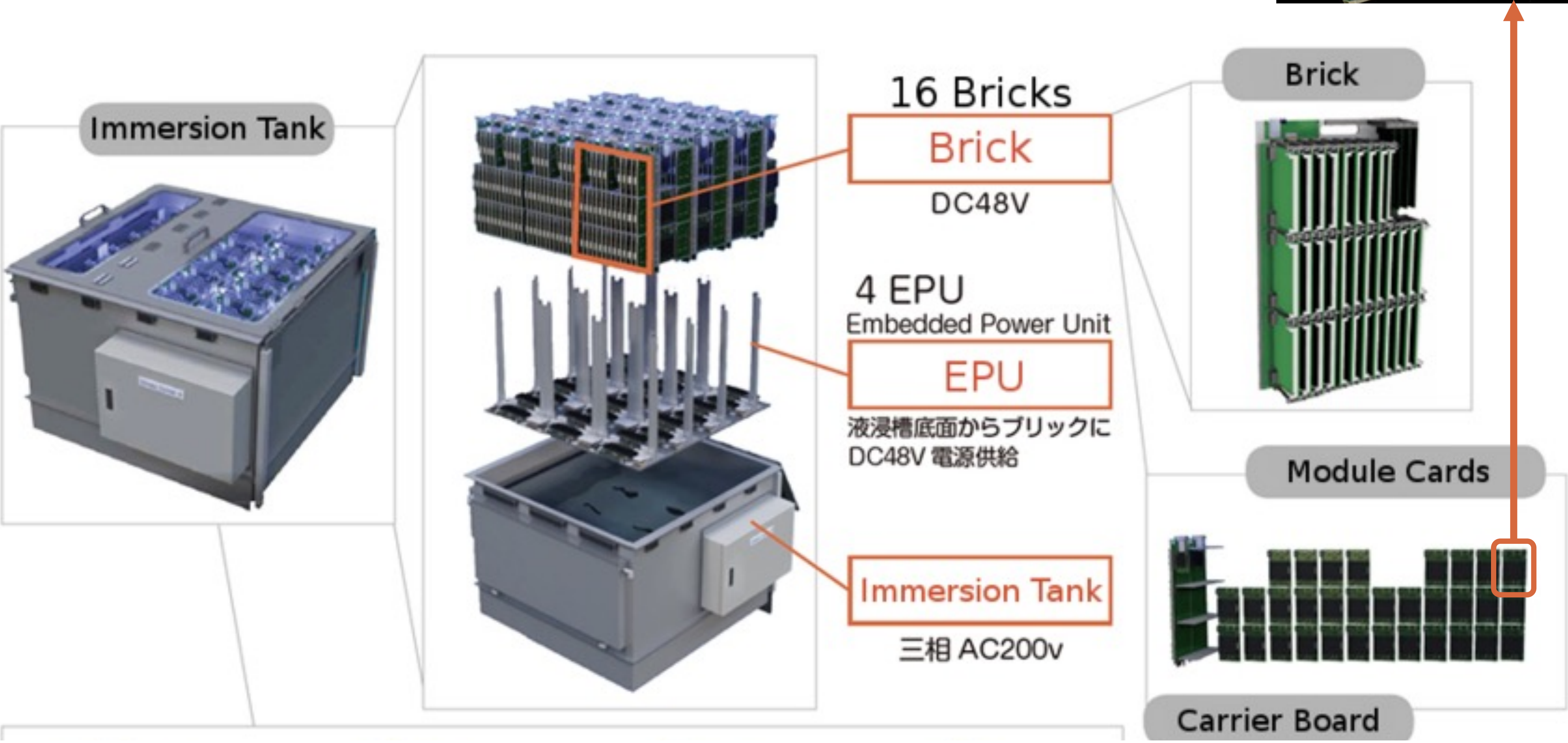


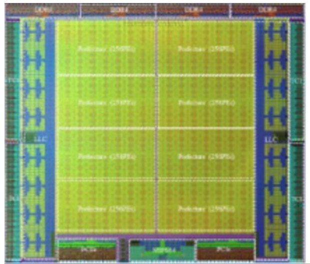
20 immersion tanks
each tank 16 bricks
each brick 32 PEZY
each PEZY
~2K 8-way SMT cores
=>
each tank ~1M cores

Gyoukou
ZettaScaler-2.2

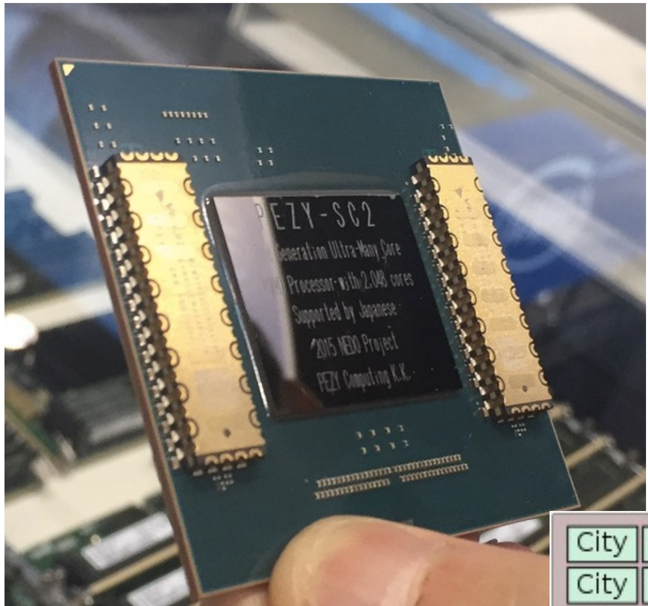








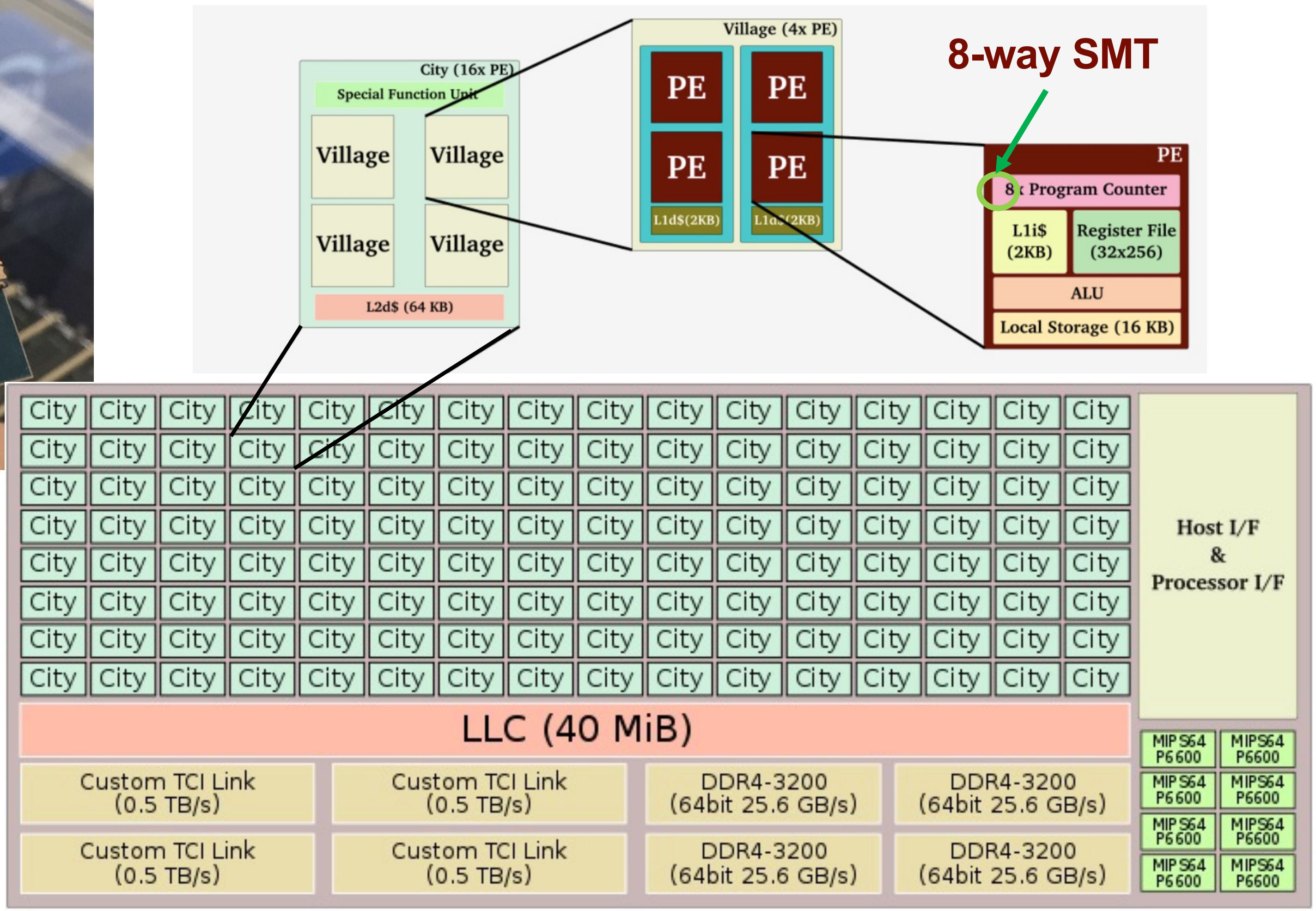
PEZY-SC2 in ZettaScaler-2.2

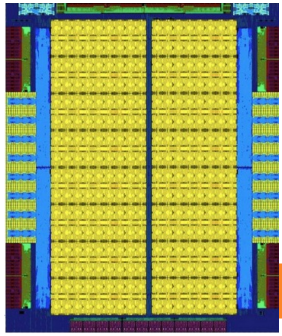


PEZY-SC2: **2 048 cores**
+ 8x MIPS cores *(2017)*

PEZY-SC3: 8 192 cores
(due 2019, arrived 2021…)

PEZY-SC4: 16 384 cores
(due in 2020, but…)





ZettaScaler-3.0 with PEZY-SC3: estimation in 2019

1 tank => 40 nodes x **[1 AMD Epyc** (64-core) + **4 PEZY-SC3** (8192-core)**]**
1 tank has **1 312 320 cores**

Estimated ZettaScaler-3.0 Specs

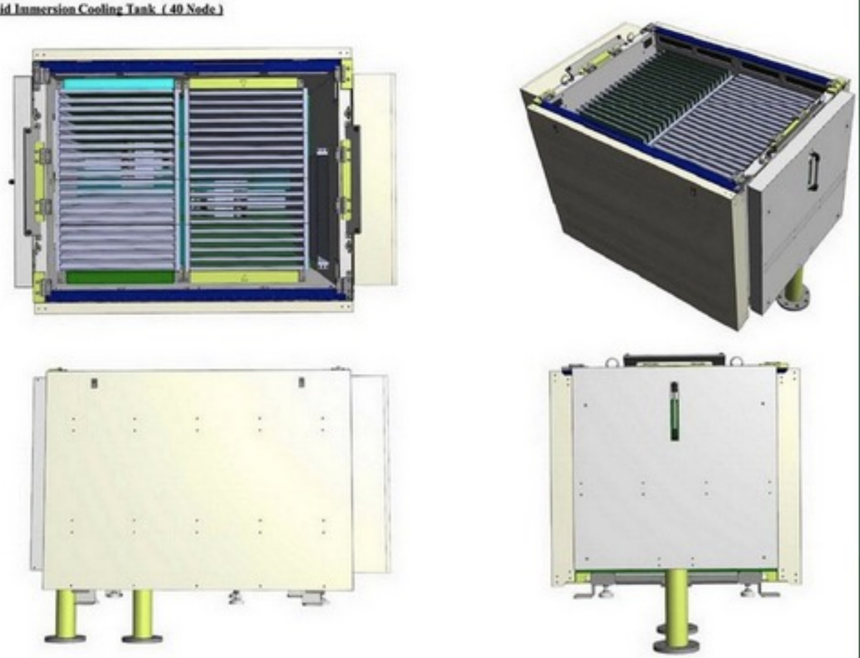
tanks configuration will provide about 100 PetaFLOPS (Rmax) and only consumes 4MW with the system cost of around $100M

ZettaScaler-3.0 single tank will have 40 nodes, 40 AMD EPYC2 (64 core), 160 of PEZY-SC3 with 48DC power

Single tank will have 3.2 PetaFLOPS (Rpeak) and 2.4 PetaFLOPS (Rmax) of DP performance

System power efficiency will be 30 GFLOPS/W or so and single tank requires 100kW range power

Liquid Immersion Cooling Tank (40 Node)



Outline

1. TOP500
   a) TOP10 lists from Nov'17 to Nov'21
   b) Country distribution over the past 25 years
   c) PU chip technology evolution in the past 25 years and since last year
   d) Evolution of the accelerators since they were available
   e) Analysis of some relevant systems and architectures

2. GREEN500
   a) TOP10 lists from Nov'17 to Nov'21
   b) Analysis of some relevant systems and architectures

3. HPCG500
   a) HPCG *vs.* HPL: an overview
   b) TOP10 lists from Nov'17 to Nov'20
   c) Analysis of some relevant systems

4. HPL-AI
   a) High-performance Linpack (HPL) and Artificial Intelligence (AI) workloads

HPCG is a self-contained C++ program with MPI and OpenMP support that measures the performance of basic operations in a unified code:

- Sparse matrix-vector multiplication
- Vector updates
- Global dot products
- Local symmetric Gauss-Seidel smoother
- Sparse triangular solve (part of Gauss-Seidel smoother)

HPCG benchmark

From Wikipedia, the free encyclopedia

The **HPCG** (**high performance conjugate gradient**) **benchmark** is a supercomputing benchmark test proposed by Michael Heroux from Sandia National Laboratories, and Jack Dongarra and Piotr Luszczek from the University of Tennessee.[1][2] It is intended to model the data access patterns of real-world applications such as sparse matrix calculations, thus testing the effect of limitations of the memory subsystem and internal interconnect of the supercomputer on its computing performance.[3] Because it is internally I/O bound, HPCG testing generally achieves only a tiny fraction of the peak FLOPS of the computer.[4]

HPCG is intended to complement benchmarks such as the LINPACK benchmarks that put relatively little stress on the internal interconnect.[5] The source of the HPCG benchmark is available on GitHub.[6]

TOP500: HPCG vs. HPL

HPCG Benchmark

The High Performance Conjugate Gradients (HPCG) Benchmark project is an effort to create a new metric for ranking HPC systems. HPCG is intended as a complement to the High Performance LINPACK (HPL) benchmark, currently used to rank the TOP500 computing systems. The computational and data access patterns of HPL are still representative of some important scalable applications, but not all. HPCG is designed to exercise computational and data access patterns that more closely match a different and broad set of important applications, and to give incentive to computer system designers to invest in capabilities that will have impact on the collective performance of these applications.

HPCG is a complete, stand-alone code that measures the performance of basic operations in a unified code:

- Sparse matrix-vector multiplication.
- Vector updates.
- Global dot products.
- Local symmetric Gauss-Seidel smoother.
- Sparse triangular solve (as part of the Gauss-Seidel smoother).
- Driven by multigrid preconditioned conjugate gradient algorithm that exercises the key kernels on a nested set of coarse grids.
- Reference implementation is written in C++ with MPI and OpenMP support.

http://www.hpcg-benchmark.org

Top 10 HPCG systems Nov’17

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	HPCG (TFlop/s)
1	10	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	11,280.4	602.736
2	2	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P , NUDT National Super Computer Center in Guangzhou China	3,120,000	33,862.7	54,902.4	580.109
3	7	**Trinity** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	979,968	14,137.3	43,902.6	546.124
4	3	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	361,760	19,590.0	25,326.3	486.398
5	1	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	480.8
6	9	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	24,913.5	385.479
7	8	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	27,880.7	355.442
8	6	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom , IBM DOE/NNSA/LLNL United States	1,572,864	17,173.2	20,132.7	330.373
9	5	**Titan** - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x , Cray Inc. DOE/SC/Oak Ridge National Laboratory United States	560,640	17,590.0	27,112.5	322.322
10	13	**TSUBAME3.0** - SGI ICE XA, IP139-SXM2, Xeon E5-2680v4 14C 2.4GHz, Intel Omni-Path, NVIDIA Tesla P100 SXM2 , HPE GSIC Center, Tokyo Institute of Technology Japan	135,828	8,125.0	12,127.1	188.6

HPCG

Top 10 systems Nov'18

HPCG List for November 2018

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	HPCG (TFlop/s)
1	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	2,397,824	143,500.0	2925.75
2	2	**Sierra** - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	1795.67
3	18	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect , Fujitsu RIKEN Advanced Institute for Computational Science (AICS) Japan	705,024	10,510.0	602.74
4	6	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/NNSA/LANL/SNL United States	979,072	20,158.7	546.12
5	7	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	508.85
6	5	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray Inc. Swiss National Supercomputing Centre (CSCS) Switzerland	387,872	21,230.0	496.98
7	3	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	480.85
8	13	**Nurion** - Cray CS500, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Cray Inc. Korea Institute of Science and Technology Information Korea, South	570,020	13,929.3	391.45
9	14	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	385.48
10	12	**Cori** - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray Inc. DOE/SC/LBNL/NERSC United States	622,336	14,014.7	355.44

HPCG

Top systems Nov'19

HPCG List for November 2019

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	HPCG (TFlop/s)
1	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
2	2	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband , IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	1795.67
3	7	**Trinity** - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect , Cray/HPE DOE/NNSA/LANL/SNL United States	979,072	20,158.7	546.12
4	8	**AI Bridging Cloud Infrastructure (ABCI)** - PRIMERGY CX2570 M4, Xeon Gold 6148 20C 2.4GHz, NVIDIA Tesla V100 SXM2, Infiniband EDR , Fujitsu National Institute of Advanced Industrial Science and Technology (AIST) Japan	391,680	19,880.0	508.85
5	6	**Piz Daint** - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100 , Cray/HPE Swiss National Supercomputing Centre (CSCS) Switzerland	387,872	21,230.0	496.98
6	3	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway , NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	480.85
7	14	**Nurion** - Cray CS500, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Cray/HPE Korea Institute of Science and Technology Information Korea, South	570,020	13,929.3	391.45
8	15	**Oakforest-PACS** - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path , Fujitsu Joint Center for Advanced High Performance Computing Japan	556,104	13,554.6	385.48
26	198	**Astra** - Apollo 70, Marvell ThunderX2 ARM CN9975-2000 28C 2GHz, 4xEDR Infiniband , HPE Sandia National Laboratories United States	143,640	1,833.0	90.90

HPCG

Top systems Nov'20

Rank	TOP500 Rank	System	Cores	Rmax (TFlop/s)	HPCG (TFlop/s)
1	1	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442,010.0	16004.50
2	2	**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
3	3	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	1795.67
4	5	**Selene** - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	555,520	63,460.0	1622.51
5	7	**JUWELS Booster Module** - Bull Sequana XH2000 , AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite, Atos Forschungszentrum Juelich (FZJ) Germany	449,280	44,120.0	1275.36
6	10	**Dammam-7** - Cray CS-Storm, Xeon Gold 6248 20C 2.5GHz, NVIDIA Tesla V100 SXM2, InfiniBand HDR 100, HPE Saudi Aramco Saudi Arabia			
7	8	**HPC5** - PowerEdg NVIDIA Tesla V100			
8	19	**TOKI-SORA** - PRIMEHPC FX1000, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu Japan Aerospace eXploration Agency Japan			
10	33	**Plasma Simulator** - SX-Aurora TSUBASA A412-8, Vector Engine Type10AE 8C 1.58GHz, Infiniband HDR 200, NEC National Institute for Fusion Science (NIFS) Japan	34,560	7,892.7	529.16

5.5x

Outline

1. TOP500
    a) TOP10 lists from Nov’17 to Nov’21
    b) Country distribution over the past 25 years
    c) PU chip technology evolution in the past 25 years and since last year
    d) Evolution of the accelerators since they were available
    e) Analysis of some relevant systems and architectures

2. GREEN500
    a) TOP10 lists from Nov’17 to Nov’21
    b) Analysis of some relevant systems and architectures

3. HPCG500
    a) HPCG *vs.* HPL: an overview
    b) TOP10 lists from Nov’17 to Nov’20
    c) Analysis of some relevant systems

4. HPL-AI
    a) High-performance Linpack (HPL) and Artificial Intelligence (AI) workloads

HPL-AI Mixed-Precision Benchmark

HPL-AI

The HPL-AI benchmark seeks to highlight the emerging convergence of high-performance computing (HPC) and artificial intelligence (AI) workloads. While traditional HPC focused on simulation runs for modeling phenomena in physics, chemistry, biology, and so on, the mathematical models that drive these computations require, for the most part, 64-bit accuracy. On the other hand, the machine learning methods that fuel advances in AI achieve desired results at 32-bit and even lower floating-point precision formats. This lesser demand for accuracy fueled a resurgence of interest in new hardware platforms that deliver a mix of unprecedented performance levels and energy savings to achieve the classification and recognition fidelity afforded by higher-accuracy formats.

HPL-AI strives to unite these two realms by delivering a blend of modern algorithms and contemporary hardware while simultaneously connecting to the solver formulation of the decades-old HPL framework of benchmarking the largest supercomputing installations in the world. The solver method of choice is a combination of LU factorization and iterative refinement performed afterwards to bring the solution back to 64-bit accuracy. The innovation of HPL-AI lies in dropping the requirement of 64-bit computation throughout the entire solution process and instead opting for low-precision (likely 16-bit) accuracy for LU, and a sophisticated iteration to recover the accuracy lost in factorization. The iterative method guaranteed to be numerically stable is the generalized minimal residual method (GMRES), which uses application of the L and U factors to serve as a preconditioner. The combination of these algorithms is demonstrably sufficient for high accuracy and may be implemented in a way that takes advantage of the current and upcoming devices for accelerating AI workloads.

https://icl.bitbucket.io/hpl-ai/

HPL-AI

HPL-AI List (Nov'21)

November 2021

Rank	Site	Computer	Cores	HPL-AI (Eflop/s)	TOP500 Rank
1	RIKEN	Fugaku	7,630,848	2.000	1
2	DOE/SC/ORNL	Summit	2,414,592	1.411	2
3	NVIDIA	Selene	555,520	0.630	6
4	DOE/SC/LBNL	Perlmutter	761,856	0.590	5
5	FZJ	JUWELS BM	449,280	0.470	8
6	University of Florida	HiPerGator	138,880	0.170	31
7	SberCloud	Christofari Neo	98,208	0.123	44

https://icl.bitbucket.io/hpl-ai/

1. TOP500

a) TOP10 lists from Nov'17 to Nov'21
b) Country distribution over the past 25 years
c) PU chip technology evolution in the past 25 years and since last year
d) Evolution of the accelerators since they were available
e) Analysis of some relevant systems and architectures

2. GREEN500

a) TOP10 lists from Nov'17 to Nov'21
b) Analysis of some relevant systems

3. HPCG500

a) HPCG *vs.* HPL: an overview
b) TOP10 lists from Nov'17 to Nov'20
c) Analysis of some relevant systems

4. HPL-AI

a) High-performance Linpack (HPL) and Artificial Intelligence (AI) workloads