Master Informatics Eng.

20120/21

A.J.Proença

Top HPC systems in TOP500 lists

(most images are borrowed)

WIKIPEDIA
The Free Encyclopedia

What is TOP500?

TOP500

From Wikipedia, the free encyclopedia
(Redirected from Top500)

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 56th published in November 2020. 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 212 supercomputers, leading the second place (United States) by a record margin of 113. Ranked by performance the most powerful supercomputers are located in the US (669 petaFLOPS). Followed by Japan (594 petaFLOPS) and China (564 petaFLOPS).

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.

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/

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

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]

Outline

1. TOP500

a) TOP10 lists from Nov'17 to Nov'20
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) TOP5 list in Nov'20
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

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,			
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,			
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

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
   i. Find & identify these 8 supercomputing centres
   ii. Characterize the architecture of Deucalion in MACC

Analysis of the key systems in 2020

1. **#1** in Nov'20: **Fugaku** *(Fujitsu A64FX, 48 cores)*, follow-up of #1 in Jun'11, **K-Computer** *(SPARC64 VIIIfx, 8 cores)*
2. **#2** in Nov'20, #1 in Nov'18: **Summit** *(IBM POWER9, 22 cores + NVidia Volta GV100)*; +**Sierra** were the follow-up of #1 in Jun'12, Sequoia *(IBM POWER BGQ, 16 cores)*
3. **#4** in Nov'20, #1 in Nov'17: **TaihuLight** *(Sunway SW26010, 260 c)*
4. **#5** in Nov'20: **Selene** *(AMD Epyc Rome 64 c + NVidia A100)*
5. **#6** in Nov'20: **Tianhe-2A** (MilkyWay-2A) *(Intel Xeon 2592v2, 12c + Matrix-2000)*, follow-up of #1 in Nov'10, **Tianhe-1A** (MilkyWay-1A) *(Xeon, 6 c + NVidia Fermi)*
6. **#33** in Nov'20: **SX-Aurora TSUBASA** *(AMD Epyc Rome + NEC VE)*

1

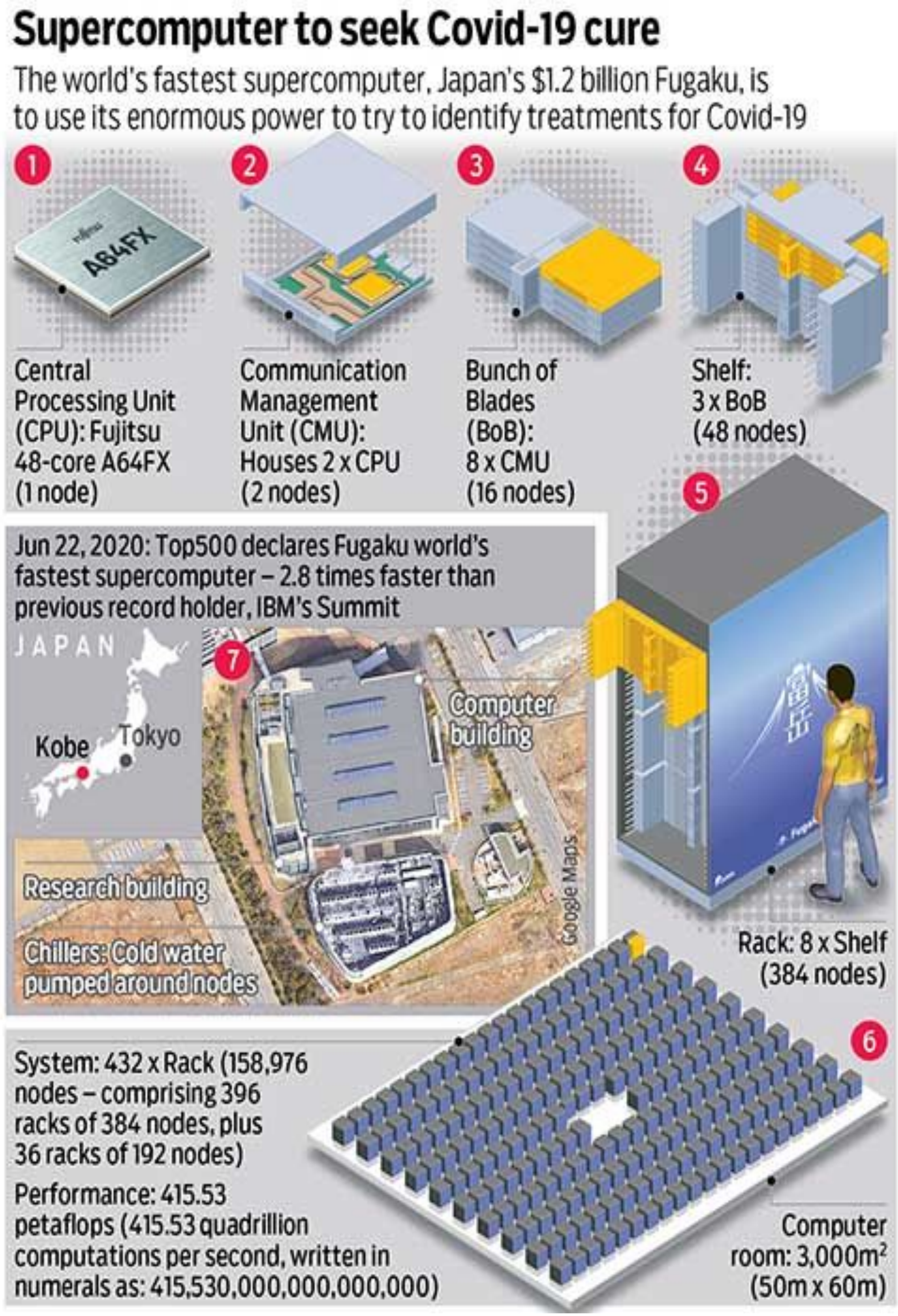
富岳

Fugaku

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

TOP500 The List.

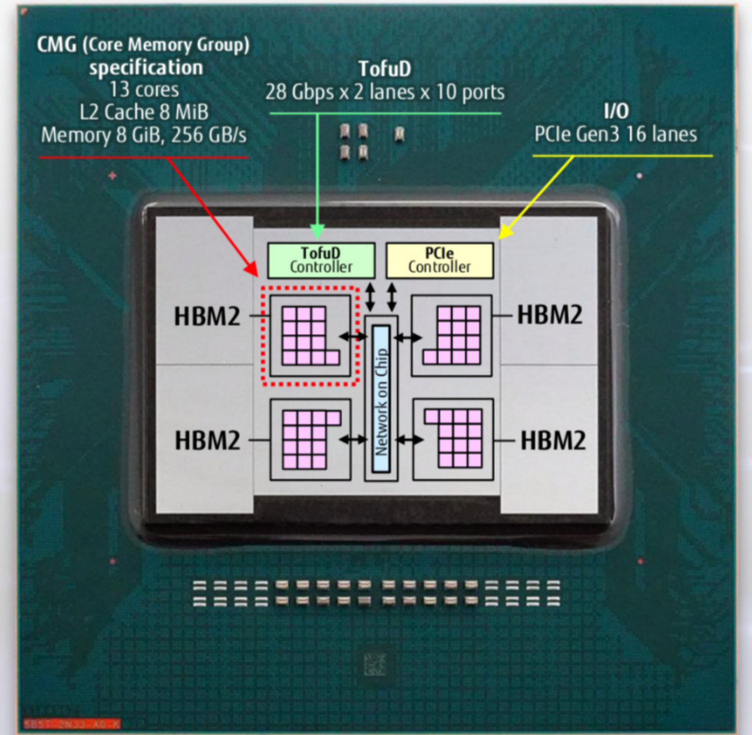
since Jun'20



Fujitsu A64FX in Fugaku

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

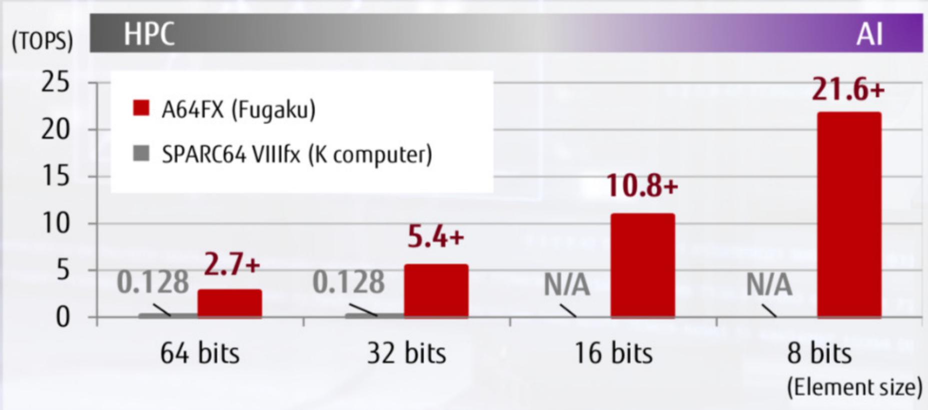
FUJITSU



■ 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

Copyright 2019 FUJITSU LIMITED

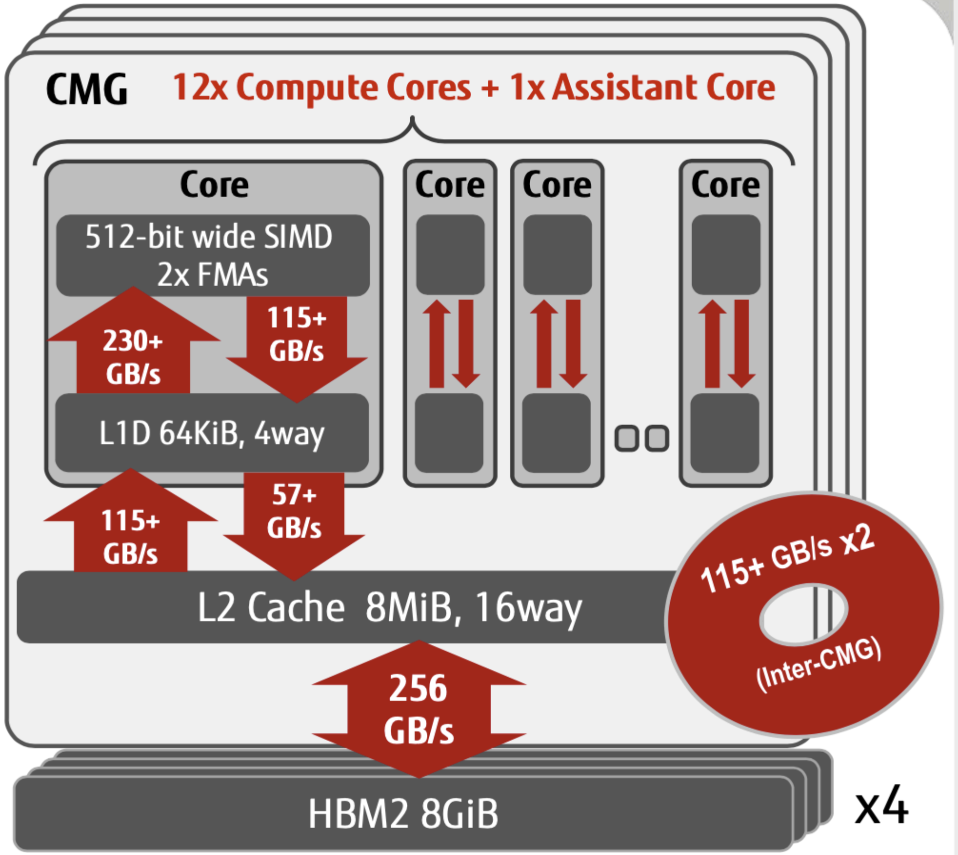
Fujitsu A64FX in Fugaku

Fujitsu-designed CPU Core w/ High Memory Bandwidth

FUJITSU

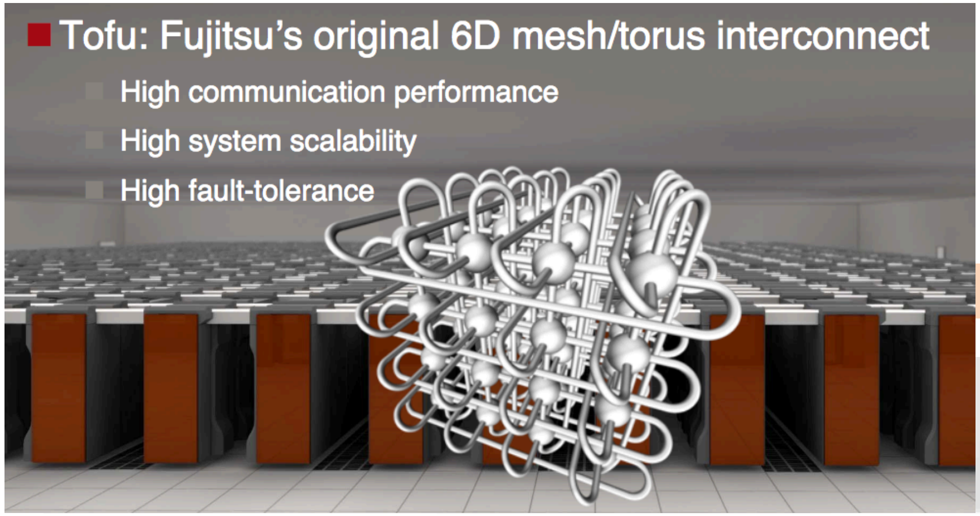
- A64FX out-of-order controls in cores, caches, and memories achieve superior throughput

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

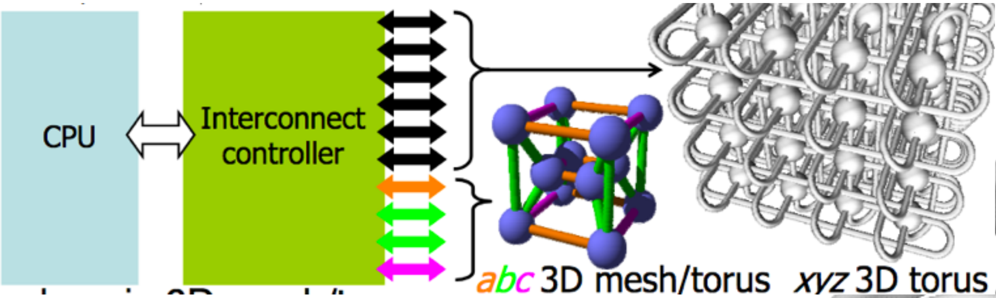


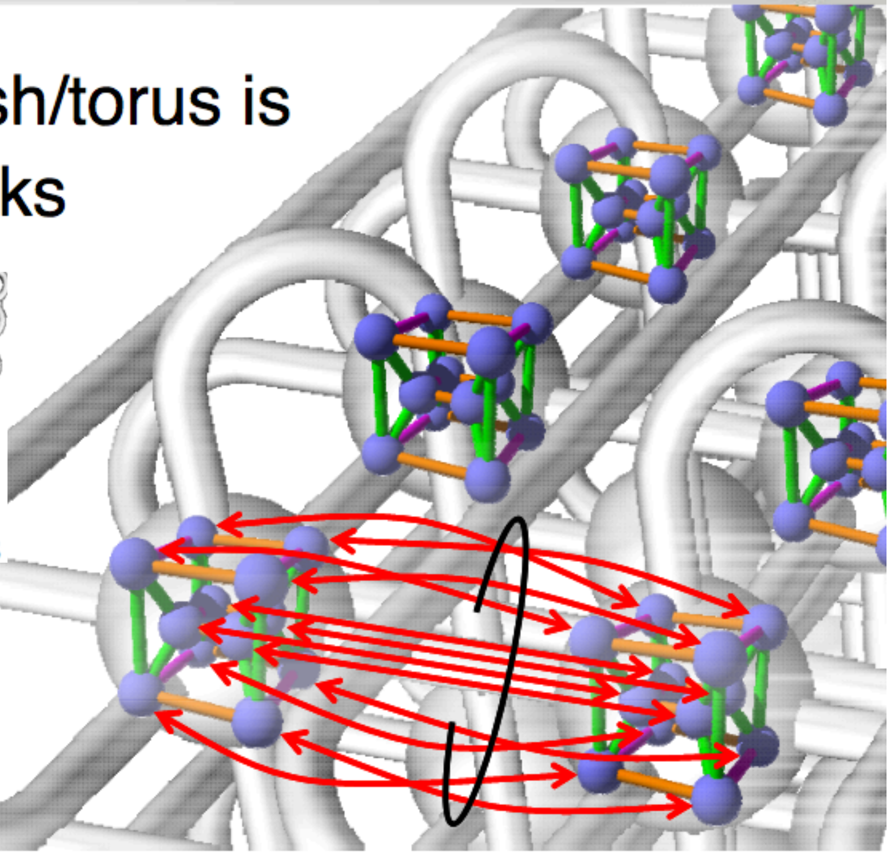
Post-K Activities, ISC19, Frankfurt 3 Copyright 2019 FUJITSU LIMITED

Tofu3: 6D mesh/torus interconnect



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





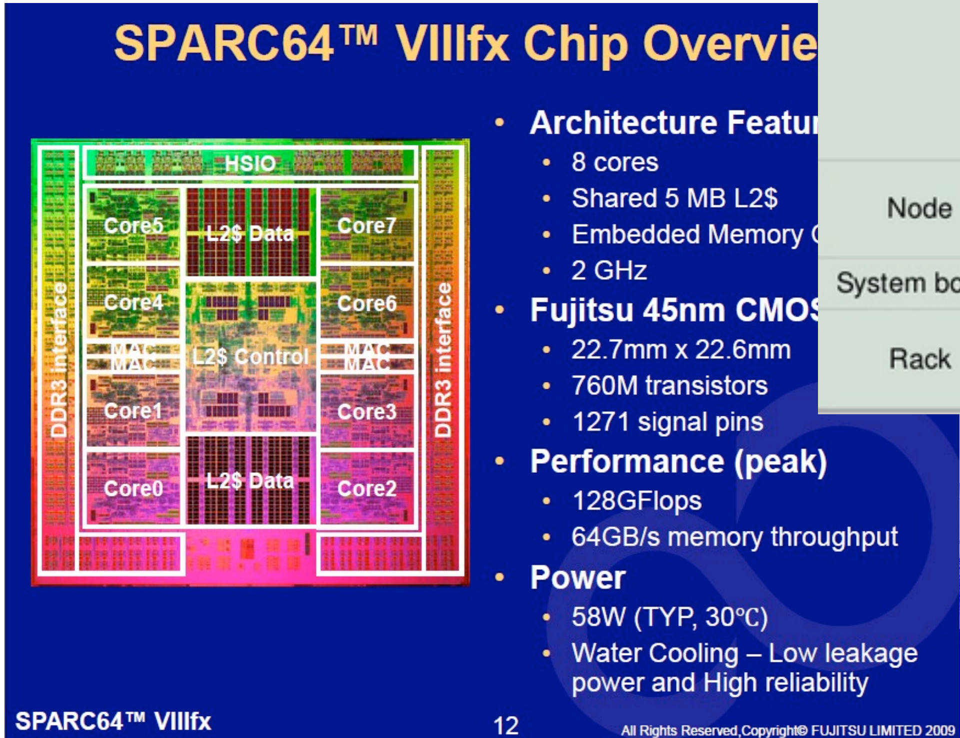
Fujitsu K computer

(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

TOP500 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)

TOP500 The List.

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

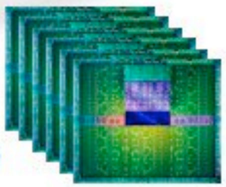
Summit Overview

OpenPOWER

Components

IBM POWER9
- 22 Cores
- 4 Threads/core
- NVLink

NVIDIA GV100
- 7 TF
- 16 GB @ 0.9 TB/s
- NVLink



Compute Node

2 x POWER9
6 x NVIDIA GV100
NVMe-compatible PCIe 1600 GB SSD

25 GB/s EDR IB- (2 ports)
512 GB DRAM- (DDR4)
96 GB HBM- (3D Stacked)
Coherent Shared Memory

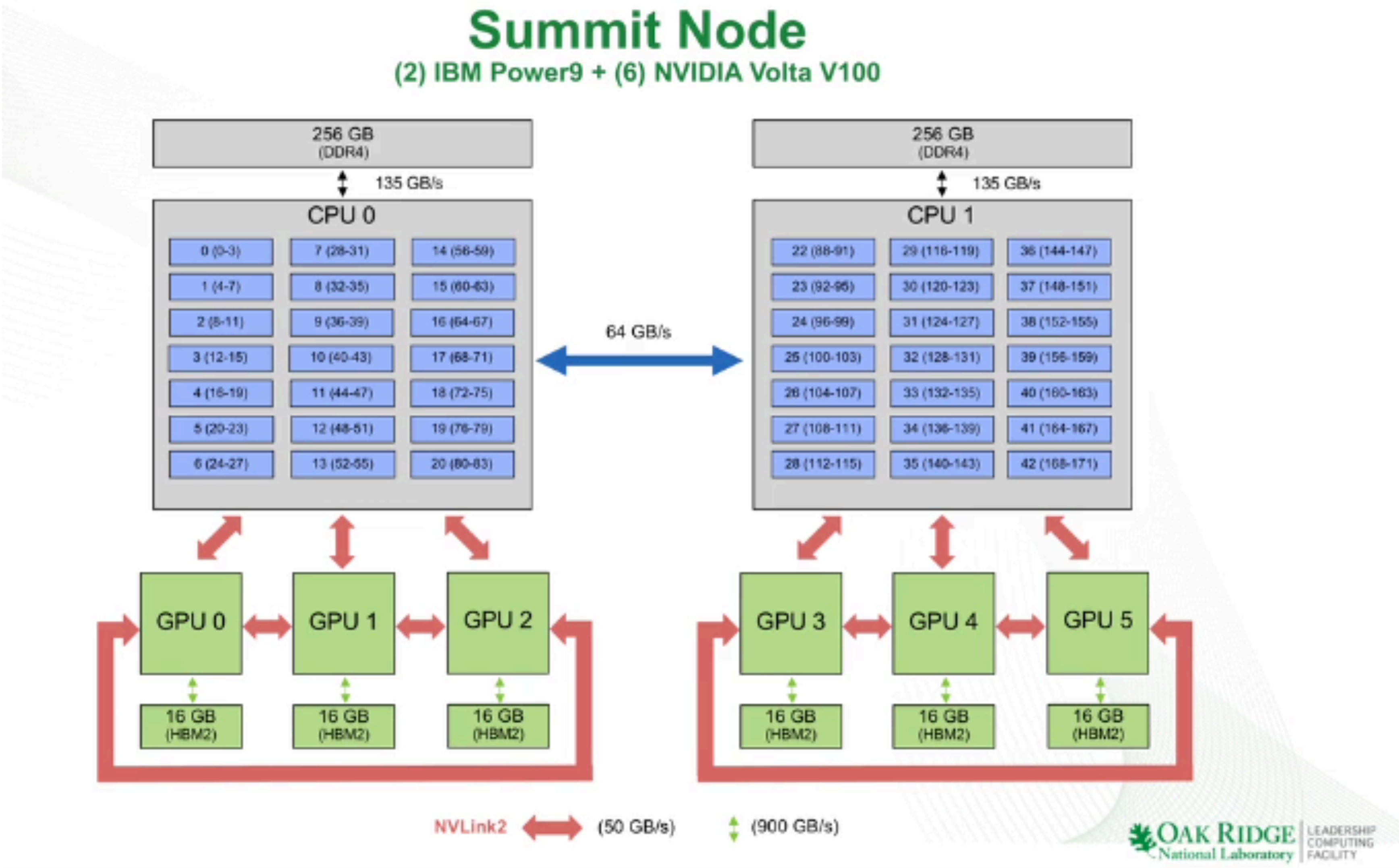
Compute Rack

18 Compute Servers
Warm water (70°F direct-cooled components)
RDHX for air-cooled components

Compute System

10.2 PB Total Memory
256 compute racks
4,608 compute nodes
Mellanox EDR IB fabric
200 PFLOPS
~13 MW

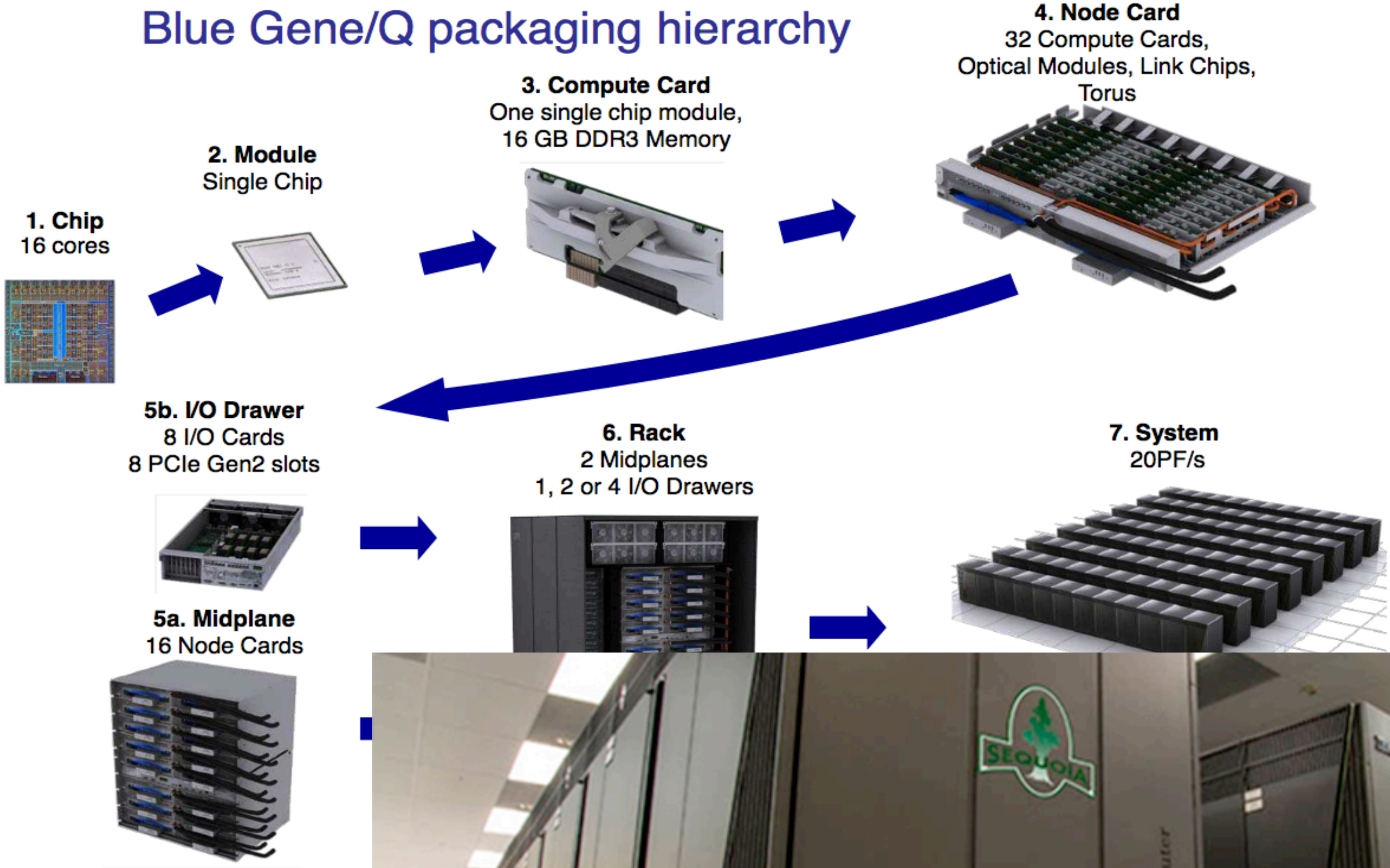
Summit/Sierra node architecture



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

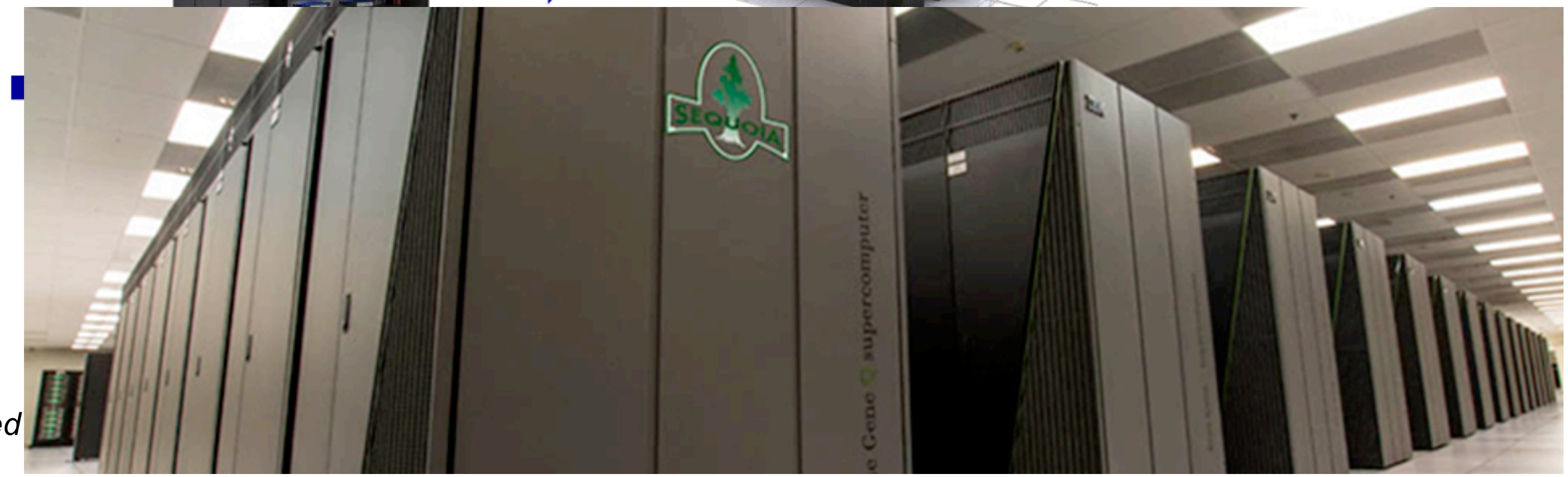
IBM Power BlueGene/Q Compute (Sequoia)

Blue Gene/Q packaging hierarchy



Ref: SC2010

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



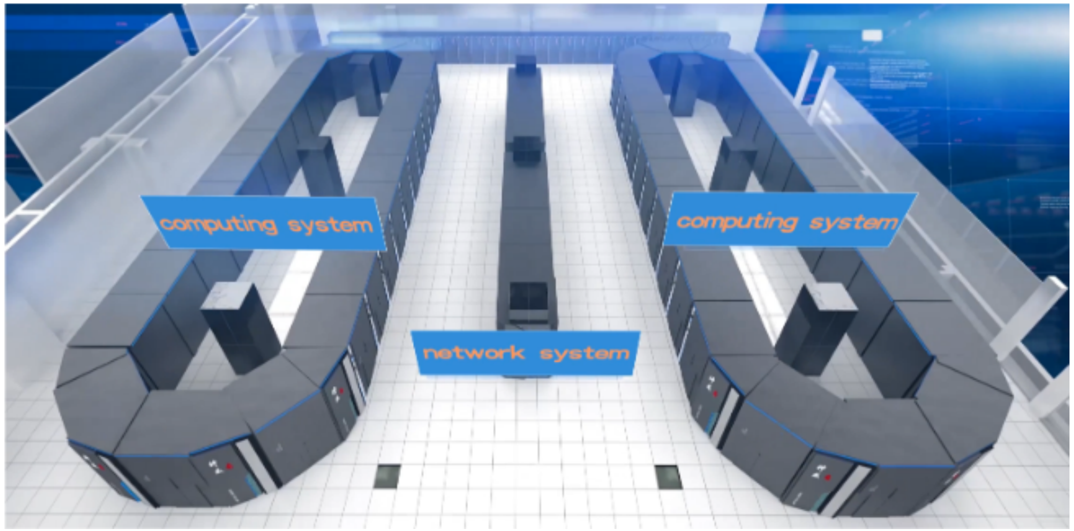
3

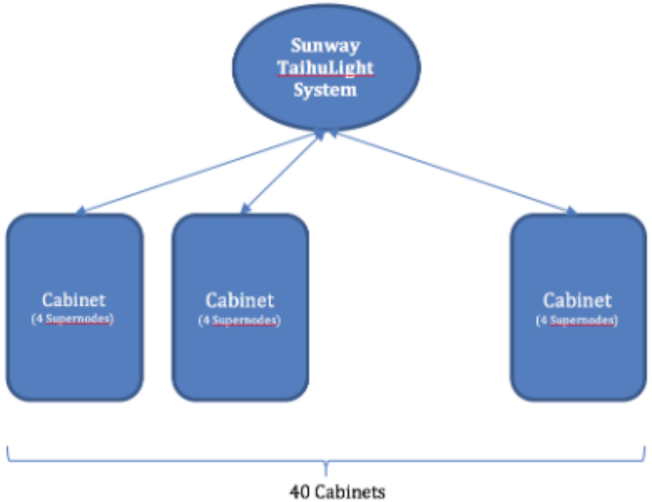
Sunway TaihuLight
(#1 in June'16 TOP500)

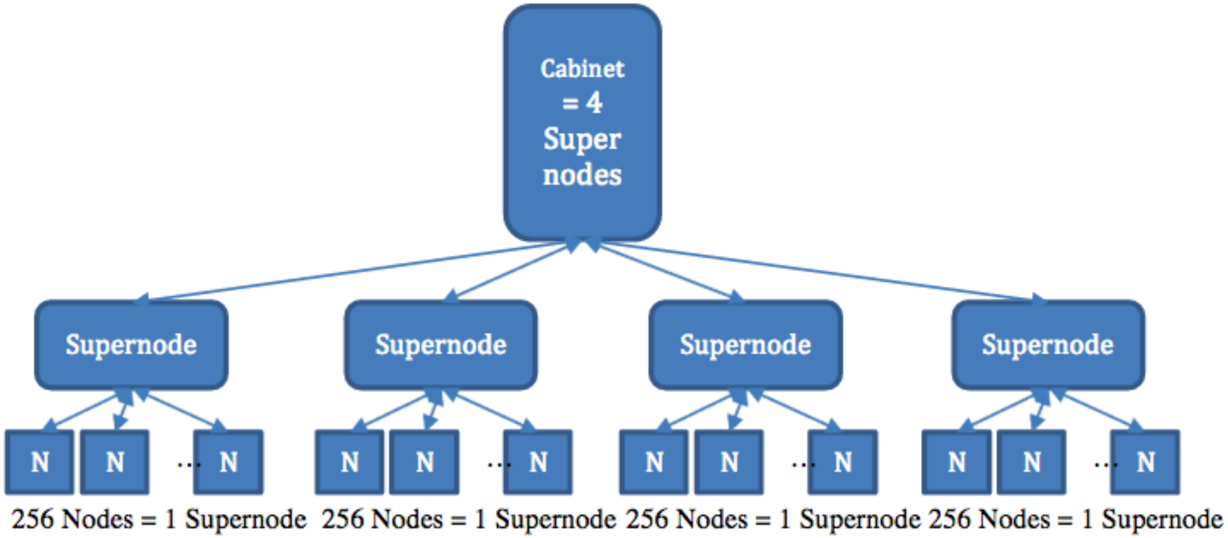
4 **Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC
National Supercomputing Center in Wuxi
China

since Jun'20

Overview of the Sunway TaihuLight System

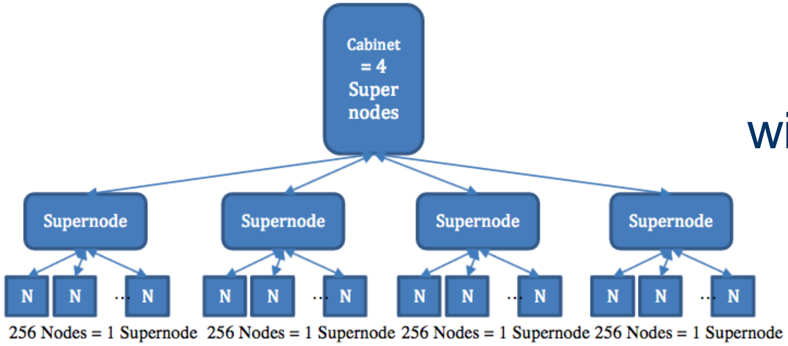




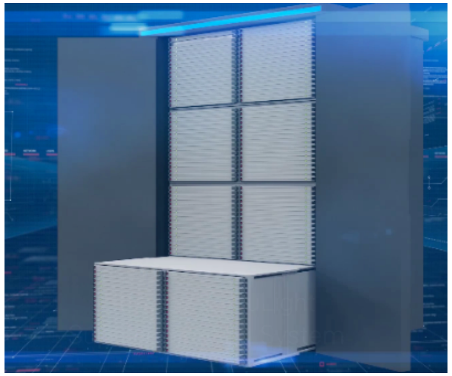


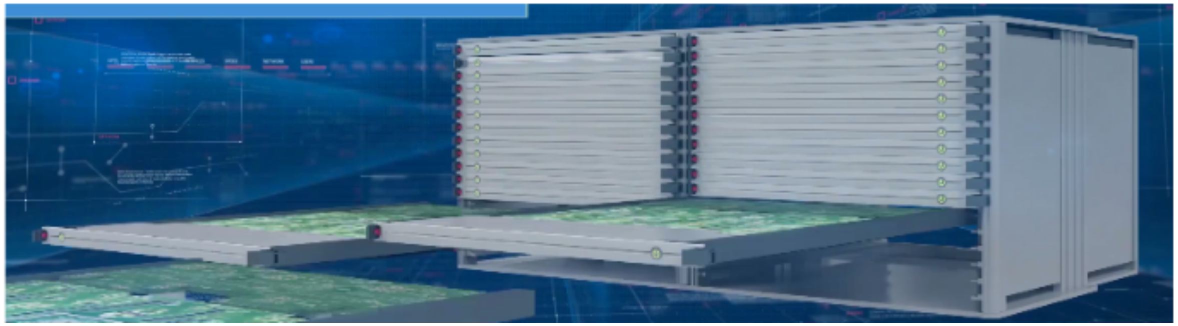
http://www.netlib.org/utk/people/JackDongarra/PAPERS/sunway-report-2016.pdf

Sunway TaihuLight
(#1 in June'16 TOP500)

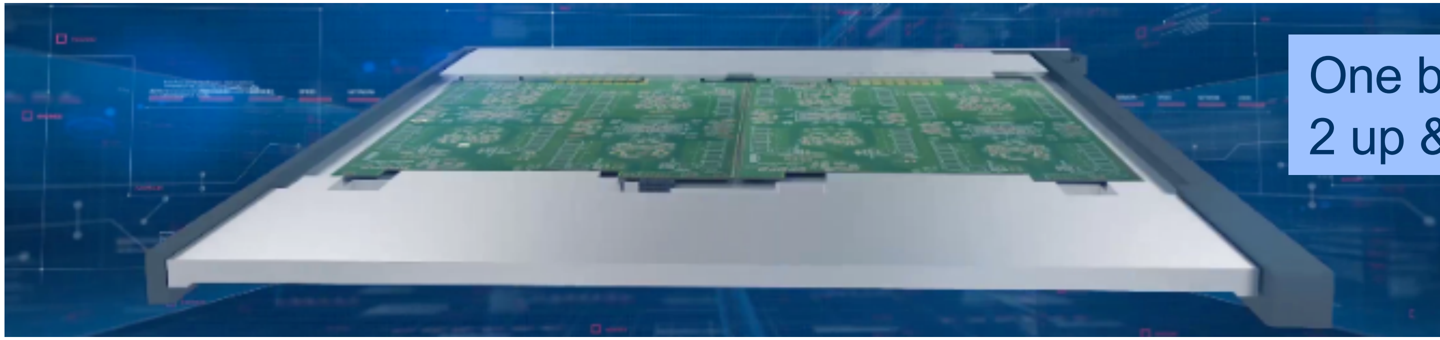


One cabinet
with 4 Supernodes





One Supernode
with 32 boards



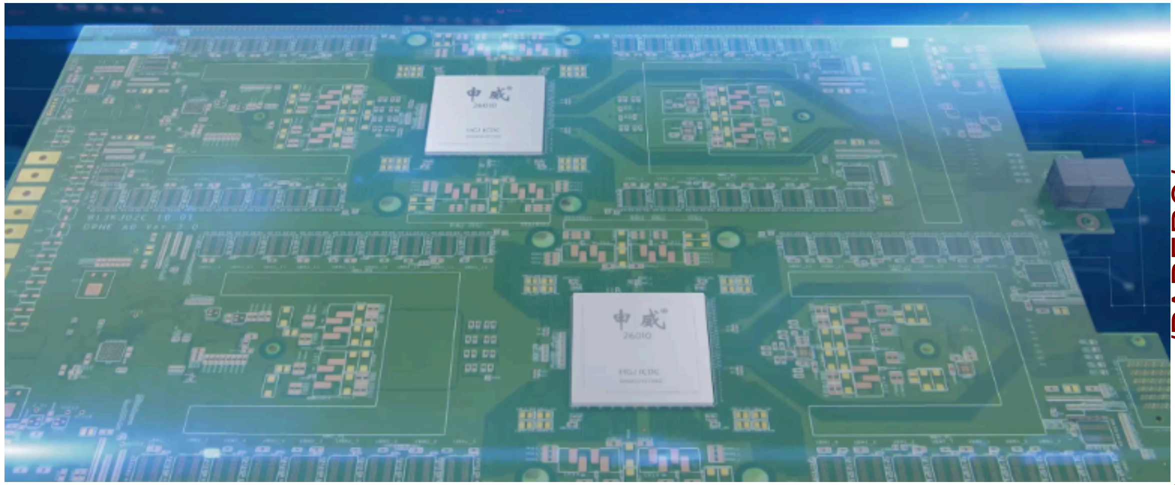
One board with 4 cards,
2 up & 2 down

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

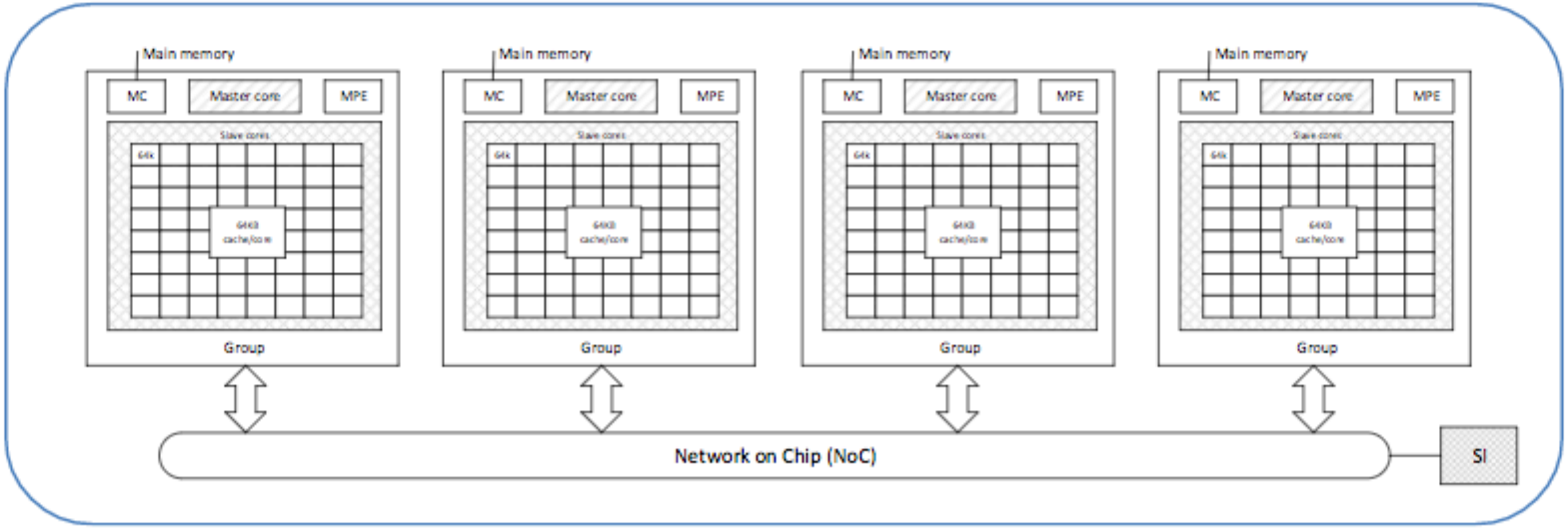
Sunway TaihuLight

(#1 in June'16 TOP500)

One card with two PU devices
(two SW26010 chips)



SW26010: the 4x64-core 64-bit RISC processor (w/ **256-bit vector** instructions & only cache L1)



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

4

Overview of Tianhe-2A

(#1 in June'13 TOP500)

6 TOP500 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'20**

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/

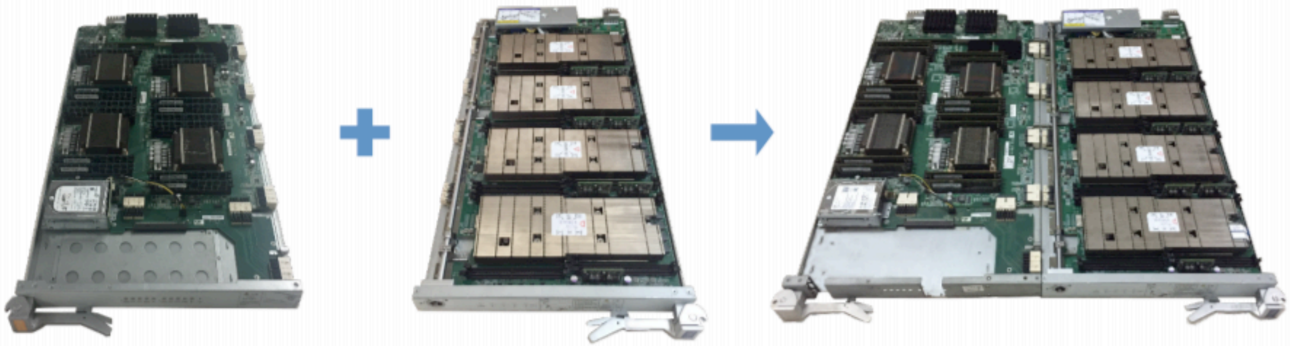
Overview of Tianhe-2A

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

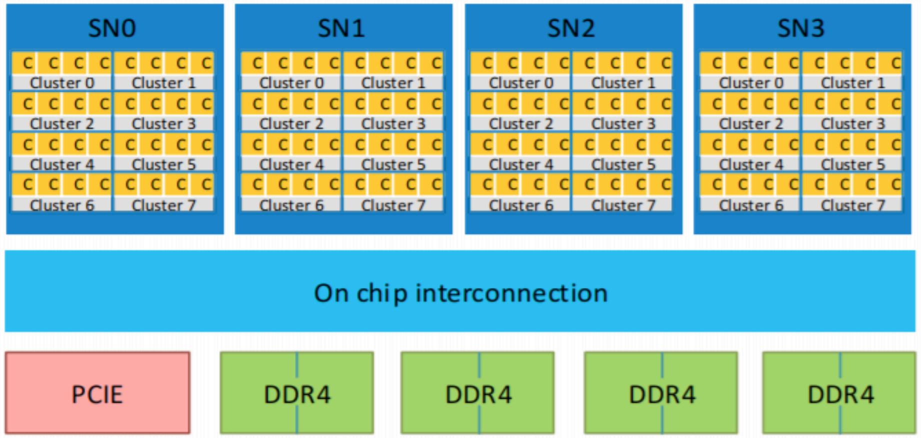
https://www.nextplatform.com/2018/02/14/look-whats-store-chinas-tianhe-2a-supercomputer/

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

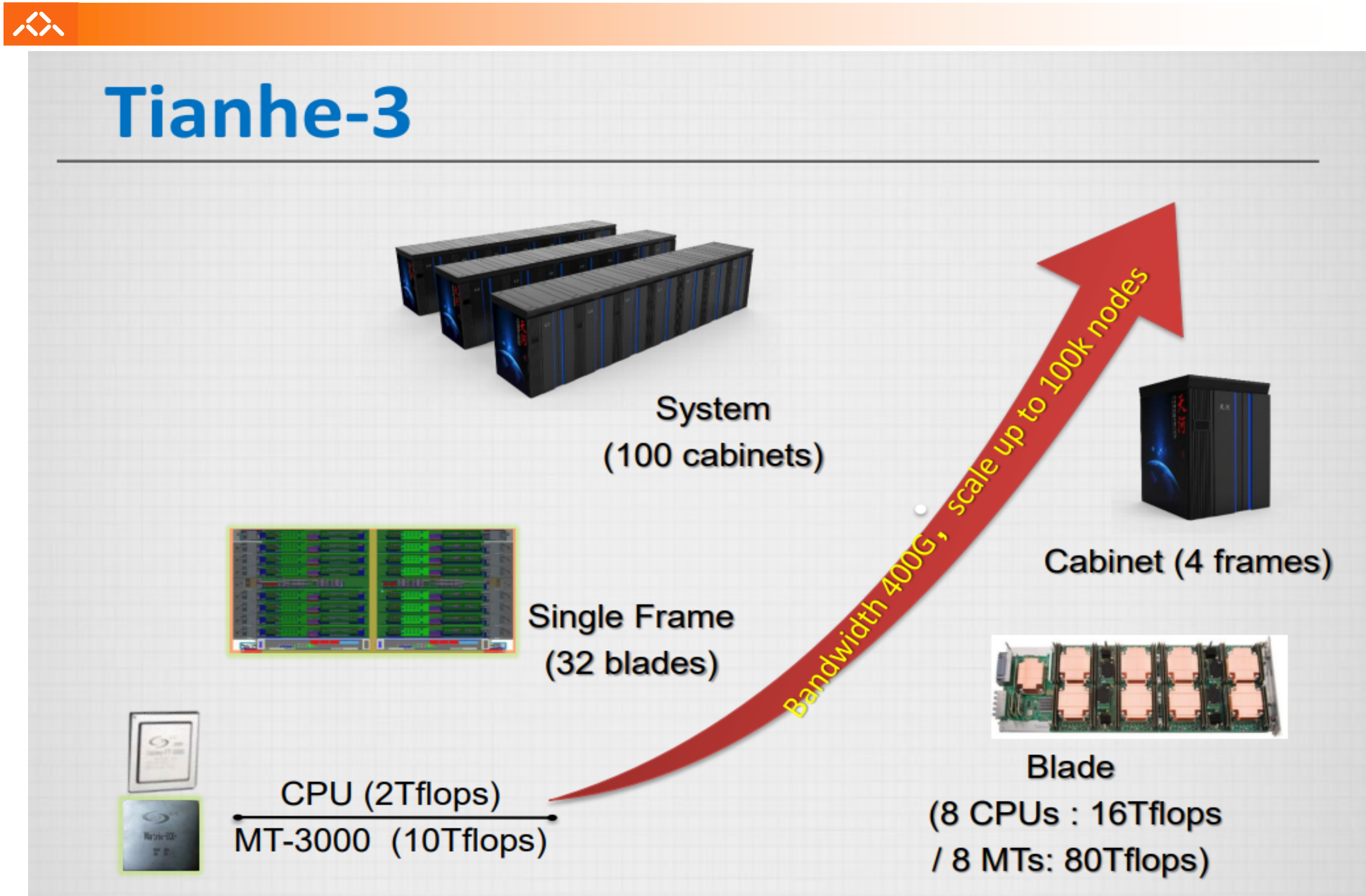
    4 SNs x 8 clusters x 4cores x 16 flops x 1.2 GHz = 2.4576 Tflops
  - Peak power dissipation: ~240w
  - Interface
    - 8 DDR4-2400 channels
    - X16 PCIE 3.0 EP Port



15

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

Next: Tianhe-3
with Fujitsu A64FX ARM-SVE + Matrix-3000 accelerators



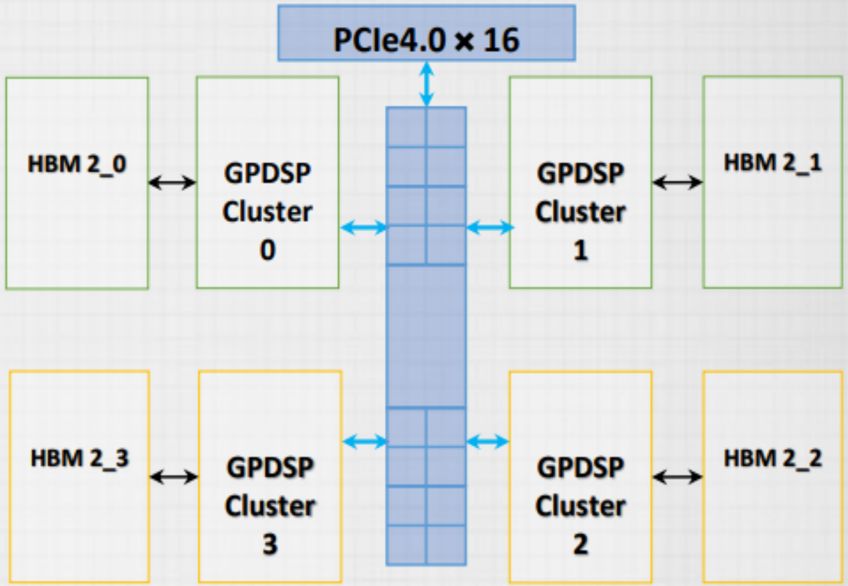
https://www.nextplatform.com/2019/05/02/china-fleshes-out-exascale-design-for-tianhe-3/

NUDT Matrix-3000 accelerator

National University of Defense Technology (NUDT)

Matrix-3000

- GPDSP
- Cores>=96, > 10 Tflops
- HBM2
- PCIe Gen4
- Support half precision



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

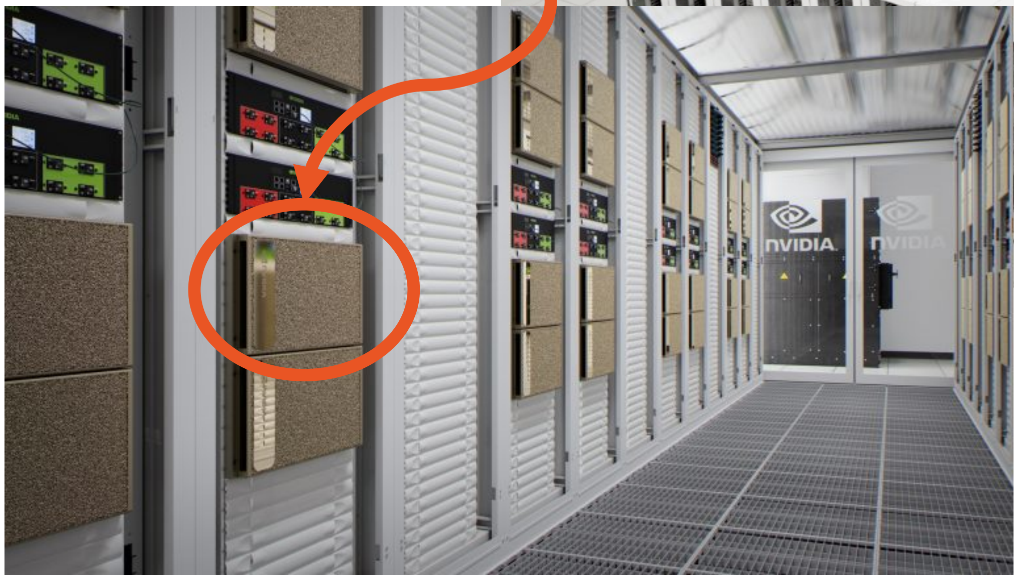
5

NVidia Selene: 280 DGX A100 nodes

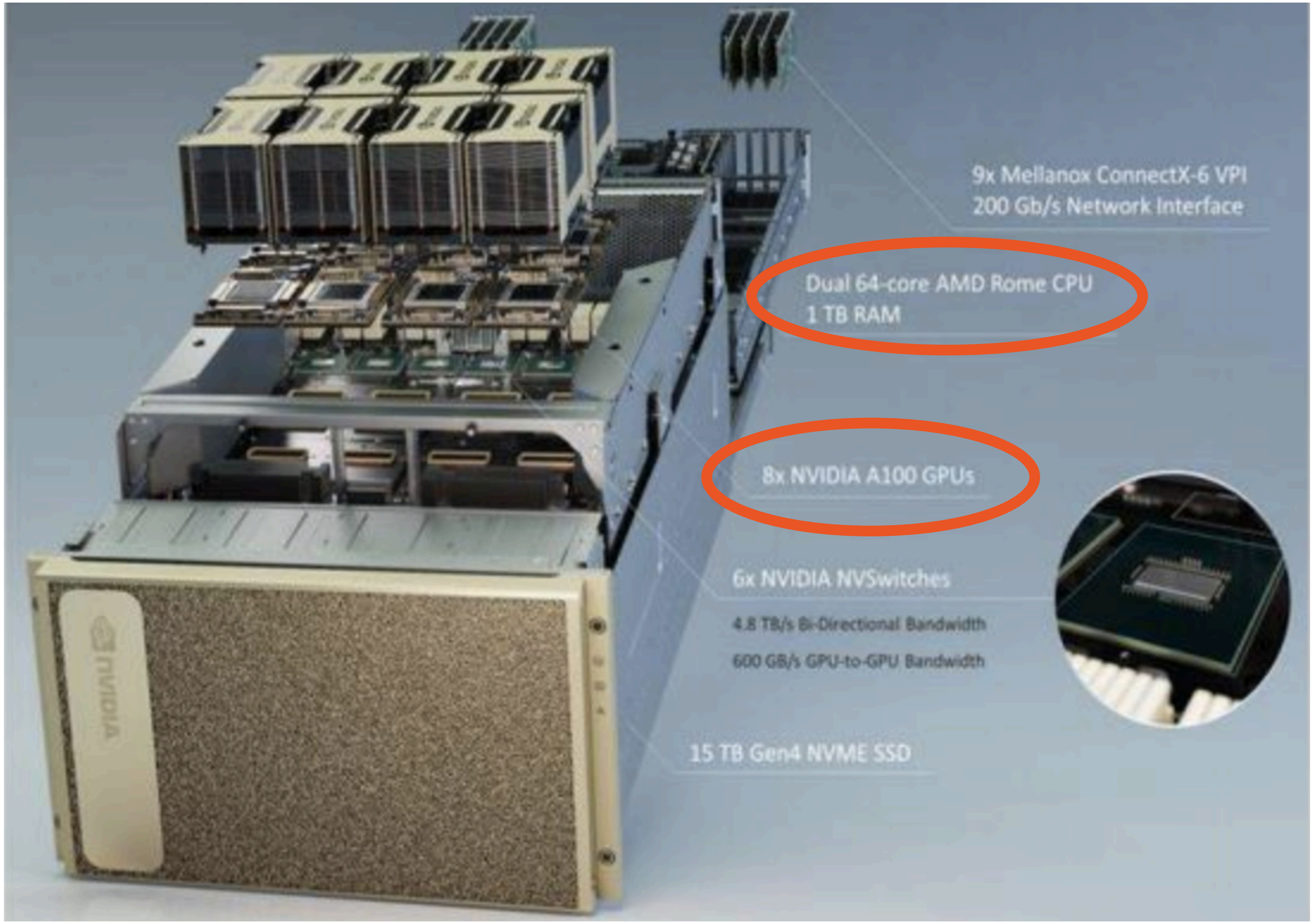
5 Selene - NVIDIA DGX A100, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia
NVIDIA Corporation
United States

Nov'20

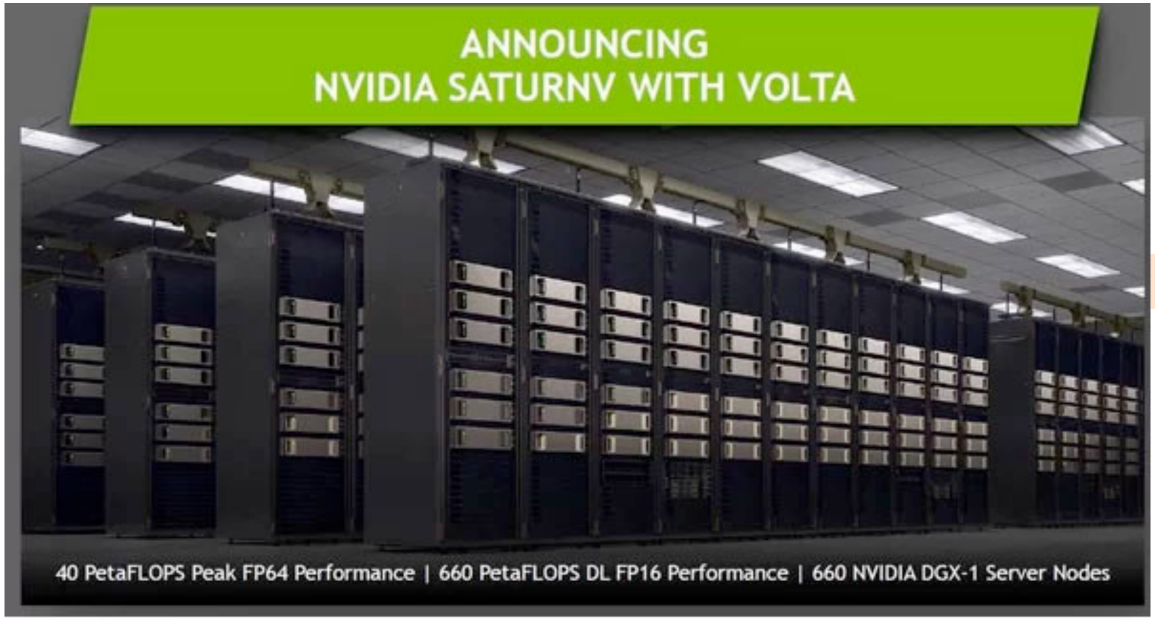
DGX A100 node

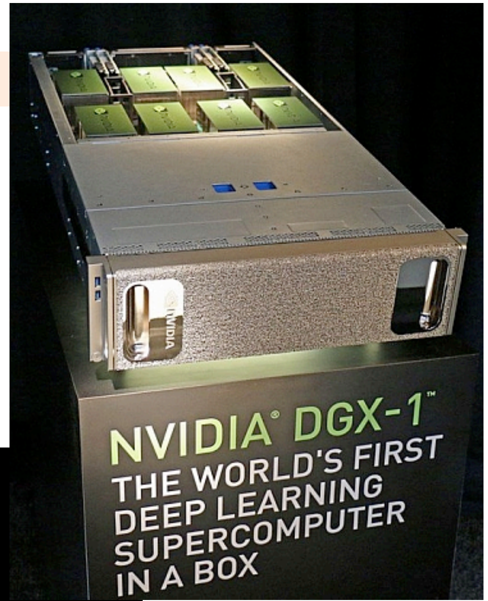


NVidia DGX A100 node

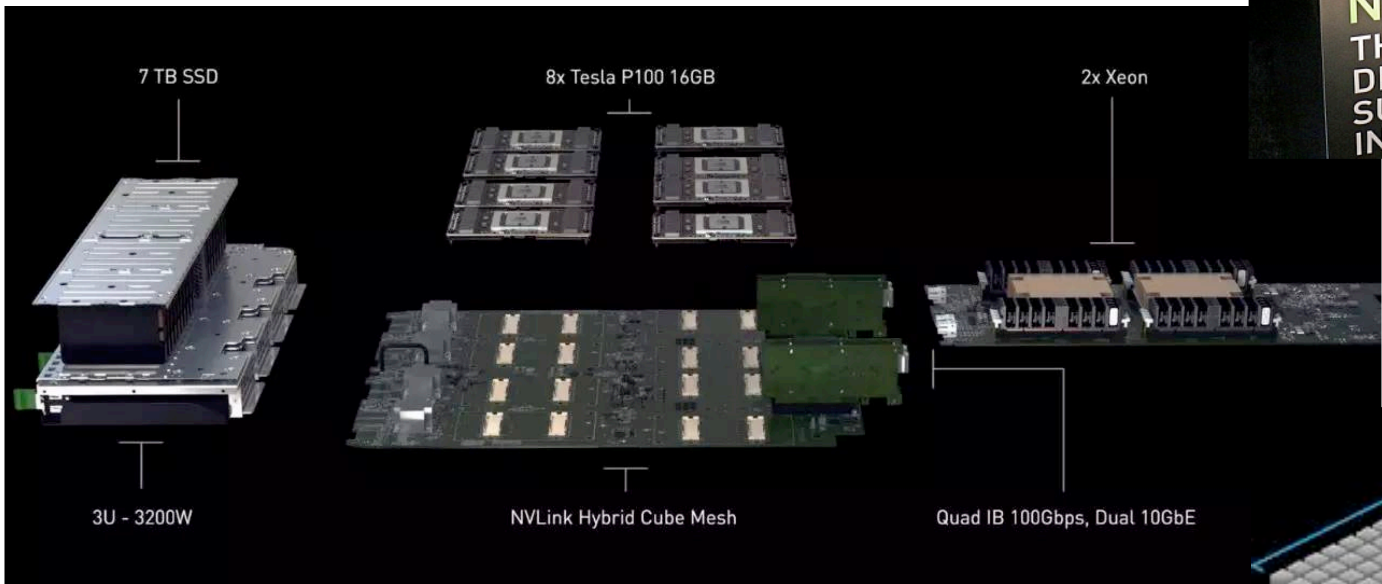


NVidia DGX-1 SaturnV: before Selene



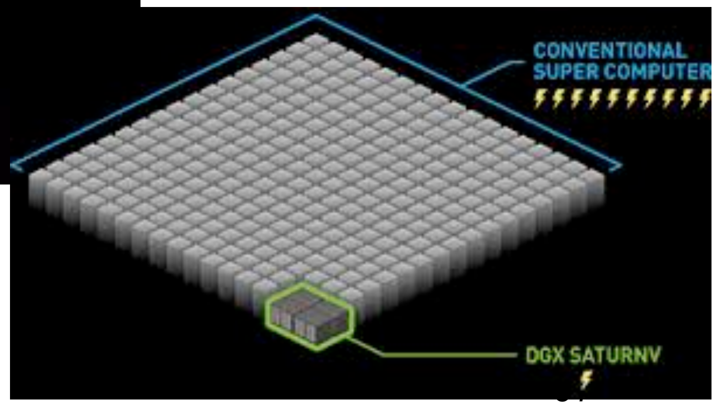


$149,000



2	374	DGX SaturnV Volta - NVIDIA DGX-1 Volta36, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla V100 , Nvidia NVIDIA Corporation United States	22,440

Nov'18 Green500

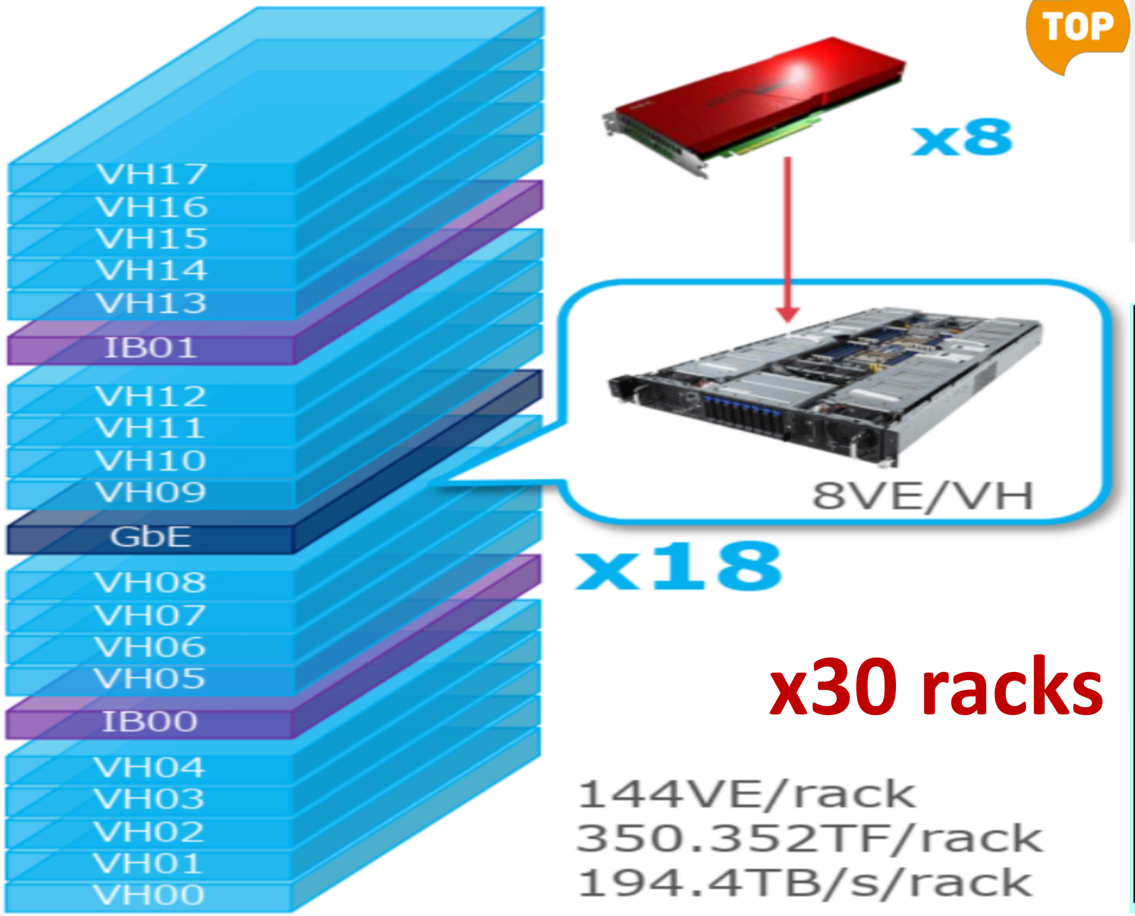


6

NEC SX-Aurora TSUBASA: the 412-8 model

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

Nov'20



Performance of SX-Aurora TSUBASA A412-8	
Total main memory capacity	202TB
Total computing performance	10.5PFlops
Number of vector engines (VE)	4,320
Number of cores per VE	8
Main memory capacity per VE	48GB
Computational performance per VE	2.4TFlops
Transfer speed between nodes (bidirectional)	10TB / s

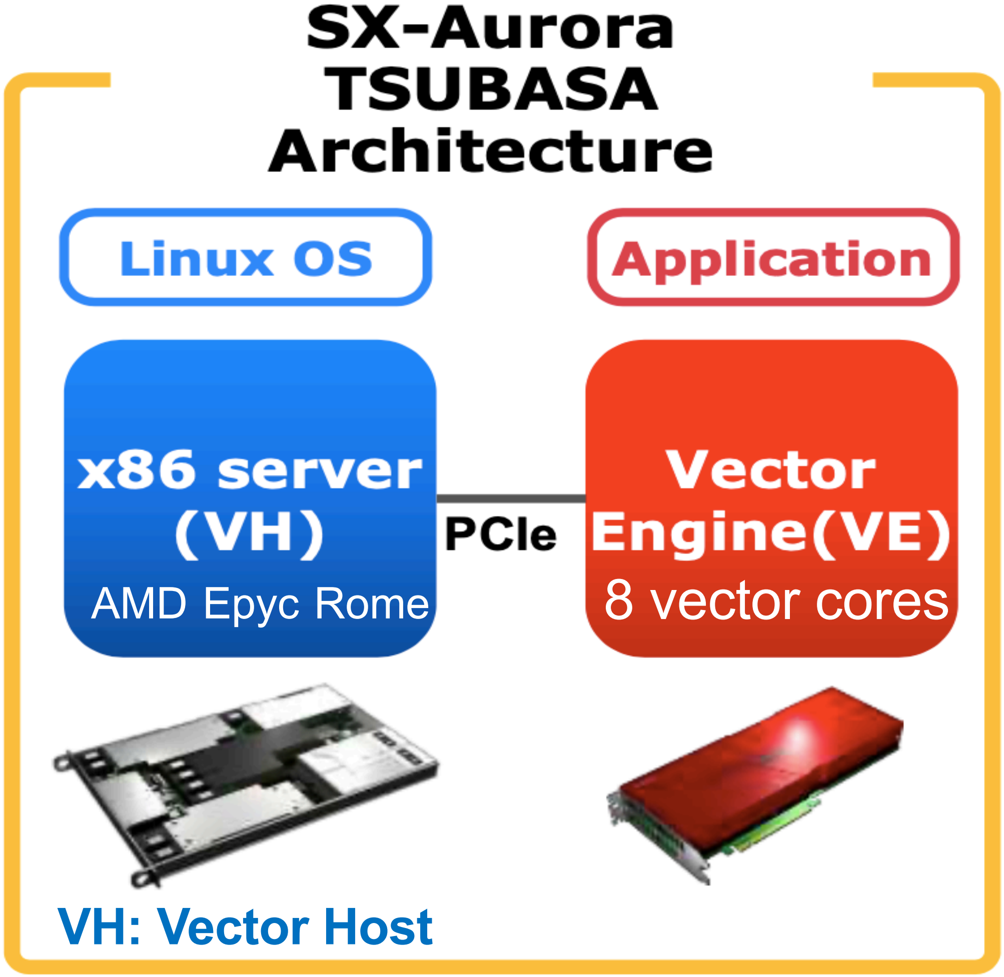
#1 The Earth Simulator, built from SX-6 nodes, was the fastest supercomputer from June 2002 to June 2004 on the LINPACK benchmark,

Architecture of NEC SX-Aurora TSUBASA:

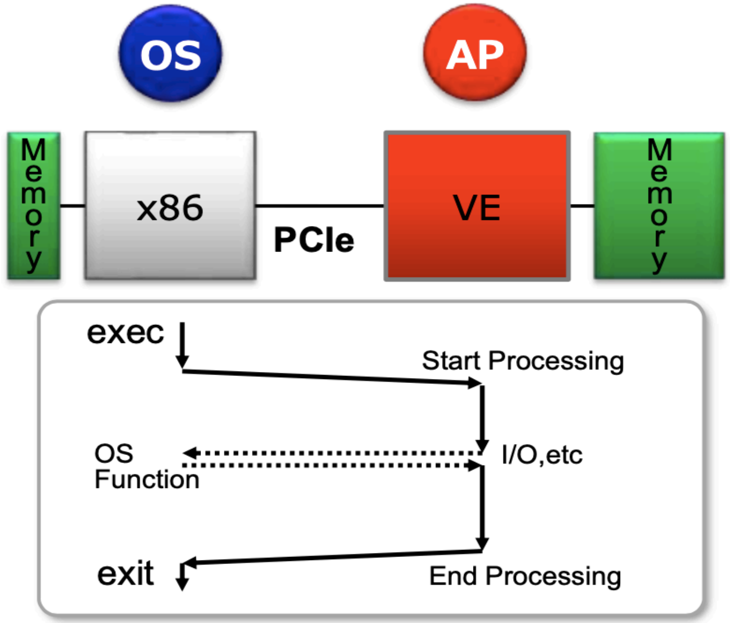
NEC supercomputers	
SX series	SX-1 · SX-2 · SX-3 · SX-4 · SX-5 · SX-6 · SX-7 · SX-8 · SX-9 · SX-ACE · **SX-Aurora TSUBASA**
SX-related topics	SUPER-UX · Earth Simulator

Earth Simulator

#1 from June 2002 till June 2004

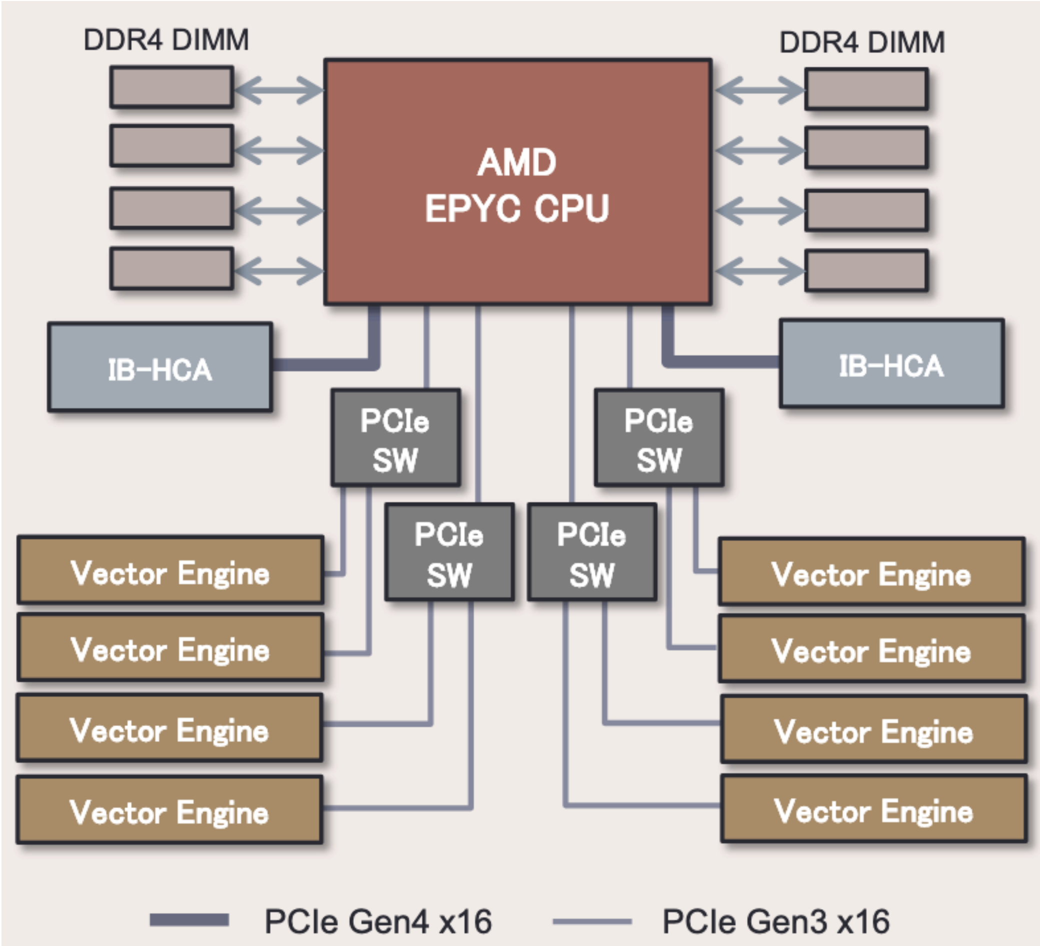


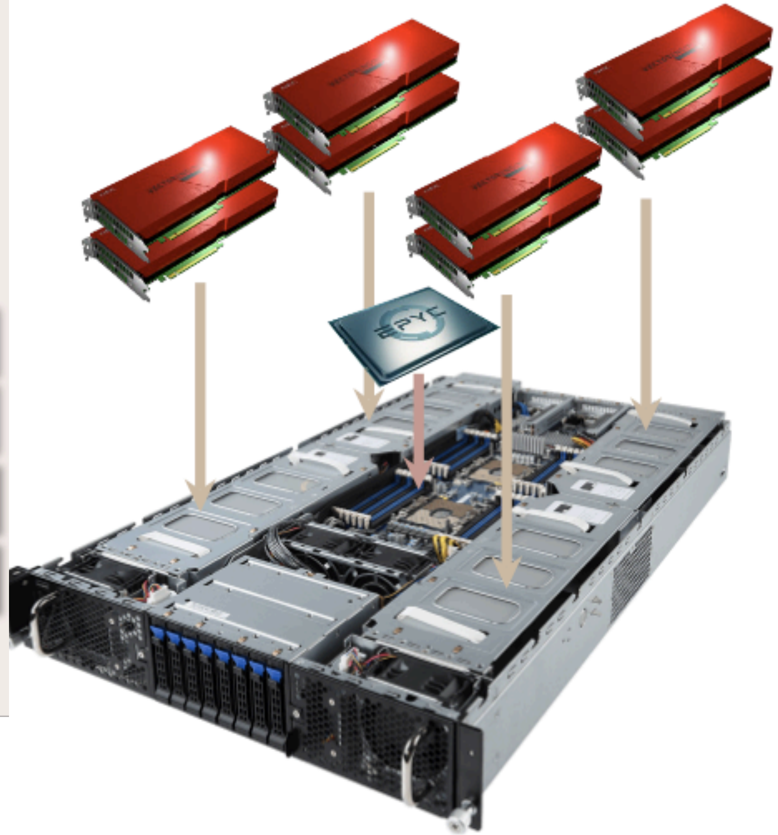
Aurora Architecture (OS offload model)



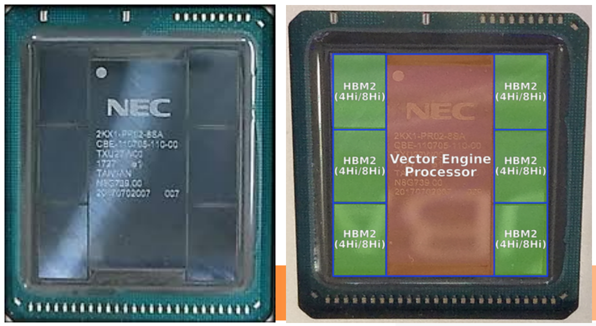
Whole AP is executed on VE

NEC SX-Aurora TSUBASA: the 412-8 model

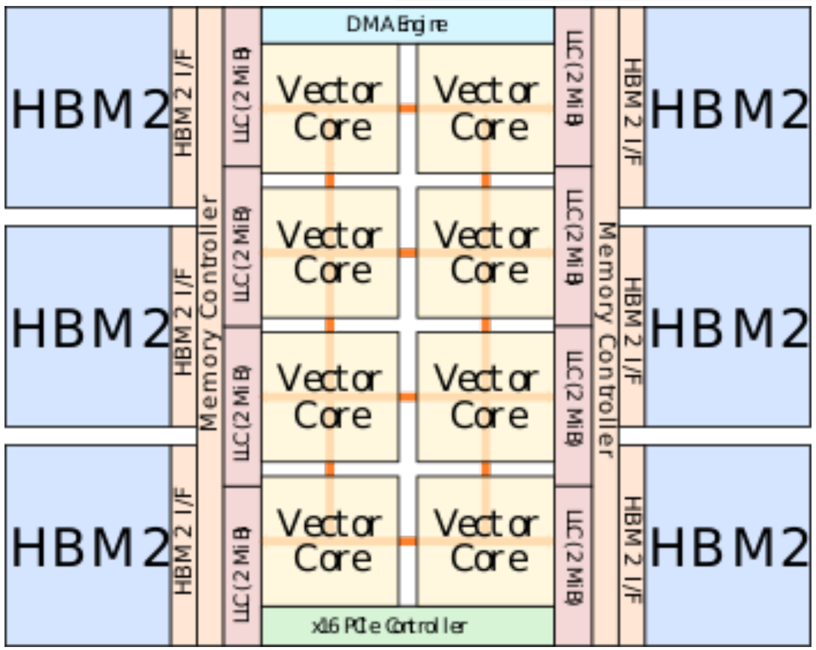




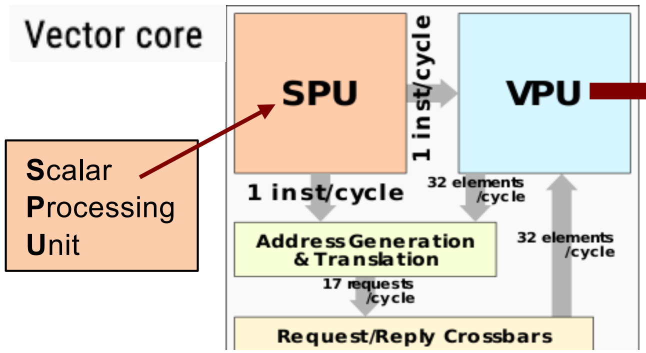
Architecture of NEC SX-Aurora SoC



Overview of the SX-Aurora

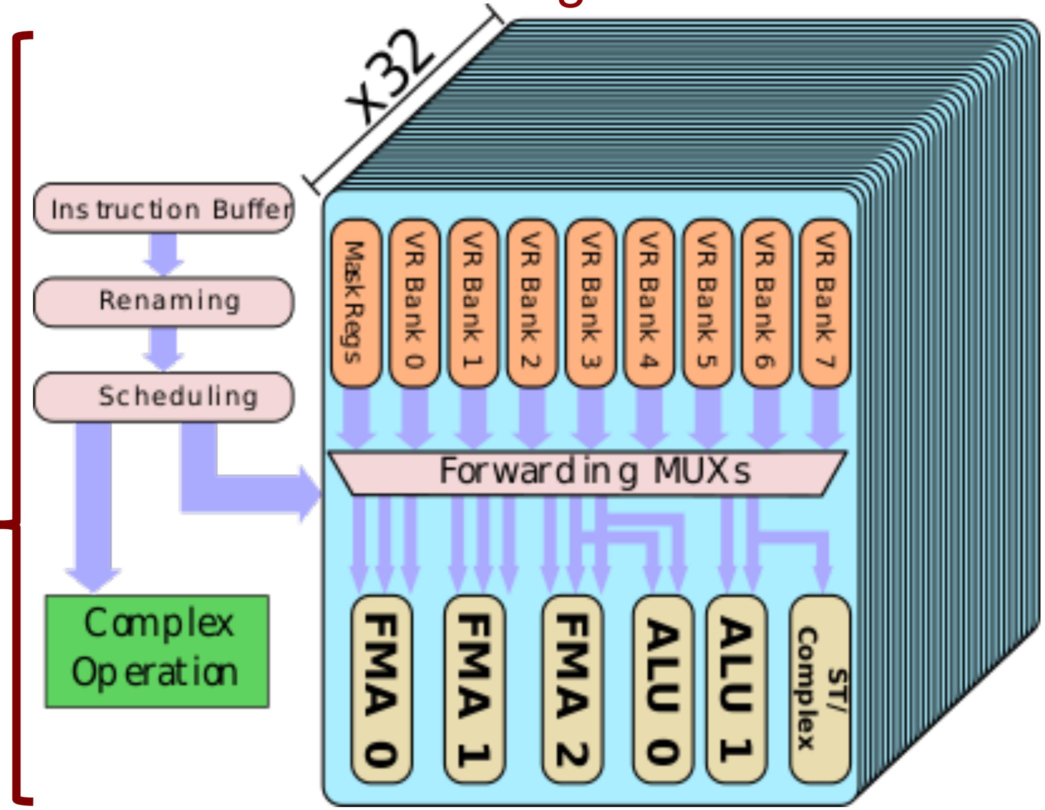


Vector core

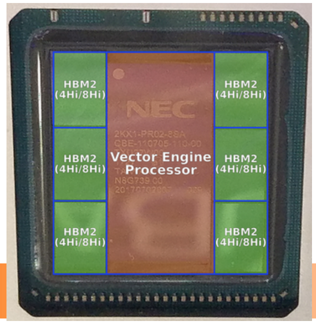


The Vector Processing Unit (**VPU**):

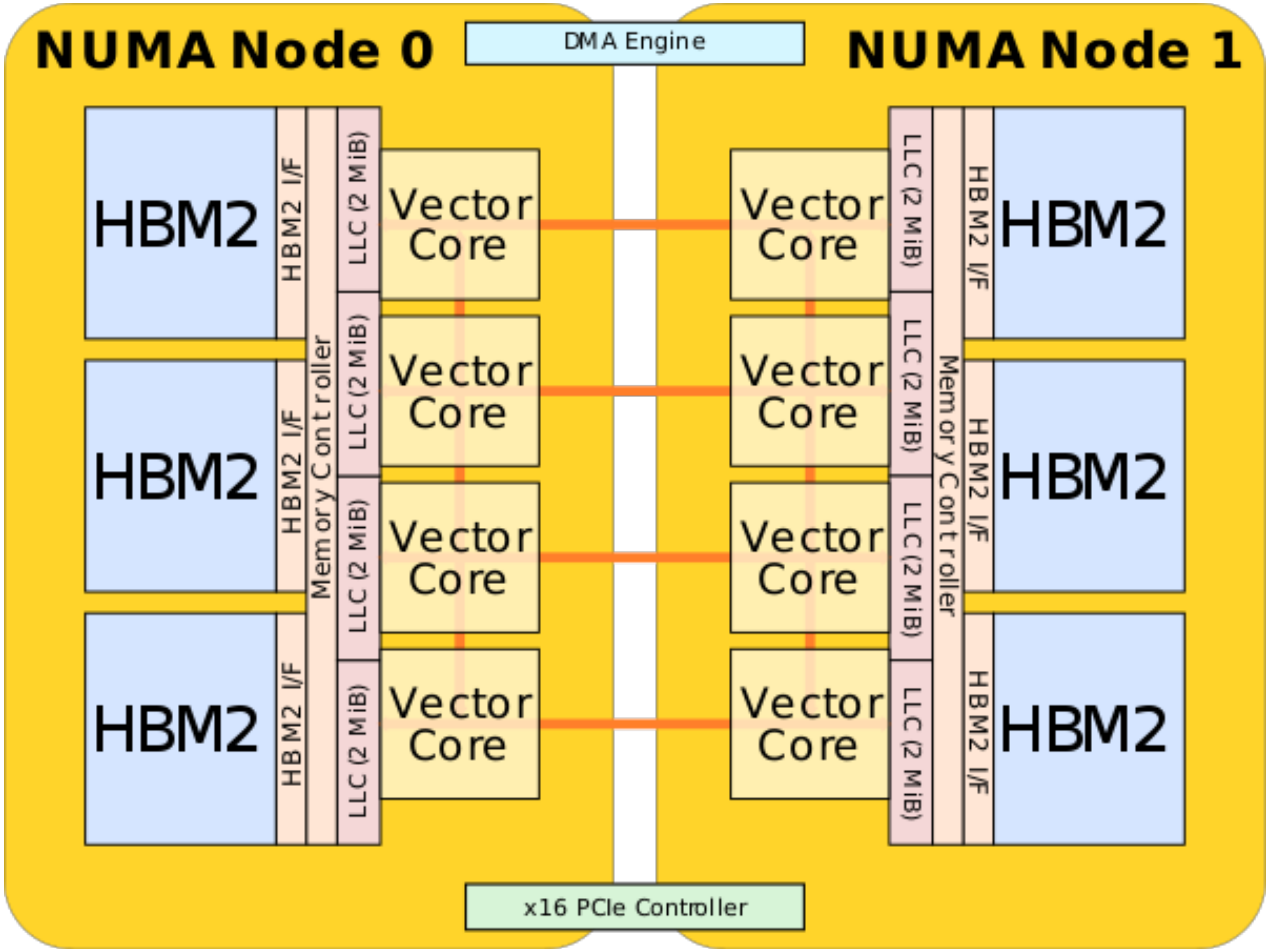
- 32 Vector Parallel Pipelines (VPP)
- 8 Vector Reg Banks
- 8*32 64-bit Registers



https://en.wikichip.org/wiki/nec/microarchitectures/sx-aurora



NUMA mode in NEC SX-Aurora



By default:
NUMA disabled

When enabled:
mem & LLC
not shared
between nodes

https://en.wikichip.org/wiki/nec/microarchitectures/sx-aurora

Outline

1. TOP500
   a) TOP10 lists from Nov’17 to Nov’20
   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) TOP5 list in Nov’20
   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

The Green500 list

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

The GREEN 500

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 MN-3 system

#1 in Jun'20 Green500

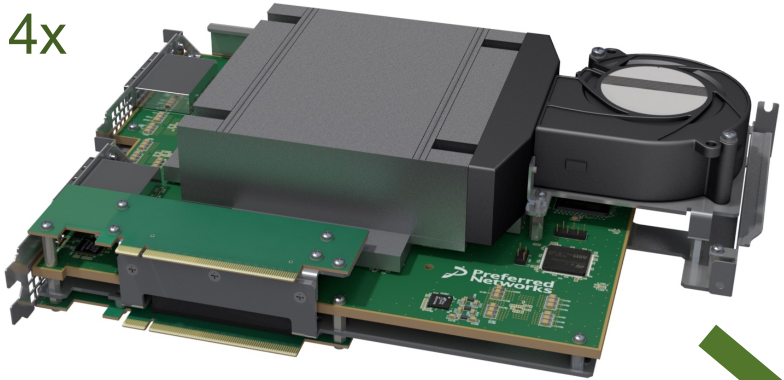
2	330	**MN-3** - MN-Core Server, Xeon Platinum 8260M 24C 2.4GHz, Preferred Networks MN-Core, MN-Core DirectConnect, Preferred Networks Preferred Networks Japan

The GREEN 500 *Nov'20*

2x

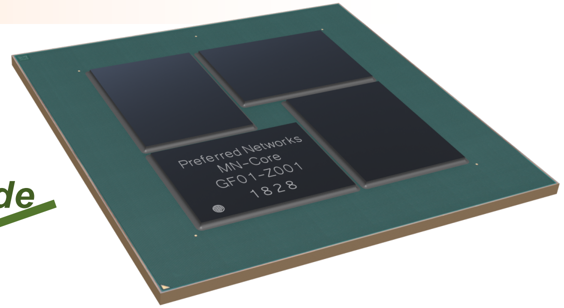
Xeon 8260M 24c

+ 4x



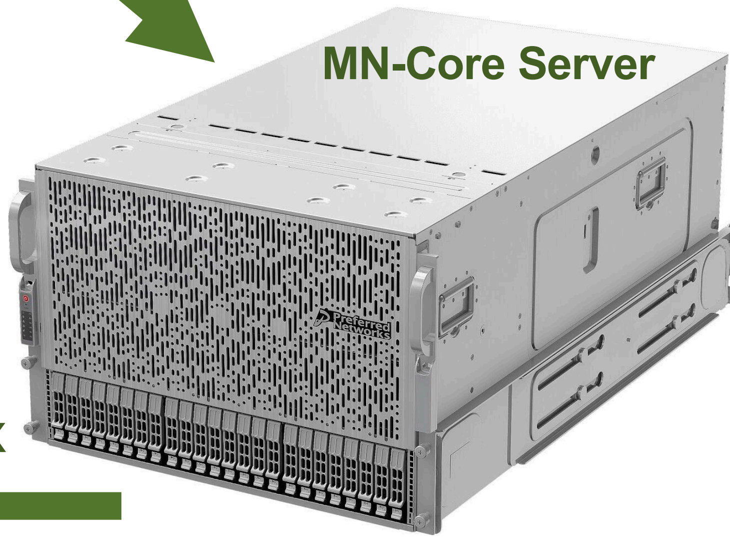
MN-Core Board

inside

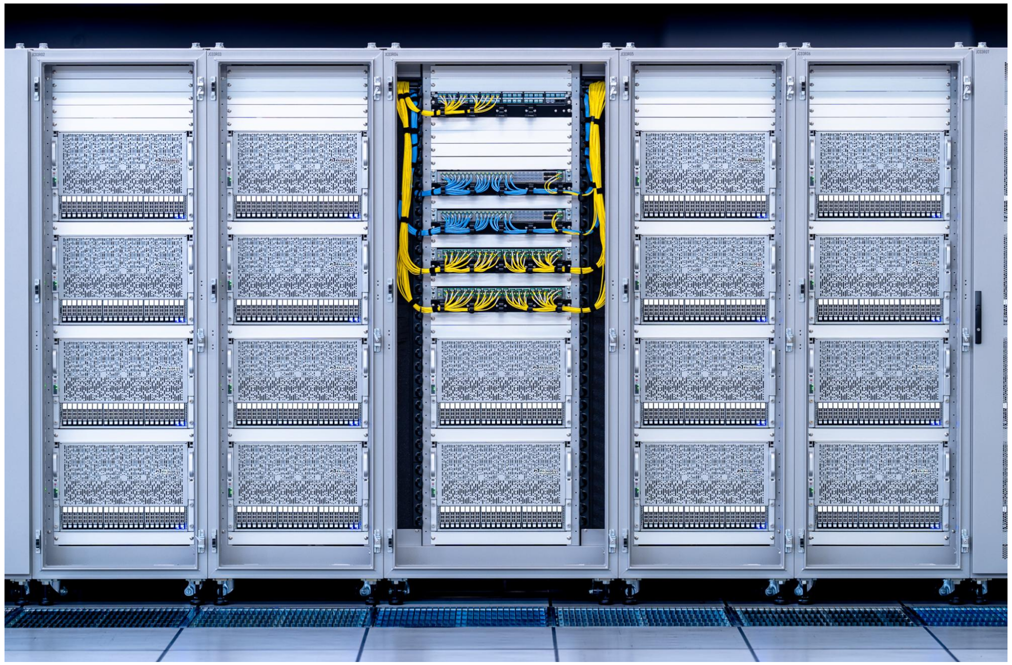


MN-Core accelerator

MN-Core Server

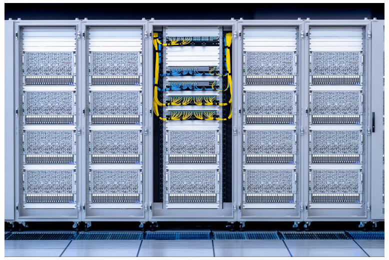


40x

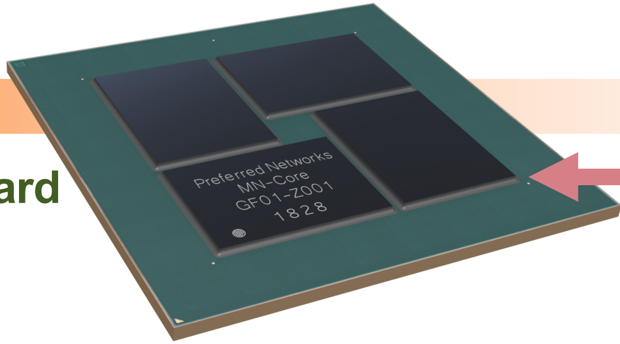


MN-3 *(partial view)*

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

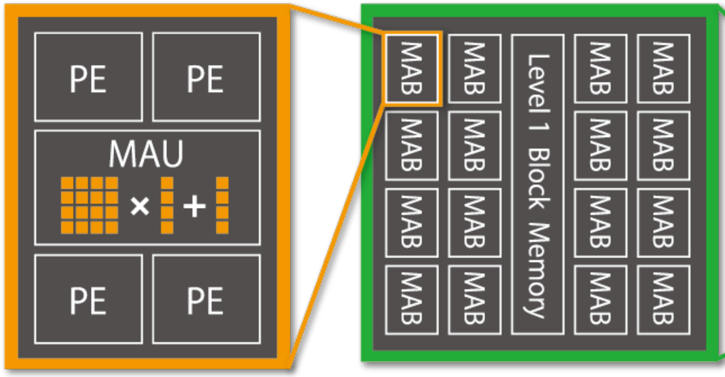


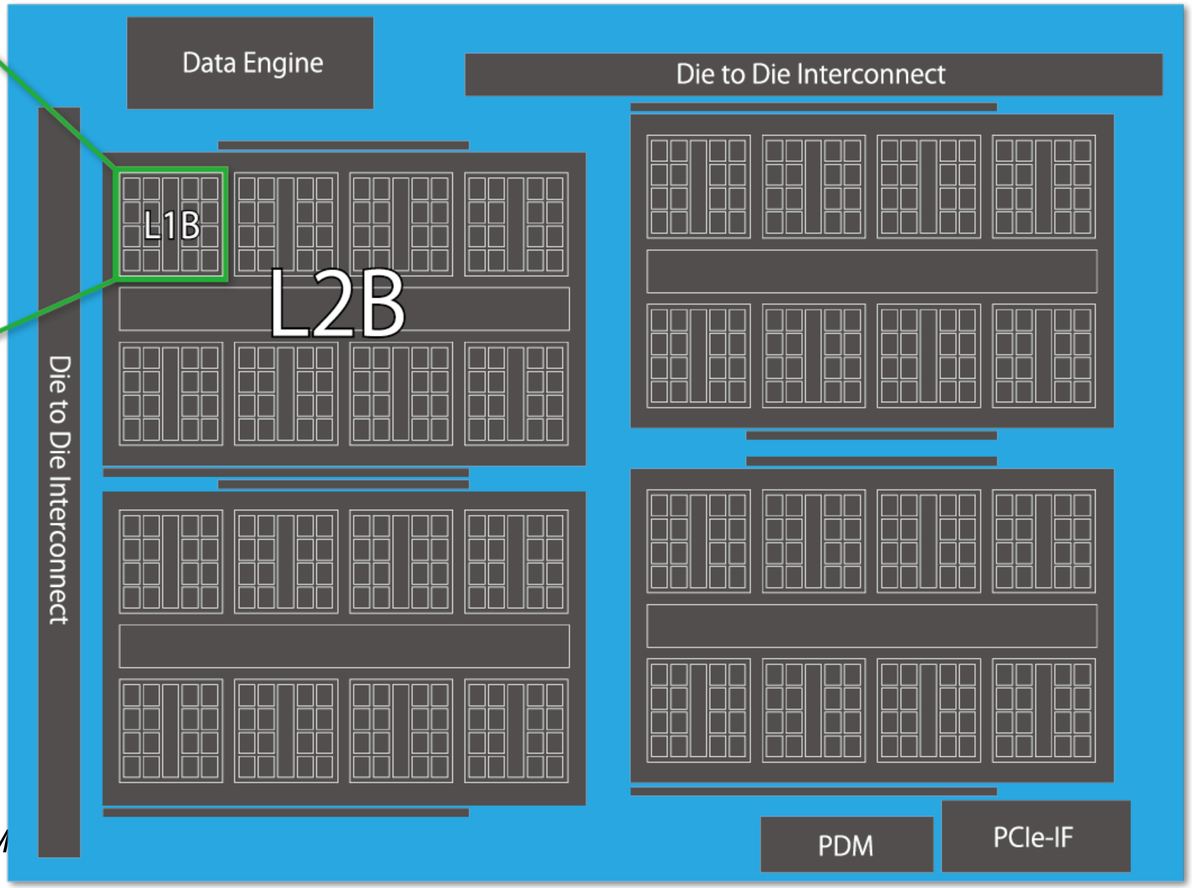
MN-Core architecture



MN-Core Board

4 dies
each die

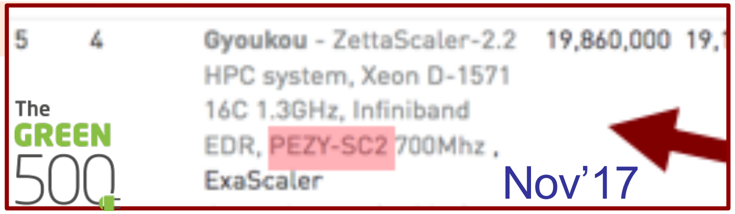




MAU: Matrix Arithmetic Unit
PE: Processing Element
MAB: Matrix Arithmetic Block
L1B: Level 1 Block w/ 16 MABs
L2B: Level 2 Block w/ 128 MABs
1 chip/die with 512 MABs
1 MN-Core package w/ 2048 MABs

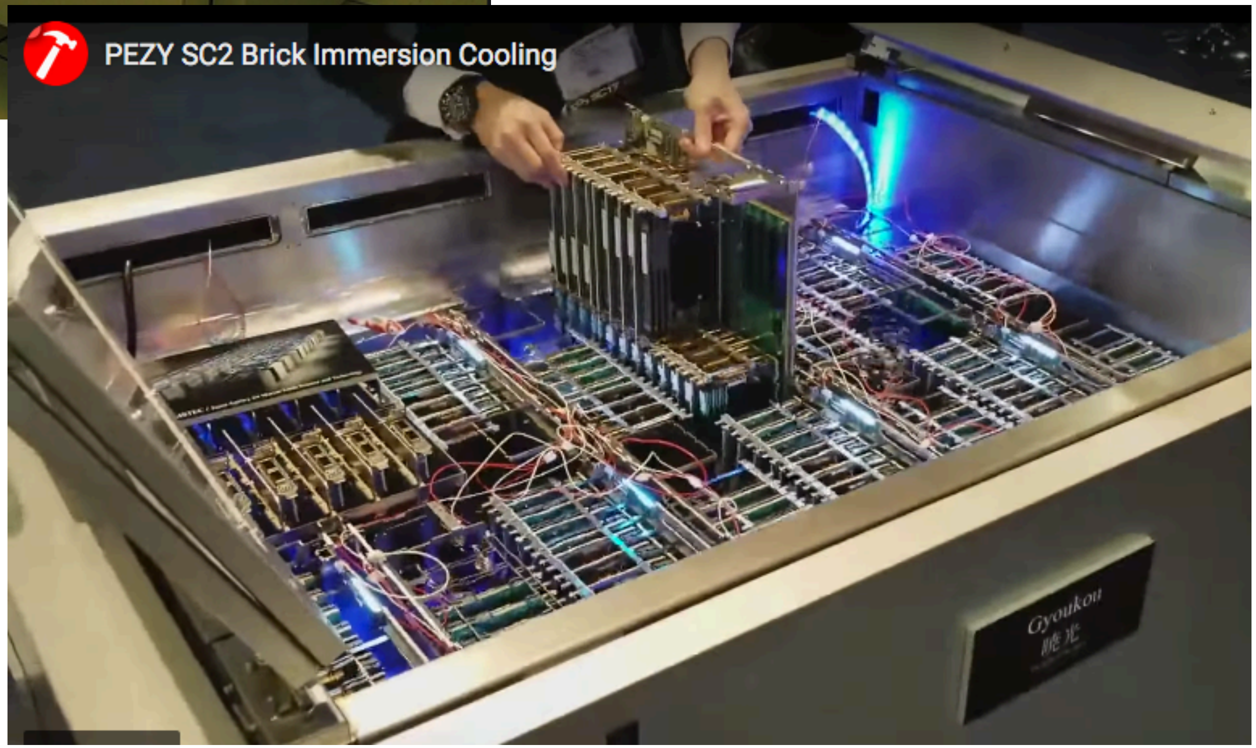
Gyoukou ZettaScaler-2.2

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



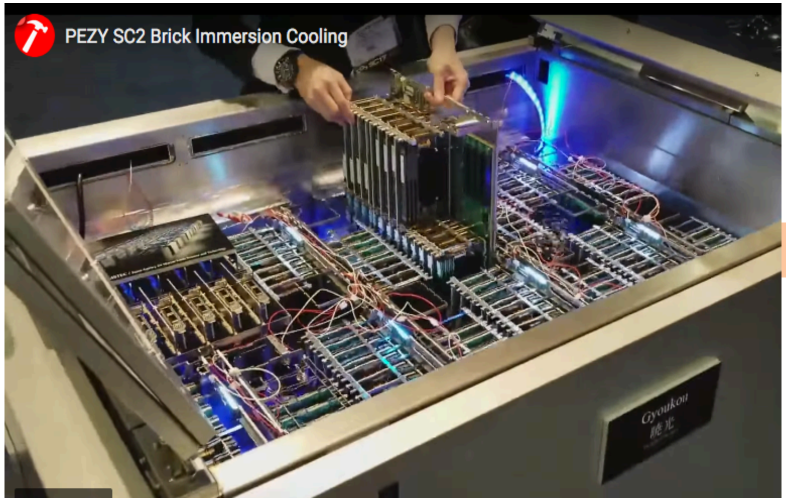
Nov'17

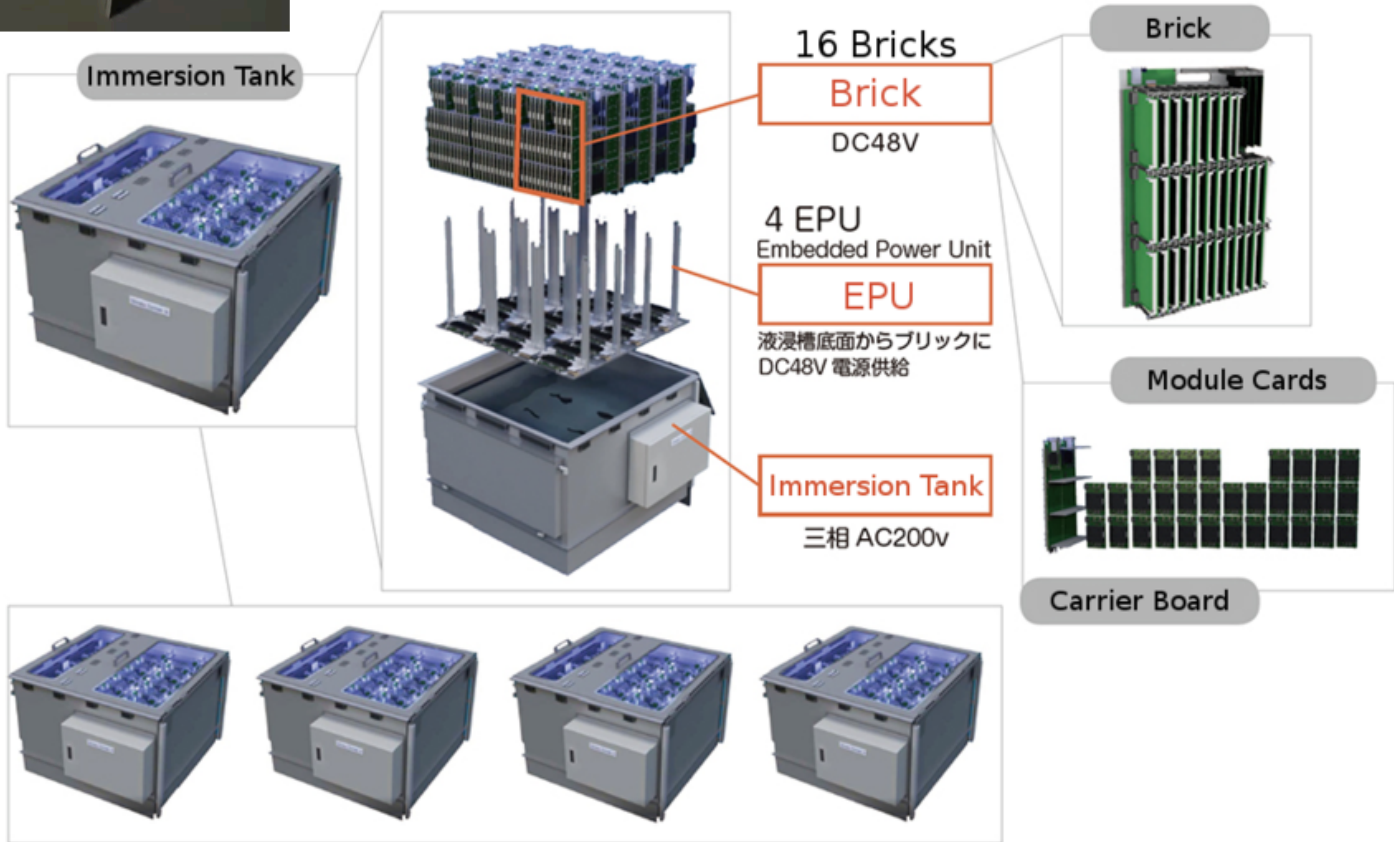
NA-1, #3 in Jun'20 is similar to Gyoukou, but w/less cores



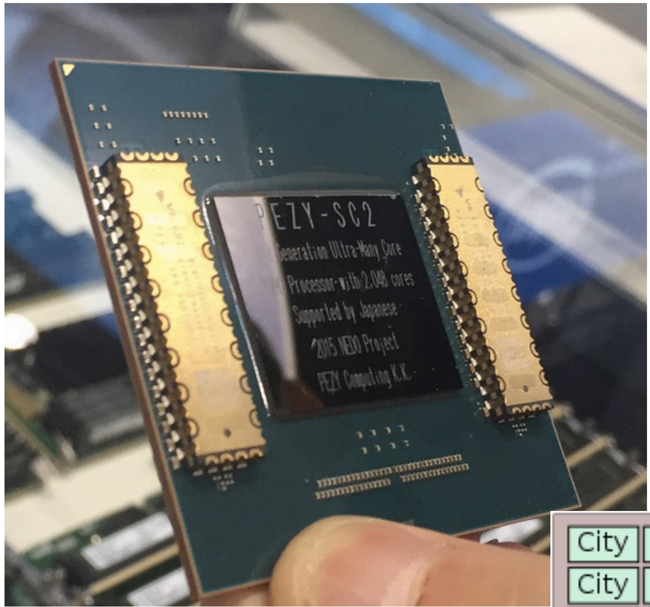
20 immersion tanks
each tank 16 bricks
each brick 32 PEZY
each PEZY
~2K 8-way SMT cores
=>
each tank ~1M cores

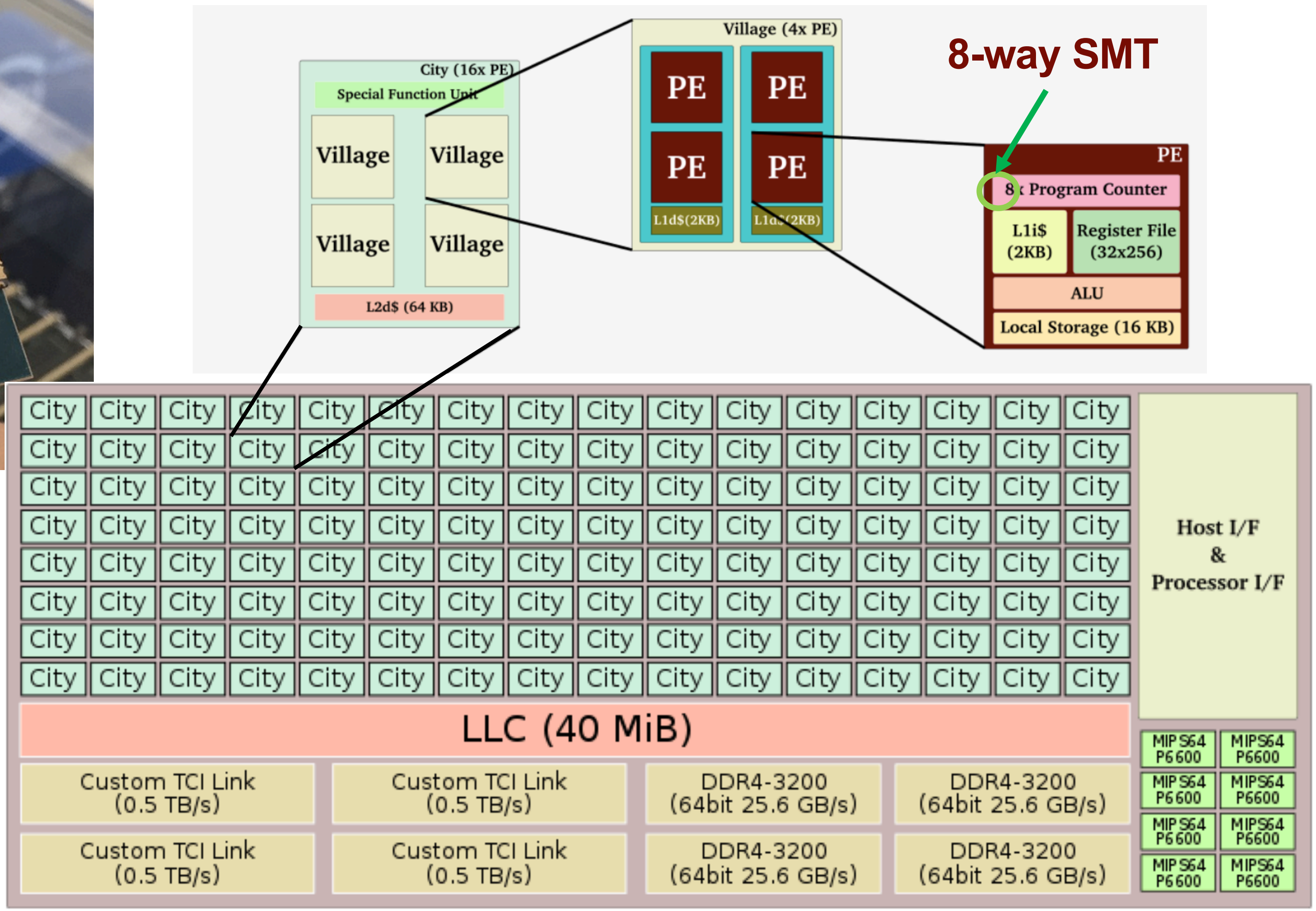
Gyoukou ZettaScaler-2.2





PEZY-SC2 in ZettaScaler-2.2





PEZY-SC2: **2048 cores**
+ 8x MIPS cores *(2017)*

PEZY-SC3: 8192 cores
(due in 2019, but…)

PEZY-SC4: 16384 cores
(due in 2020, but…)

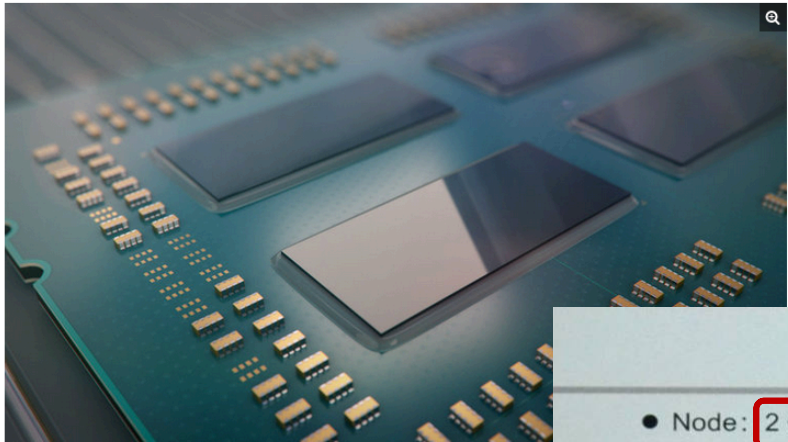
Advanced Computing System (PreE)

(the Hygon Dhyana x86 device in a Sugon cluster)

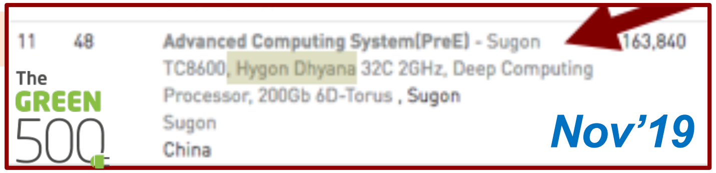
tom's HARDWARE

China Finds Zen: Begins Production Of x86 Processors Based On AMD's IP

by Paul Alcorn July 6, 2018 at 1:18 PM

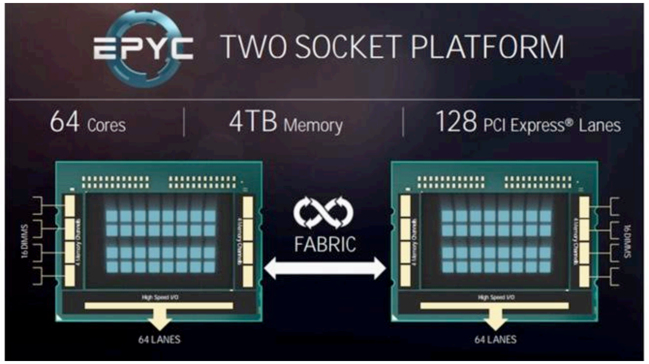


海光 HYGON



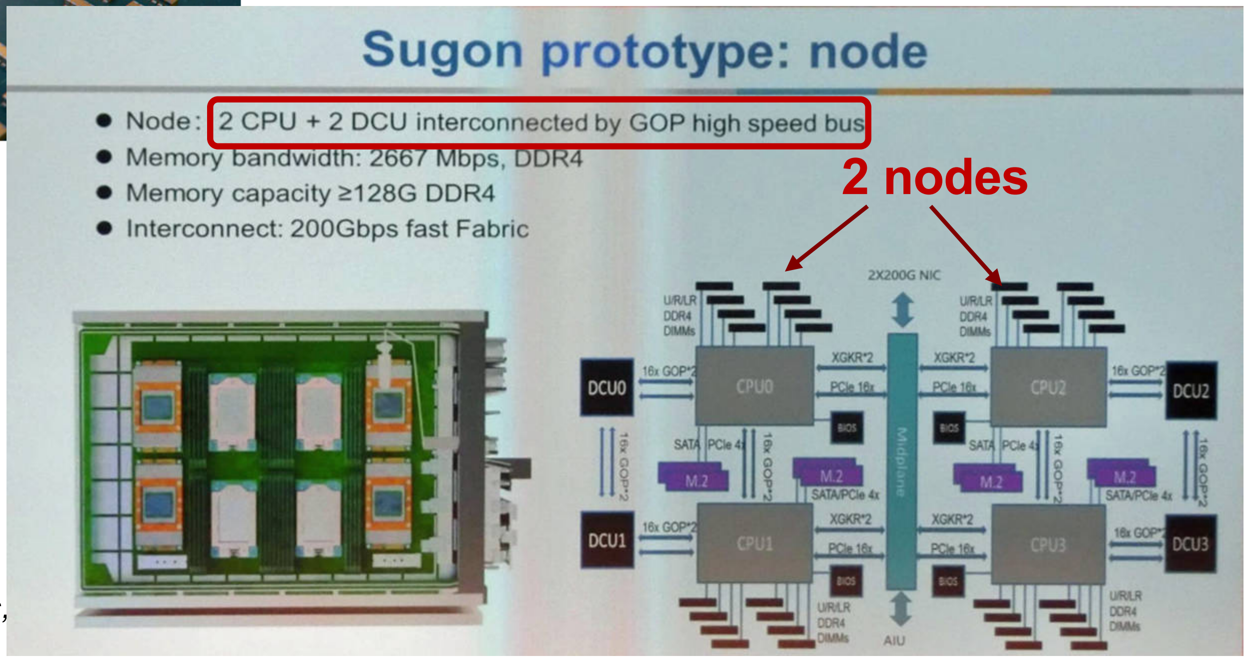
Hygon Dhyana:
32-core 2.0GHz

Similar to:
AMD EPYC 7501



"CPU":
32-core Hygon Dhyana

DCU:
GPU accelerator



Outline

1. TOP500
    a) TOP10 lists from Nov’17 to Nov’20
    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) TOP5 list in Nov’20
    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

TOP500: HPCG vs. HPL

HPCG UPDATE: ISC'17

Jack Dongarra
Michael Heroux
Piotr Luszczek

HPCG Snapshot

- High Performance Conjugate Gradients (HPCG).
- Solves *Ax=b*, *A* large, sparse, *b* known, *x* computed.
- An optimized implementation of PCG contains essential computational and communication patterns that are prevalent in a variety of methods for discretization and numerical solution of PDEs
- Patterns:
  - Dense and sparse computations.
  - Dense and sparse collectives.
  - Multi-scale execution of kernels via MG (truncated) V cycle.
  - Data-driven parallelism (unstructured sparse triangular solves).
- Strong verification (via spectral properties of PCG).

HPCG UPDATE: ISC'17

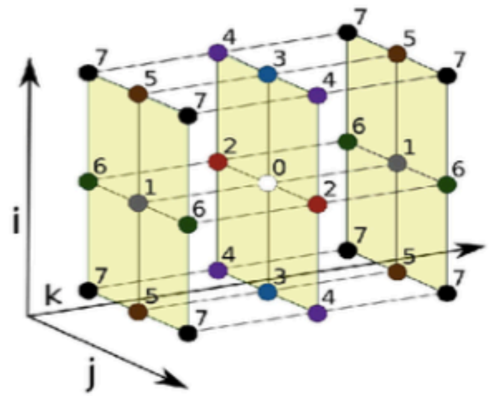
TOP500: HPCG vs. HPL

Jack Dongarra
Michael Heroux
Piotr Luszczek

hpcg-benchmark.org

Model Problem Description

- Synthetic discretized 3D PDE (FEM, FVM, FDM).
- Zero Dirichlet BCs, Synthetic RHS s.t. solution = 1.
- Local domain: $(n_x \times n_y \times n_z)$
- Process layout: $(np_x \times np_y \times np_z)$
- Global domain: $(n_x * np_x) \times (n_y * np_y) \times (n_z * np_z)$
- Sparse matrix:
  - 27 nonzeros/row interior.
  - 8 – 18 on boundary.
  - Symmetric positive definite.



27-point stencil operator

TOP500: HPCG vs. HPL

HPCG UPDATE: ISC'17

Jack Dongarra
Michael Heroux
Piotr Luszczek

hpcg-benchmark.org 5

Merits of HPCG

- Includes major communication/computational patterns.
  - Represents a minimal collection of the major patterns.
- Rewards investment in:
  - High-performance collective ops.
  - Local memory system performance.
  - Low latency cooperative threading.
- Detects/measures variances from bitwise reproducibility.
- Executes kernels at several (tunable) granularities:
  - nx = ny = nz = 104 gives
  - nlocal = 1,124,864; 140,608; 17,576; 2,197
  - ComputeSymGS with multicoloring adds one more level:
    - 8 colors.
    - Average size of color = 275.
    - Size ratio (largest:smallest): 4096
  - Provide a "natural" incentive to run a big problem.
- Full performance discussion:
  - http://www.hpcg-benchmark.org -> "Performance Overview" tab.

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’20
   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’20
   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
5. GRAPH500
   a) Performance Metric (TEPS)
   b) Breadth-First Search (BFS) & Single Source Shortest Paths (SSSP)

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'20)

November 2020

November 2020

Rank	Site	Computer	Cores	HPL-AI (Eflop/s)	TOP500 Rank	HPL Rma (Eflop/s)
1	RIKEN	Fugaku	7,299,072	2.0	1	0.442
2	ORNL	Summit	2,414,592	0.55	2	0.149
3	NVIDIA, USA	Selene	277,760	0.25	5	0.063
4	FZJ	JUWELS_BM	449,280	0.11	7	0.044
5	Nagoya	Flow	110,592	0.03	41	0.007

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

The GRAPH500 benchmark

- Description

  The benchmark used in Graph500 stresses the communication subsystem of the system, instead of counting double precision floating-point.[1] It is based on a breadth-first search in a large undirected graph (a model of Kronecker graph with average degree of 16). There are three computation kernels in the benchmark: the first kernel is to generate the graph and compress it into sparse structures CSR or CSC (Compressed Sparse Row/Column); the second kernel does a parallel BFS search of some random vertices (64 search iterations per run); the third kernel runs a single-source shortest paths (SSSP) computation. Six possible sizes (Scales) of graph are defined: toy (2^{26} vertices; 17 GB of RAM), mini (2^{29}; 137 GB), small (2^{32}; 1.1 TB), medium (2^{36}; 17.6 TB), large (2^{39}; 140 TB), and huge (2^{42}; 1.1 PB of RAM).[4]

- Performance metric: traversed edges per second (TEPS)
- List of benchmark problems
  - Breadth-First Search (BFS)
  - Single-Source Shortest Paths (SSSP)

https://en.wikipedia.org/wiki/Graph500

The GRAPH500 BFS list
(Nov'20)

	RANK	MACHINE	VENDOR	INSTALLATION SITE	LOCATION	COUNTRY	YEAR	NUMBER OF NODES	NUMBER OF CORES	SCALE	GTEPS
1	1	Supercomputer Fugaku	Fujitsu	RIKEN Center for Computational Science (R-CCS)	Kobe Hyogo	Japan	2020	158976	7630848	41	102956
4	2	Sunway TaihuLight	NRCPC	National Supercomputing Center in Wuxi	Wuxi	China	2015	40768	10599680	40	23755.7
19	3	TOKI-SORA	Fujitsu	Japan Aerospace eXploration Agency (JAXA)	Tokyo	Japan	2020	5760	276480	36	10813
2	4	OLCF Summit (CPU-Only)	IBM	Oak Ridge National Laboratory	Oak Ridge TN	United States	2018	2048	86016	40	7665.7
	5	SuperMUC-NG	Lenovo	Leibniz Rechenzentrum	Garching	Germany	2018	4096	196608	39	6279.47
	6	NERSC Cori - 1024 haswell partition	Cray	NERSC/LBNL	DOE/SC /LBNL/NERSC	United States	2017	1024	32768	37	2562.16
6	7	Tianhe-2 (MilkyWay-2)	National University of Defense Technology	Changsha China	Changsha China	China	2013	8192	196608	36	2061.48

https://graph500.org/

The GRAPH500 SSSP list
(Nov'20)

RANK	MACHINE	VENDOR	INSTALLATION SITE	LOCATION	COUNTRY	YEAR	NUMBER OF NODES	NUMBER OF CORES	SCALE	GTEPS
1	SuperMUC-NG	Lenovo	Leibniz Rechenzentrum	Garching	Germany	2018	4096	196608	37	1053.93
2	NERSC Cori - 1024 haswell partition	Cray	NERSC/LBNL	DOE/SC /LBNL/NERSC	United States	2017	1024	32768	36	558.833
3	Nurion	Cray	Korea Institute of Science and Technology Information	Daejeon	Korea Republic Of	2018	1024	65536	36	337.239
4	NERSC Cori - 512 KNL partition	Cray	NERSC/LBNL	DOE/SC /LBNL/NERSC	United States	2017	512	32768	35	229.188
5	Undisclosed Cray XE6	Cray	National Computing Facility	University	United States	2013	512	16384	34	134.173
6	Undisclosed Cray XE6	Cray	National Computing Facility	University	United States	2013	512	8192	31	12.88
7	Xeon Server	Dell	industry	BoiseID	United States	2017	1	40	23	3.09
8	RAIDEN CPU subsystem	Fujitsu	RIKEN AIP	Tokyo	Japan	2016	32	1024	33	2.66

https://graph500.org/

1. TOP500
   a) TOP10 lists from Nov’17 to Nov’20
   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’20
   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
5. GRAPH500
   a) Performance Metric (TEPS)
   b) Breadth-First Search (BFS) & Single Source Shortest Paths (SSSP)