QCD Workshop, 14th Dec 2019 @ R-CCS, Kobe, Japan

Update of “Fugaku” and FLAGSHIP2020 project

Yuetsu Kodama
Architecture Development Team
FLAGSHIP 2020 project
RIKEN Center for Computational Science (R-CCS)

FLAGSHIP2020 Project “Fugaku”

Missions

- Building the Japanese national flagship supercomputer Fugaku (a.k. a post K), and
- Developing wide range of HPC applications, running on Fugaku, in order to solve social and science issues in Japan (application development proj will be over at the end of march)

Overview of Fugaku architecture

Node: Manycore architecture

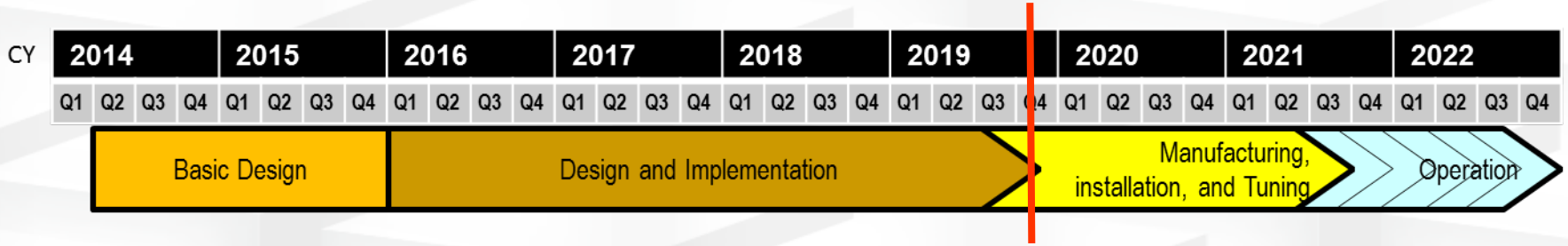
- Armv8-A + SVE (Scalable Vector Extension)
- SIMD Length: 512 bits
- # of Cores: 48 + (2/4 for OS) (3.07 TF / 48 core@2.0GHz)
- Co-design with application developers and high memory bandwidth utilizing on-package stacked memory (HBM2) 1 TB/s B/W
- Low power : 15GF/W (dgemm)

Network: TofuD

- Chip-Integrated NIC, 6D mesh/torus Interconnect

Status and Update

- March 2019: The official contract with Fujitsu to manufacture, ship, and install hardware for Fugaku is done
- RIKEN revealed #nodes > 150K
- March 2019: The Name of the system was decided as “Fugaku”
- Aug. 2019: The K computer stopped the services and shutdown (removed from the computer room)
- Oct 2019: access to the test chips was started.
- Nov. 2019: Fujitsu announce FX1000 and FX700, and business with Cray.
- Nov 2019: Fugaku clock frequency will be 2.0GHz and boost to 2.2 GHz.
- Nov 2019: Green 500 1st position!
- Oct-Nov 2019: MEXT announced the Fugaku “early access program” to begin around Q2/CY2020
- 3rd Dec. 2019: Install of “Fugaku” has begun.



Install of Fugaku has begun 3rd Dec. 2019

You can look movie from home page of Kobe Newspaper:
https://www.kobe-np.co.jp/news/sougou/201912/0012928194.shtml

We will have a tour of Computer room from 13:30 to 14:00 on 13th Dec. at a visitor hall through the window. No photography allowed !!

Green500, Nov. 2019

A64FX prototype –

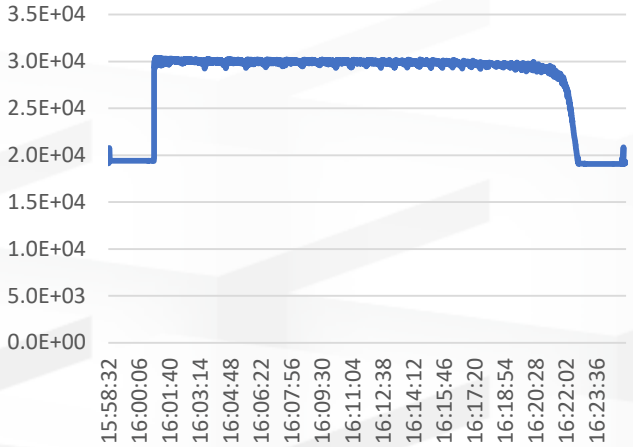
Fujitsu A64FX 48C 2GHz ranked #1 on the list

768x general purpose A64FX CPU w/o accelerators (2 racks)

- 1.9995 PFLOPS @ HPL, 84.75%
- 16.876 GF/W
- Power quality level 2

Ave. of core phase:
118.48 kW / system
Idle power:
46.92 kW / system

Power consumption log of 192-node



https://www.top500.org/green500/lists/2019/11/

KPIs on Fugaku development in FLAGSHIP 2020 project

3 KPIs (key performance indicator) were defined for Fugaku development

- **1. Extreme Power-Efficient System**
  - Maximum performance under Power consumption of 30 - 40MW (for system)
  - Approx. 15 GF/W (dgemm) confirmed by the prototype CPU

    16.8 GF/W with 768 node (2 racks) (micro-fugaku: #1 Green500 Nov 2019)
- **2. Effective performance of target applications**
  - It is expected to exceed 100 times higher than the K computer's performance in some applications
  - 125 times faster in GENESIS (MD application), 120 times faster in NICAM+LETKF (climate simulation and data assimilation) were estimated
- **3. Ease-of-use system for wide-range of users**

Target Application's Performance

- **Performance Targets**
  - 100 times faster than K for some applications (tuning included)
  - 30 to 40 MW power consumption

https://postk-web.r-ccs.riken.jp/perf.html

Predicted Performance of 9 Target Applications

As of 2019/05/14

Area	Priority Issue	Performance Speedup over K	Application	Brief description
Health and longevity	1. Innovative computing infrastructure for drug discovery	**x125+**	**GENESIS**	MD for proteins
	2. Personalized and preventive medicine using big data	**x8+**	**Genomon**	Genome processing (Genome alignment)
Disaster prevention and Environment	3. Integrated simulation systems induced by earthquake and tsunami	**x45+**	**GAMERA**	Earthquake simulator (FEM in unstructured & structured grid)
	4. Meteorological and global environmental prediction using big data	**x120+**	**NICAM+ LETKF**	Weather prediction system using Big data (structured grid stencil & ensemble Kalman filter)
Energy issue	5. New technologies for energy creation, conversion / storage, and use	**x40+**	**NTChem**	Molecular electronic (structure calculation)
	6. Accelerated development of innovative clean energy systems	**x35+**	**Adventure**	Computational Mechanics System for Large Scale Analysis and Design (unstructured grid)
Industrial competitiveness enhancement	7. Creation of new functional devices and high-performance materials	**x30+**	**RSDFT**	Ab-initio program (density functional theory)
	8. Development of innovative design and production processes	**x25+**	**FFB**	Large Eddy Simulation (unstructured grid)
Basic science	9. Elucidation of the fundamental laws and evolution of the universe	**x25+**	**LQCD**	Lattice QCD simulation (structured grid Monte Carlo)

KPIs on Fugaku development in FLAGSHIP 2020 project

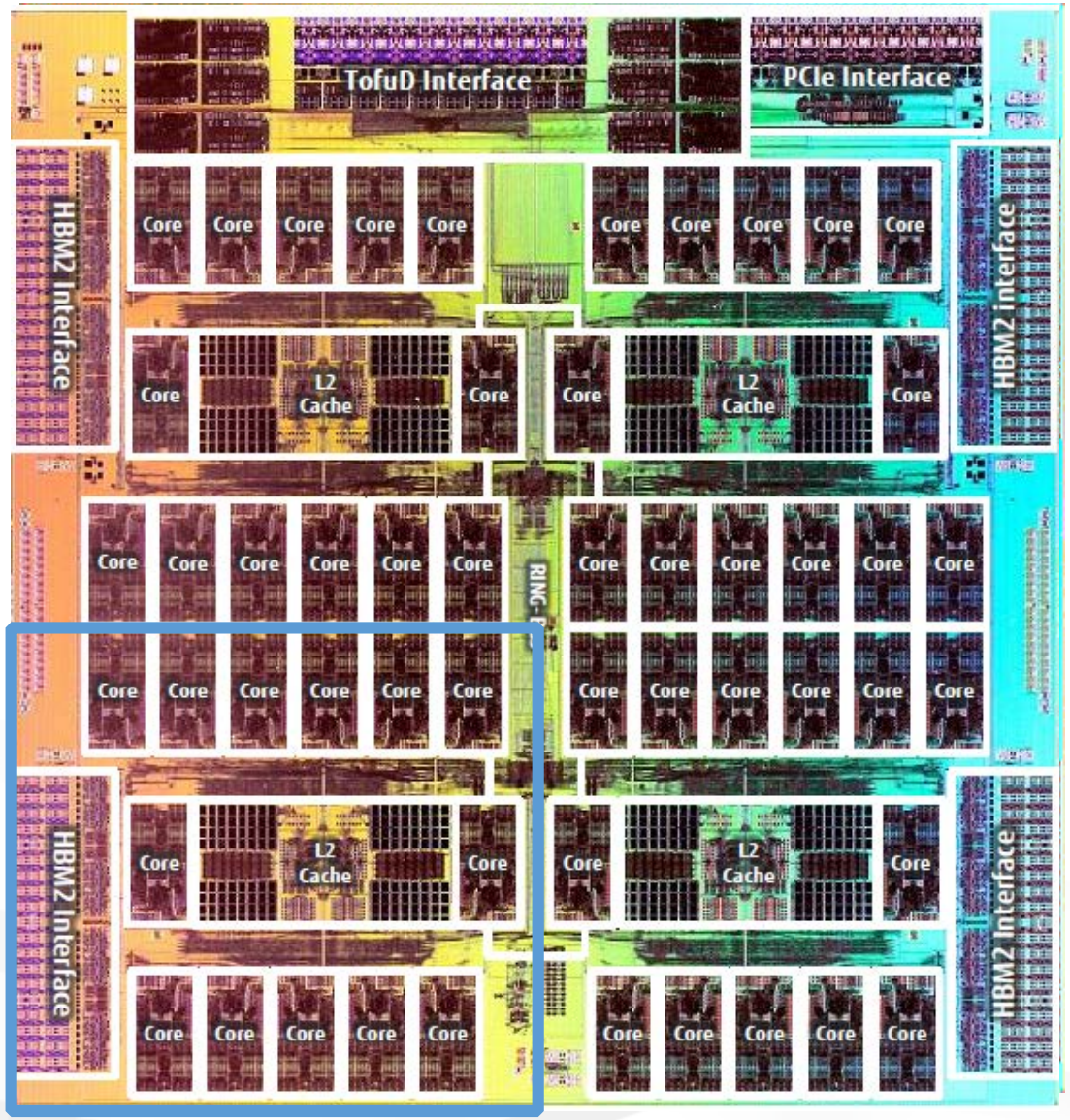
3 KPIs (key performance indicator) were defined for Fugaku development

- **1. Extreme Power-Efficient System**
  - Maximum performance under Power consumption of 30 - 40MW (for system)
  - Approx. 15 GF/W (dgemm) confirmed by the prototype CPU

- **2. Effective performance of target applications**
  - It is expected to exceed 100 times higher than the K computer's performance in some applications
  - 125 times faster in GENESIS (MD application), 120 times faster in NICAM+LETKF (climate simulation and data assimilation) were estimated

- **3. Ease-of-use system for wide-range of users**
  - Shared memory system with high-bandwidth on-package memory must make existing OpenMP-MPI program ported easily.
  - No programming effort for accelerators such as GPUs is required.
  - Co-design with application developers

CPU-Die

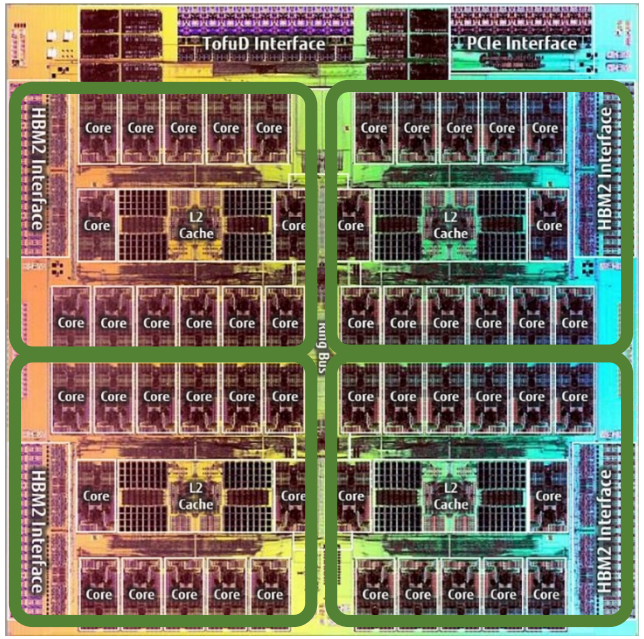


CMG: Core Memory Group

CPU A64FX

Architecture	Armv8.2-A SVE (512 bit SIMD)	
Core	48 cores for compute and 2/4 for OS activities	
	Normal: 2.0 GHz	DP: 3.072 TF, SP: 6.144 TF, HP: 12.288 TF
	Boost: 2.2 GHz	DP: 3.379 TF, SP: 6.758 TF, HP: 13.516 TF
Cache L1	64 KiB, 4 way, 256 GB/s(load), 128 GB/s (store) @ 2.0GHz	
Cache L2	CMG(NUMA): 8 MiB, 16way Node: 4096 GB/s Core: 128 GB/s (load), 64 GB/s (store) @ 2.0GHz	
Memory	HBM2 32 GiB, 1024 GB/s	
Interconnect	TofuD (28 Gbps x 2 lane x 10 port)	
I/O	PCIe Gen3 x 16 lane	
Technology	7nm FinFET	

4 NUMA Nodes



Performance

Stream triad: 830 GB/s (>80% efficiency)

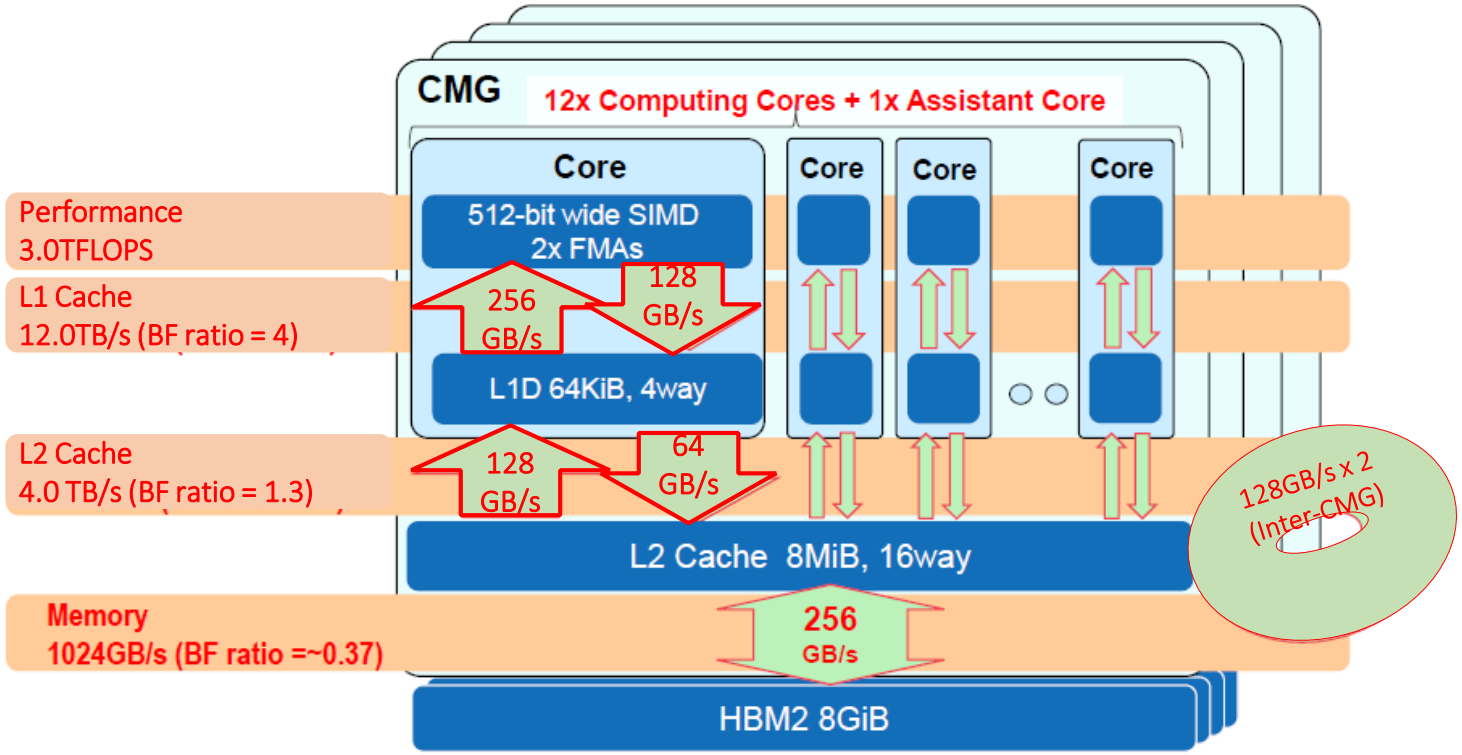
Dgemm: 2.7 TF (>90% efficiency)

ref. Toshio Yoshida, "Fujitsu High Performance CPU for the Post-K Computer," IEEE Hot Chips: A Symposium on High Performance Chips, San Jose, August 21, 2018.

Courtesy of FUJITSU LIMITED

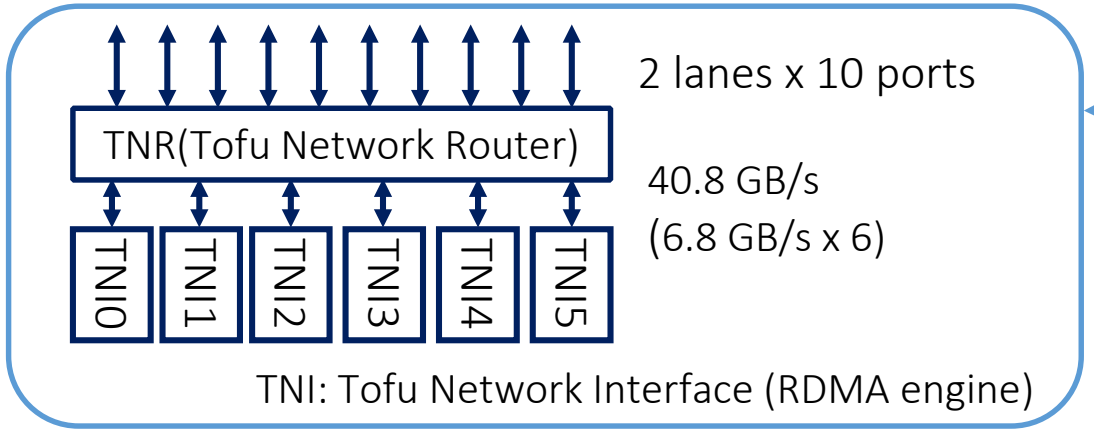
High Bandwidth

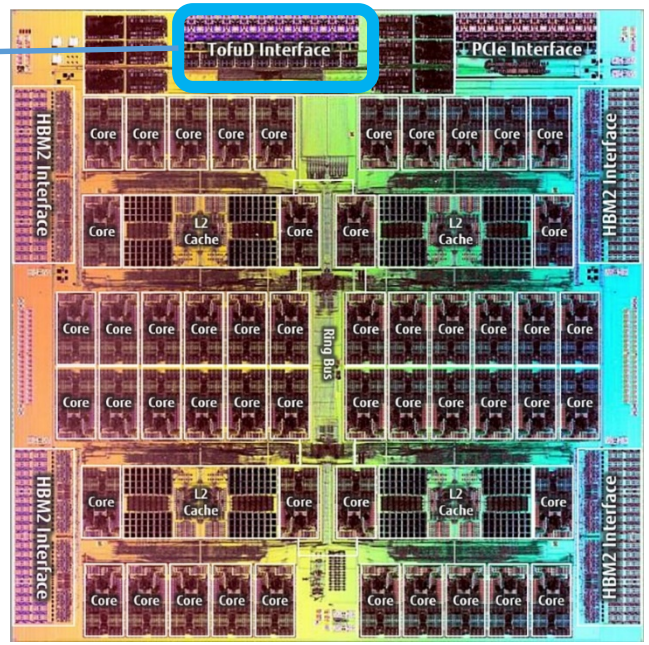
- Extremely high bandwidth in caches and memory
  - A64FX has out-of-order mechanisms in cores, caches and memory controllers. It maximizes the capability of each layer's bandwidth



@2.0GHz

TofuD Interconnect



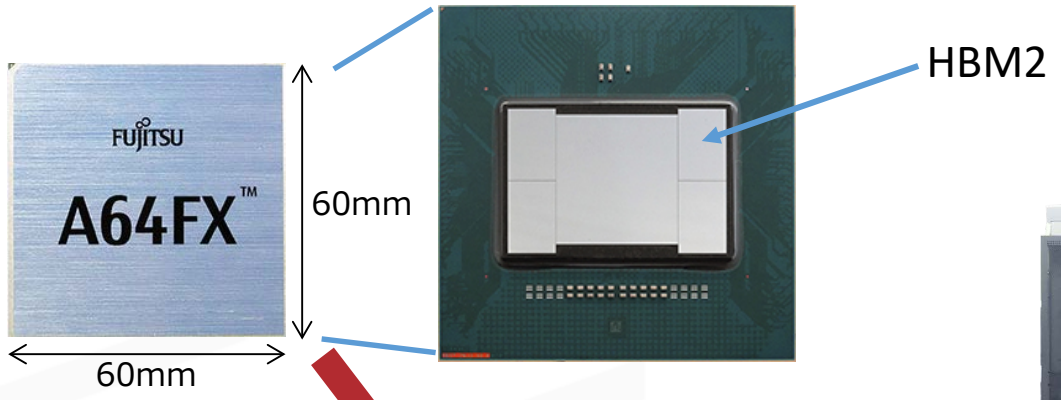


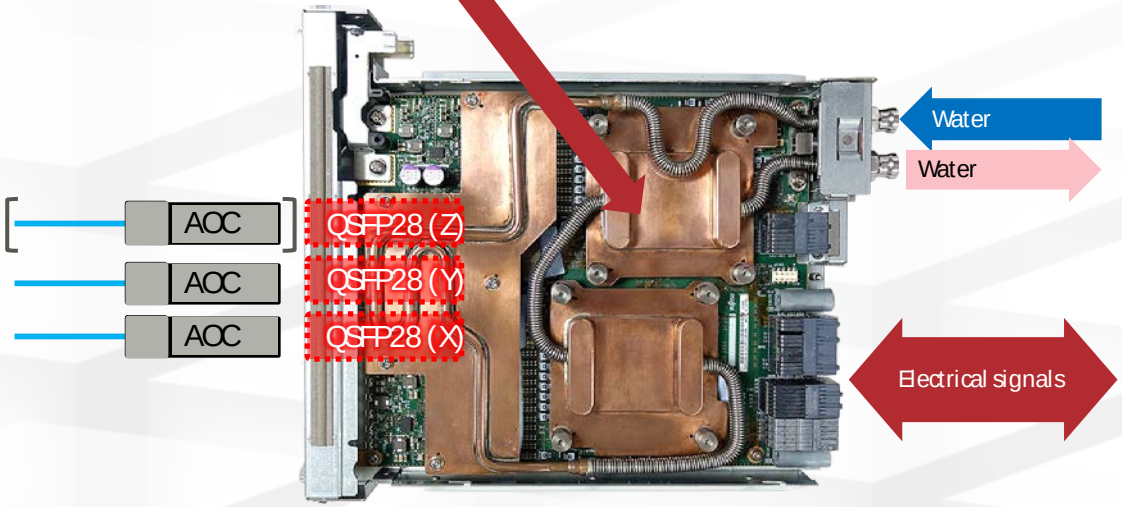
- 6 RDMA Engines
- Hardware barrier support
- Network operation offloading capability

8B Put latency	0.49 – 0.54 usec
1MiB Put throughput	6.35 GB/s

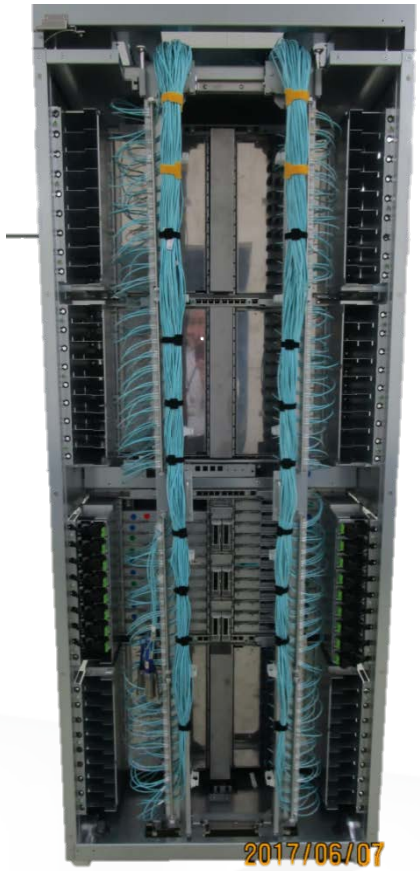
rf. Yuichiro Ajima, et al. , “The Tofu Interconnect D,” IEEE Cluster 2018, 2018.

Fugaku prototype board and rack





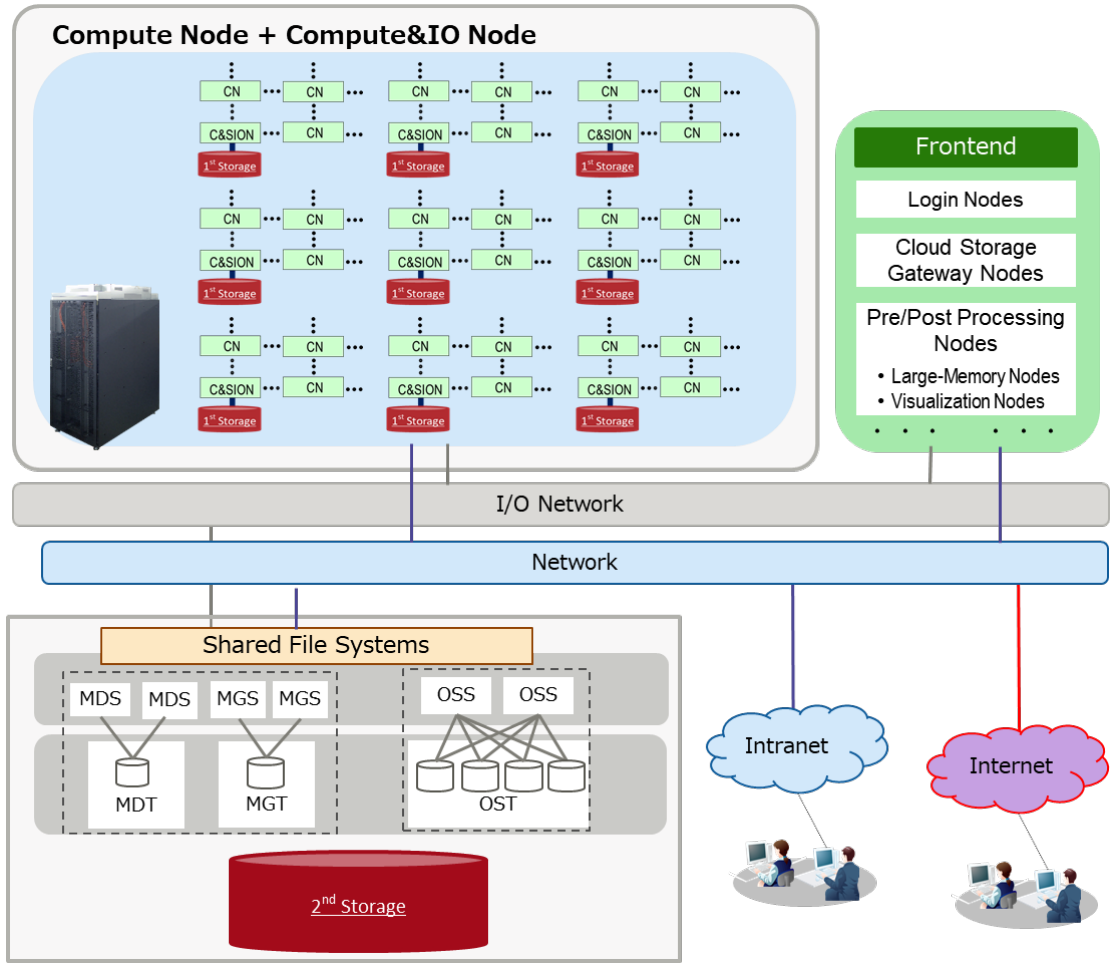
2 CPU / CMU



Shelf: 48 CPUs (24 CMU)
Rack: 8 shelves = 384 CPUs (8x48)

Fugaku System Configuration

- **150k+ node**
- **Two types of nodes**
  - Compute Node and Compute & I/O Node connected by Fujitsu TofuD, 6D mesh/torus Interconnect
- **3-level hierarchical storage system**
  - 1st Layer
    - One of 16 compute nodes, called Compute & Storage I/O Node, has SSD about 1.6 TB
    - Services
      - ~ Cache for global file system
      - ~ Temporary file systems
        - Local file system for compute node
        - Shared file system for a job
  - 2nd Layer
    - Fujitsu FEFS: Lustre-based global file system
  - 3rd Layer
    - Cloud storage services



Advances from the K computer

	K computer	Fugaku	ratio
# core	8	48	
Si tech. (nm)	45	7	
Core perf. (GFLOPS)	16	> 64	4
Chip(node) perf. (TFLOPS)	0.128	>3.0	24
Memory BW (GB/s)	64	1024	
B/F (Bytes/FLOP)	0.5	0.4	
#node / rack	96	384	4
Rack perf. (TFLOPS)	12.3	>1179.6	96
#node/system	82,944	> 150,000	
System perf.(DP PFLOPS)	10.6	> 460.8	43

Si Tech

SVE

CMG&Si Tech

HBM

More than **7.5 M General-purpose cores**!

- SVE increases core performance
- Silicon tech. and scalable architecture (CMG) to increase node performance
- HBM enables high bandwidth

Conclusion

- **3 KPIs (key performance indicator) are achieved in Fugaku**
  - ✓ 1. Extreme Power-Efficient System
  - ✓ 2. Effective performance of target applications
  - ✓ 3. Ease-of-use system for wide-range of users
- **Early access program will begin around Q2/CY2020**
- **We are now evaluating on prototype system**
  - ✓ Keys to get high performance are how to use wide SIMD, as well as how to use many core, and how to use high memory bandwidth with cache hierarchy.