CEA-RIKEN Summer School, 13th June 2019 @ MDLS, Paris

The post-K project and Fujitsu ARM-SVE enabled A64FX processor

for energy-efficiency and sustained application performance

Mitsuhisa Sato Team Leader of Architecture Development Team

Deputy project leader, FLAGSHIP 2020 project

Deputy Director, RIKEN Center for Computational Science (R-CCS)

Professor (Cooperative Graduate School Program), University of Tsukuba

The name of our system (a.k.a post-K) was announced as “Fugaku” (May 23, 2019)

富岳 (Fugaku)

II

Mt. Fuji

FLAGSHIP2020 Project

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

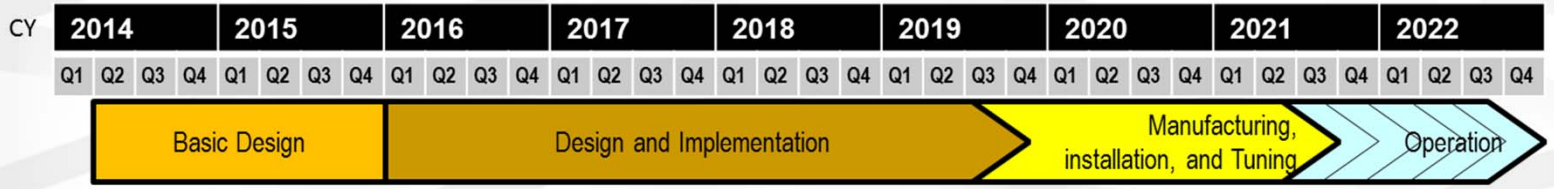
Planned Budget (from 2014FY to 2020FY)

- 110 billion JPY (about 1 billion US$ if 1US$=110JPY, total) includes:
  - Research and development, and manufacturing of the Fugakusystem
  - Development of applications

Project organization

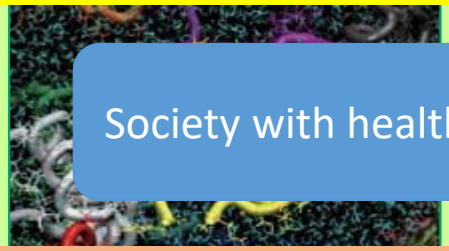
- System development
  - RIKEN is in charge of development
  - Fujitsu is vendor partner.
  - International collaborations: DOE, JLESC, CEA ..

- Applications
  - The government selected 9 social & scientific priority issues and their R&D organizations.
  - Additional projects for Exploratory Issues were selected in Jun 2016



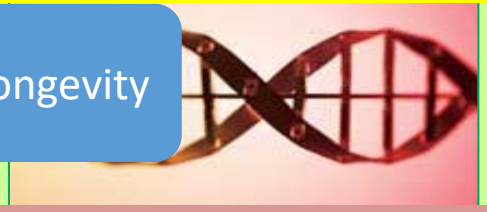
Target science: 9 Priority Issues

①Innovative Drug Discovery



RIKEN Quant. Biology Center

②Personalized and Preventive Medicine



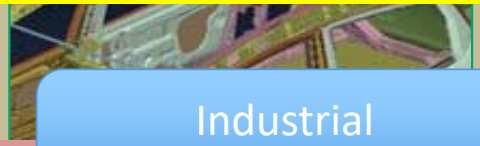
Inst. Medical Science, U. Tokyo

Society with health and longevity

③Hazard and Disaster induced by Earthquake and Tsunami

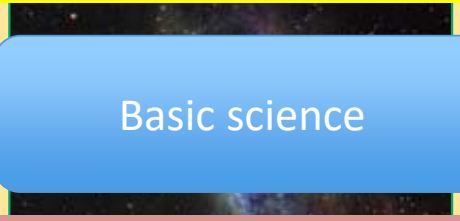
Disaster prevention and global climate

⑧ Innovative Design and Production Processes for the Manufacturing Industry in the Near Future



Industrial competitiveness

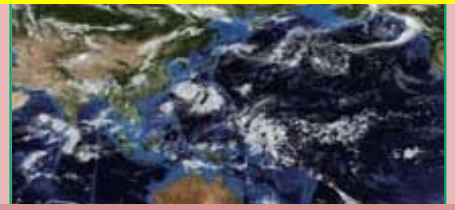
⑨Fundamental Laws and Evolution of the Universe



Cent. for Comp. Science, U. Tsukuba

Basic science

④Environmental Predictions with Observational Big Data



Center for Earth Info., JAMSTEC

⑦New Functional Devices and High-Performance

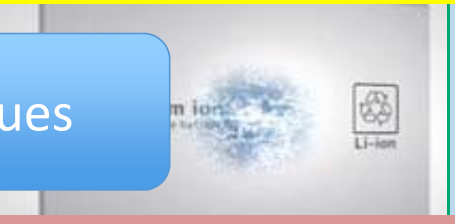
Inst. For Solid State Phys., U. Tokyo

⑥Innovative Clean Energy Systems

Grad. Sch. Engineering, U. Tokyo

Energy issues

⑤High-Efficiency Energy Creation, Conversion/Storage and Use



Inst. Molecular Science, NINS

Target science: Exploratory Issues

Interactive Models of Socio-Economic Phenomena and their Applications

Frontiers of Basic Science - challenge to extremes -

Formation of exo-planets (second Earth) and Environmental Changes of Solar Planets

Mechanisms of Neural Circuits for Human Thoughts and Artificial Intelligence

Projects (more than 10 teams) were selected in Jun 2016

FLAGSHIP2020 Project

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

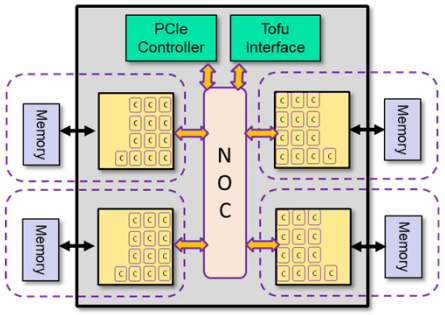
Overview of Fugaku architecture

Node: Manycore architecture

- Armv8-A + SVE (Scalable Vector Extension)
- SIMD Length: 512 bits
- # of Cores: 48 + (2/4 for OS) (> 2.7 TF / 48 core)
- 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

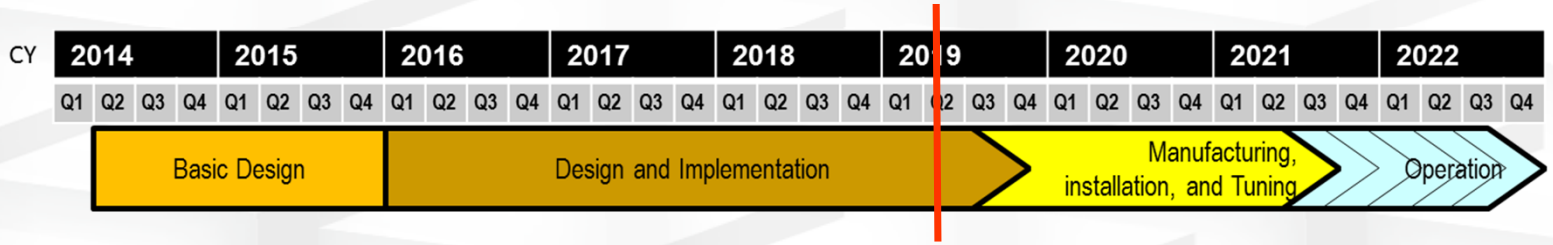


Fujitsu A64FX processor

Prototype board

Status and Update

- "Design and Implementation" completed
- The official contract with Fujitsu to manufacture, ship, and install hardware for Fugaku is done
- RIKEN revealed #nodes > 150K
- The Name of the system was decided as "Fugaku"
- RIKEN announced the Fugaku early access program to begin around Q2/CY2020



Latest Announcement from Fujitsu

Fujitsu Begins Production of Post-K

Also advances productization of commercial units based on the supercomputer technology

Fujitsu Limited

Tokyo, April 15, 2019

Fujitsu Limited today announced that, working with RIKEN, it has completed the design of Post-K, the successor to the K supercomputer. The Ministry of Education, Culture, Sports, Science and Technology (MEXT) is aiming to start the public service of Post-K around 2021 or 2022. Fujitsu has now concluded an official contract with RIKEN to manufacture, ship, and install hardware for Post-K. In addition, Fujitsu will productize a commercial supercomputer using technology created in the Post-K development process, and plans to begin global sales in the second half of fiscal 2019.

The company's efforts in the development of Post-K will be exhibited at Fujitsu Forum 2019, to be held on May 17 at the Tokyo International Forum in Japan.

https://www.fujitsu.com/global/about/resources/news/press-releases/2019/0415-01.html

KPIs on Fugaku development in FLAGSHIP 2020 project

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

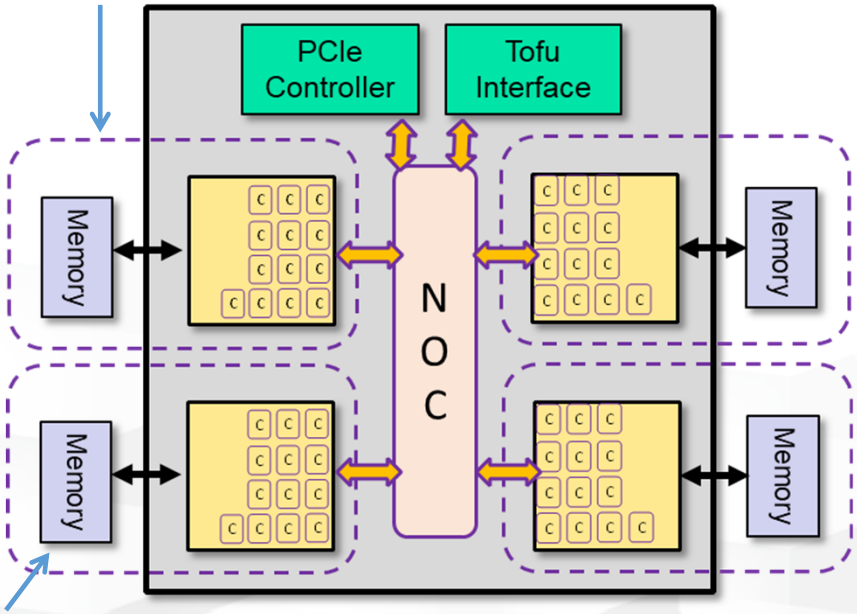
- **1. Extreme Power-Efficient System**
  - 30-40 MW at system level
- **2. Effective performance of target applications**
  - It is expected to exceed 100 times higher than the K computer's performance in some applications
- **3. Easy-of-use system for wide-range of users**

CPU Architecture: A64FX

- **Armv8.2-A (AArch64 only) + SVE (Scalable Vector Extension)**
  - FP64/FP32/FP16 (https://developer.arm.com/products/architecture/a-profile/docs)
- **SVE 512-bit wide SIMD**
- **# of Cores: 48 + (2/4 for OS)**
- Co-design with application developers and high memory bandwidth utilizing on-package stacked memory: **HBM2(32GiB)**
- Leading-edge Si-technology (7nm FinFET), low power logic design **(approx. 15 GF/W (dgemm))**, and power-controlling knobs
- PCIe Gen3 16 lanes
- Peak performance
  - > 2.7 TFLOPS (>90% @ dgemm)
  - Memory B/W 1024GB/s (>80% stream)
  - Byte per Flops: approx. 0.4

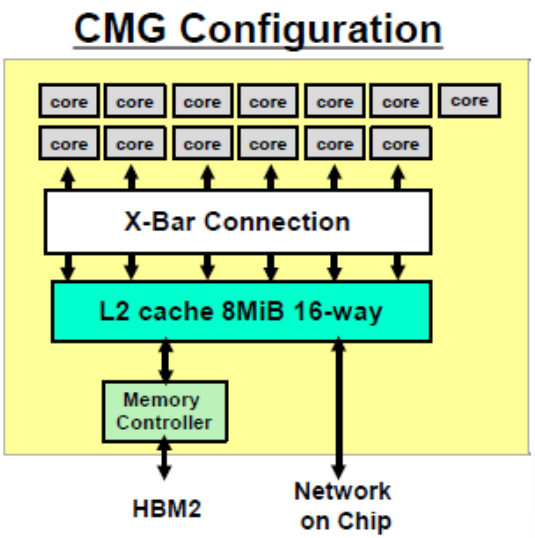
- ◆ "Common" programing model will be to run each MPI process on a NUMA node (CMG) with OpenMP-MPI hybrid programming.
- ◆ 48 threads OpenMP is also supported.

CMG(Core-Memory-Group): NUMA node
12+1 core

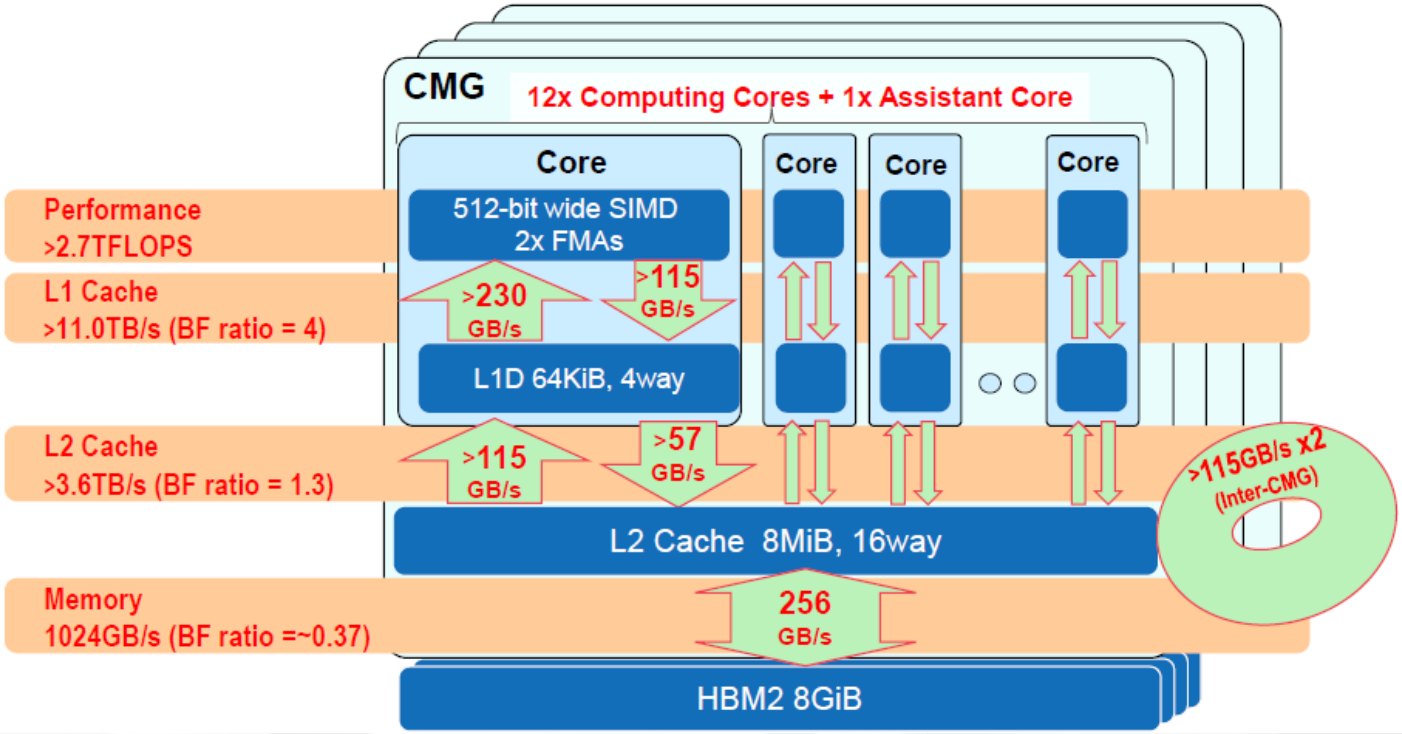


CMG (Core Memory Group)

- CMG: 13 cores (12+1) and L2 cache (8MiB 16way) and memory controller for HBM2 (8GiB)
- X-bar connection in a CMG maximize efficiency for throughput of L2 (>115 GB/s for R, >57 GB/s for W)
- Assistant core is dedicated to run OS demon, I/O, etc
- 4 CMGs support cache coherency by ccNUMA with on-chip directory (> 115GB/s x 2 for inter-CMGs)

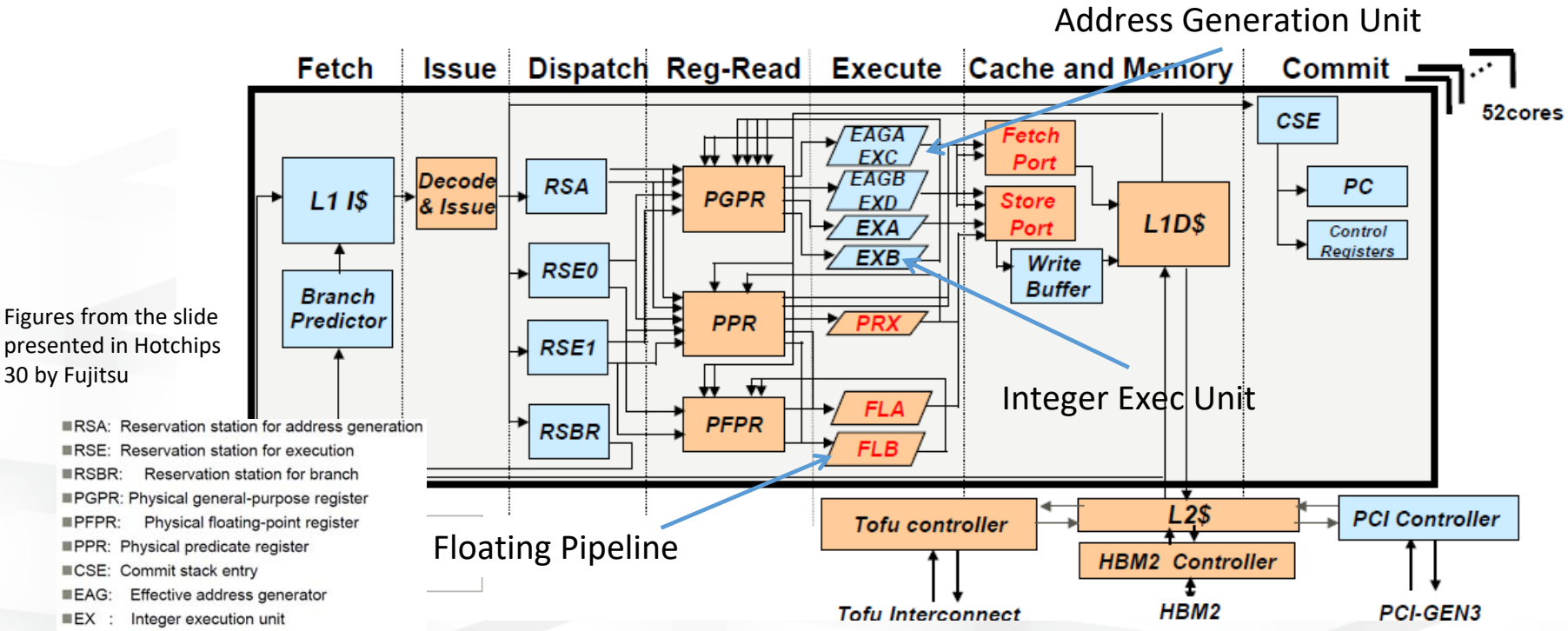


Figures from the slide presented in Hotchips 30 by Fujitsu



FX64A Core Pipeline

- **Superscalar Arch with out-of-order, branch prediction, inherited from Fujitsu SPARC**
- **L1D cache: 64 KiB, 4 ways, “Combined Gather” mechanism on L1**
- **SIMD and predicate operations**
  - 2x 512-bit wide SIMD FMA + Predicate Operation + 4x ALU (shared w/ 2x AGEN)
  - 2x 512-bit wide SIMD load or 512-bit wide SIMD store



Figures from the slide presented in Hotchips 30 by Fujitsu

Tofu interconnect D

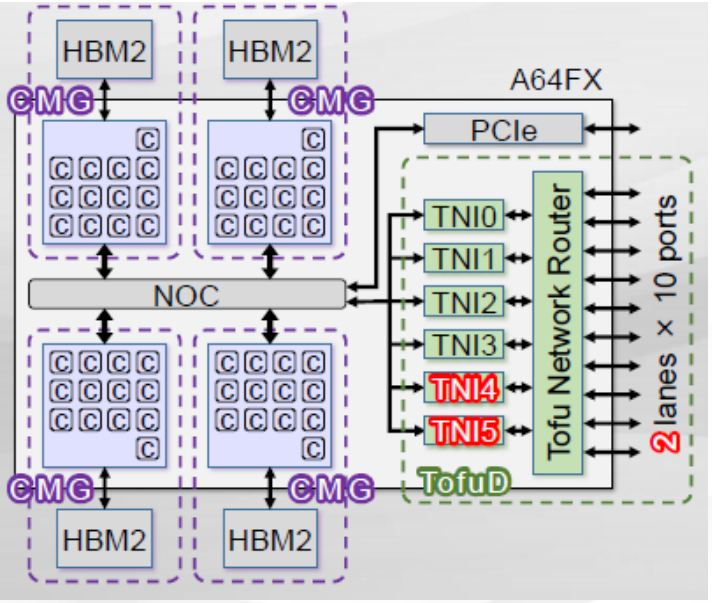
Presented in IEEE Cluster 2018
By Fujitsu

- **Direct network, 6-D Mesh/Torus**
- **28Gbps x 2 lanes x 10 ports (6.8GB/s / link)**
- **Network Interface on Chip**
  - 6 TNIs: Increased TNIs (Tofu Network Interface) achieves higher injection BW & flexible comm. Patterns
  - Memory bypassing achieves low latency

	TofuD spec
Data rate	28.05 Gbps
Link bandwidth	6.8 GB/s
Injection bandwidth	40.8 GB/s

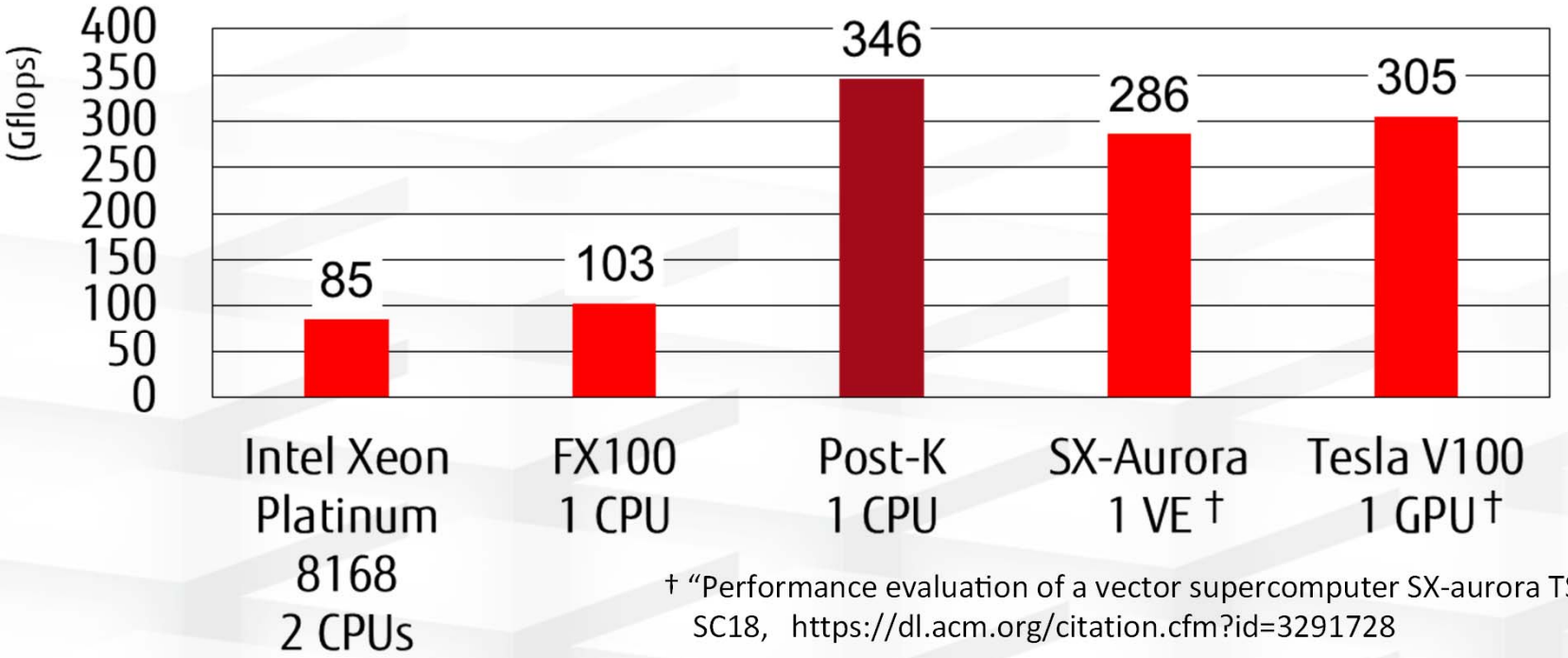
Ref) K computer: Link BW=5.0GB/s, #TNI=4

	Measured
Put throughput	6.35 GB/s
PingPong latency	0.49~0.54 μs



- **HPL & Stream**
  - > 2.5TF / node for dgemm
  - > 830GB/s /node for stream triad

- **Himeno Benchmark (Fortran90)**



† "Performance evaluation of a vector supercomputer SX-aurora TSUBASA", SC18, https://dl.acm.org/citation.cfm?id=3291728

Peak Performance

	Fugaku	K
Peak DP (double precision)	400+ Pflops (x34+)	11.3 Pflops
Peak SP (single precision)	800+ Pflops (x70+)	11.3 Pflops
Peak HP (half precision)	1600+ Pflops (x141+)	--
Total memory bandwidth	150+ PB/sec (x29+)	5.2PB/sec

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)

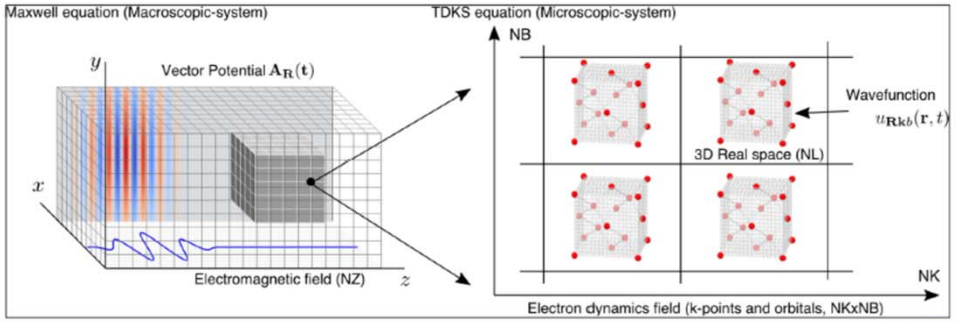
Performance study using Post-K simulator

University of Tsukuba
Center for Computational Sciences

SALMON: Electron Dynamics Simulator

- Main developers: Center for Computational Sciences, U. Tsukuba
- Coupled Maxwell-TDDFT multi-scale simulation
- Open-source application (99% Fortran, 1% C), Apache 2.0 license
  - https://salmon-tddft.jp/

SALMON



- We have been developing a cycle-level simulator for the post-K processor using gem5.
- Collaboration with U. Tsukuba
- Kernel evaluation using single core

	Post-K Simulator	KNL
Execution time [msec]	4.2	5.5
Number of L1D misses	29569	–
L1D miss rate	1.19%	–
Number of L2 misses	20	
L2 miss rate	0.01%	

- 1.3 times faster than KNL per core
- With further optimization (inst. scheduling) exec time reduced to 3.4 msec (1.6 times faster)
- This is the evaluation on L1. OpenMP Multicore execution will be much faster due to HBM memory

3.4 msec by further optimization

Low-power Design & Power Management

- **Leading-edge Si-technology (7nm FinFET)**
- **Low power logic design (15 GF/W @ dgemm)**

- **A64FX provides power management function called "Power Knob"**
  - FL pipeline usage: FLA only, EX pipeline usage : EXA only, Frequency reduction …
  - User program can change "Power Knob" for power optimization
  - "Energy monitor" facility enables chip-level power monitoring and detailed power analysis of applications

- **"Eco-mode" : FLA only with lower "stand-by" power for ALUs**
  - Reduce the power-consumption for memory intensive apps.
  - 4 apps out of 9 target applications select "eco-mode" for the max performance under the limitation of our power capacity (Even using HBM2!)
- **Retention mode: power state for de-activation of CPU with keeping network alive**
  - Large reduction of system power-consumption at idle time

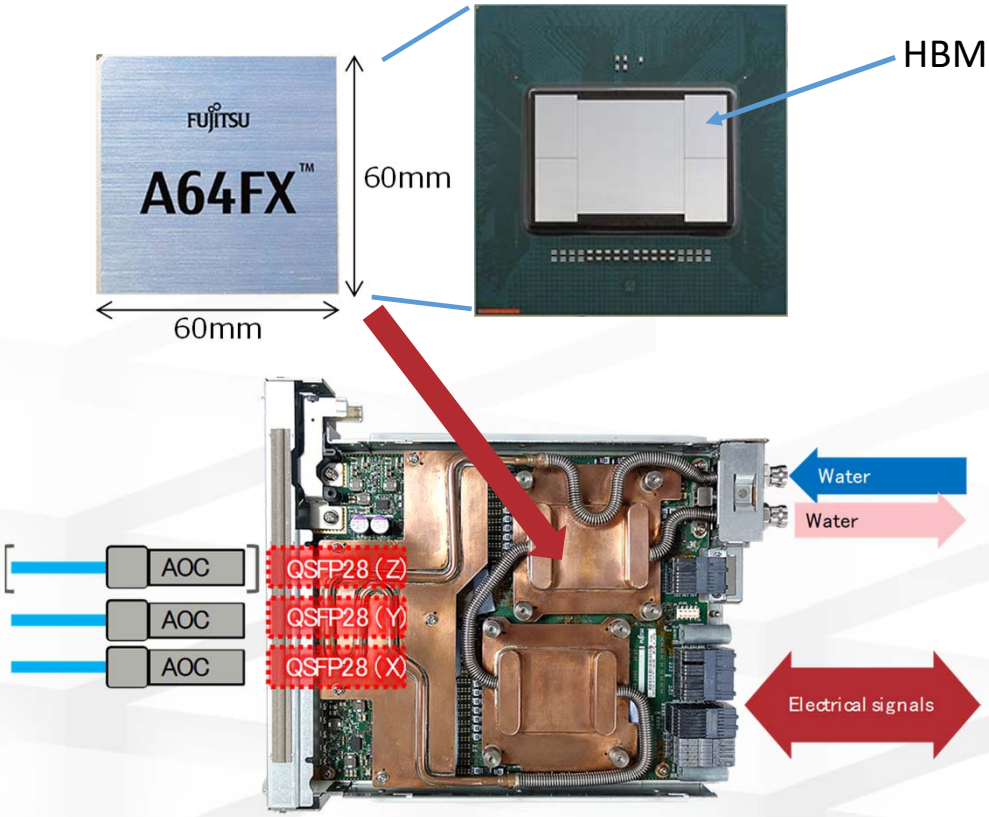
KPIs on Fugaku development in FLAGSHIP 2020 project

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

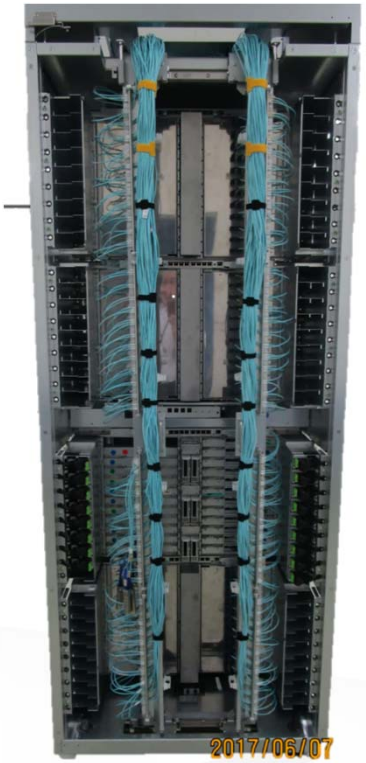
- **1. Extreme Power-Efficient System**
  - Approx. 15 GF/W (dgemm) confirmed by the prototype CPU
  - Maximum performance under Power consumption of 30 - 40MW (for system) will be achieved
- **2. Effective performance of target applications**
  - It is expected to exceed 100 times higher than the K 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. Easy-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

Fugaku prototype board and rack

- **"Fujitsu Completes Post-K Supercomputer CPU Prototype, Begins Functionality Trials", HPCwire June 21, 2018**



2 CPU / CMU



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

Advances from the K computer

	K computer	Fugaku	ratio
# core	8	48	
Si tech. (nm)	45	7	
Core perf. (GFLOPS)	16	> 56	3.5
Chip(node) perf. (TFLOPS)	0.128	>2.7	21
Memory BW (GB/s)	64	1024	
B/F (Bytes/FLOP)	0.5	0.4	
#node / rack	96	384	4
Rack perf. (TFLOPS)	12.3	1036.8	84
#node/system	82,944	> 150,000	
System perf.(DP PFLOPS)	10.6	> 405	38

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

Fugaku Programming Environment

- **Programing Languages and Compilers provided by Fujitsu**
  - Fortran2008 & Fortran2018 subset
  - C11 & GNU and Clang extensions
  - C++14 & C++17 subset and GNU and Clang extensions
  - OpenMP 4.5 & OpenMP 5.0 subset
  - Java
  - GCC, LLVM, and Arm compiler will be also available
- **Parallel Programming Language & Domain Specific Library provided by RIKEN**
  - XcalableMP PGAS Language
  - FDPS (Framework for Developing Particle Simulator)
- **Process/Thread Library provided by RIKEN**
  - PiP (Process in Process)
- **Script Languages provided by Fujitsu**
  - E.g., Python+NumPy, SciPy
- **Communication Libraries**
  - MPI 3.1 & MPI4.0 subset
    - Fujitsu MPI (Based on Open MPI), Riken MPI (Based on MPICH)
  - Low-level Communication Libraries
    - uTofu (Fujitsu), LLC(RIKEN)
- **File I/O Libraries provided by RIKEN**
  - pnetCDF, DTF, FTAR
- **Math Libraries**
  - BLAS, LAPACK, ScaLAPACK, SSL II (Fujitsu)
  - EigenEXA, Batched BLAS (RIKEN)
- **Programming Tools provided by Fujitsu**
  - Profiler, Debugger, GUI

OSS Application Porting @ Arm HPC Users Group

FUJITSU

(http://arm-hpc.gitlab.io/)

Application	Lang.	GCC	LLVM	Arm	Fujitsu
LAMMPS	C++	Modified	Modified	Modified	Modified
GROMACS	C	Modified	Modified	Modified	Modified
GAMESS*	Fortran	Modified	Modified	Modified	Modified
OpenFOAM	C++	Modified	Modified	Modified	Modified
NAMD	C++	Modified	Modified	Modified	Modified
WRF	Fortran	Modified	Modified	Modified	Modified
Quantum ESPRESSO	Fortran	Ok in as is	Ok in as is	Ok in as is	Modified
NWChem	Fortran	Ok in as is	Modified	Modified	Modified
ABINIT	Fortran	Modified	Modified	Modified	Modified
CP2K	Fortran	Ok in as is	Issues found	Issues found	Modified
NEST*	C++	Ok in as is	Modified	Modified	Modified
BLAST*	C++	Ok in as is	Modified	Modified	Modified

“PostK” performance evaluation environment

- **RIKEN is constructing “PostK” performance evaluation environment for application programmers to evaluate and estimate the performance of their applications on “PostK” and for performance turning for “postK”.**
- **The “PostK” performance evaluation environment is available on the servers installed in RIKEN. The environment includes the following tools and servers:**
  - A small-scale FX100 system and “postK” performance estimation tool:

    The estimation tool gives the performance estimation of multithreaded programs on “postK” from the profile data taken on FX100.
  - “PostK” processor simulator based on GEM-5:

    “PostK” processor simulator will give a detail performance results including estimated executing time, cache-miss, the number of instruction executed in O3. The user can understand how the compiled code for SVE is executed on “postK” processor for optimization. (Arm released GEM-5 beta0 of SVE)

    FP16 SVE will be available soon.
  - Compilers for “PostK” processor
    - Fujitsu Compilers : Fortran, C, C++. Fully-tuning for “postK” architecture.
    - Arm Compiler : LLVM-based compiler to generate code forArmv8-A + SV. C,C++ by Clang, Fortran by Flang
  - SVE emulator on Arm server, developed by Arm for fast SVE code execution.
  - Arm Severs (HPE Appollo70, Cavium ThX2)

Fugaku CPU New Innovations: Summary

1. Ultra high bandwidth using on-package memory & matching CPU core

- Recent studies show that majority of apps are memory bound, some compute bound but can use lower precision e.g. FP16
- Comparison w/mainstream CPU: much faster FPU, almost order magnitude faster memory BW, and ultra high performance accordingly
- Memory controller to sustain massive on package memory (OPM) BW: difficult for coherent memory CPU, first CPU in the world to support OPM

2. Very Green e.g. extreme power efficiency

- Power optimized design, clock gating & power knob, efficient cooling
- Power efficiency much better than CPUs, comparable to GPU systems

3. Arm Global Ecosystem & SVE contribution

- Annual processor production: x86 3-400mil, ARM 21bil, (2~3 bil high end)
- Rapid upbringing HPC&IDC Ecosystem (e.g. Cavium, HPE, Sandia, Bristol,…)
- SVE(Scalable Vector Extension) -> Arm-Fujitsu co-design, future global std.

4. High Performance on Society5.0 apps including AI

- Next gen AI/ML requires massive speedup => high perf chips + HPC massive scalability across chips
- Fujitsu A64FX processor: support for AI/ML acceleration e.g. Int8/FP16+fast memory for GPU-class convolution, fast interconnect for massive scaling
- Top performance in AI as well as other Society 5.0 apps

Thank you for your attention!
Q & A