FUJITSU

shaping tomorrow with you

POST-K SUPERCOMPUTER DEVELOPMENT

Toshiyuki Shimizu

Vice President, System Development Division

Next Generation Technical Computing Unit

March 12th, 2019

Agenda

FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

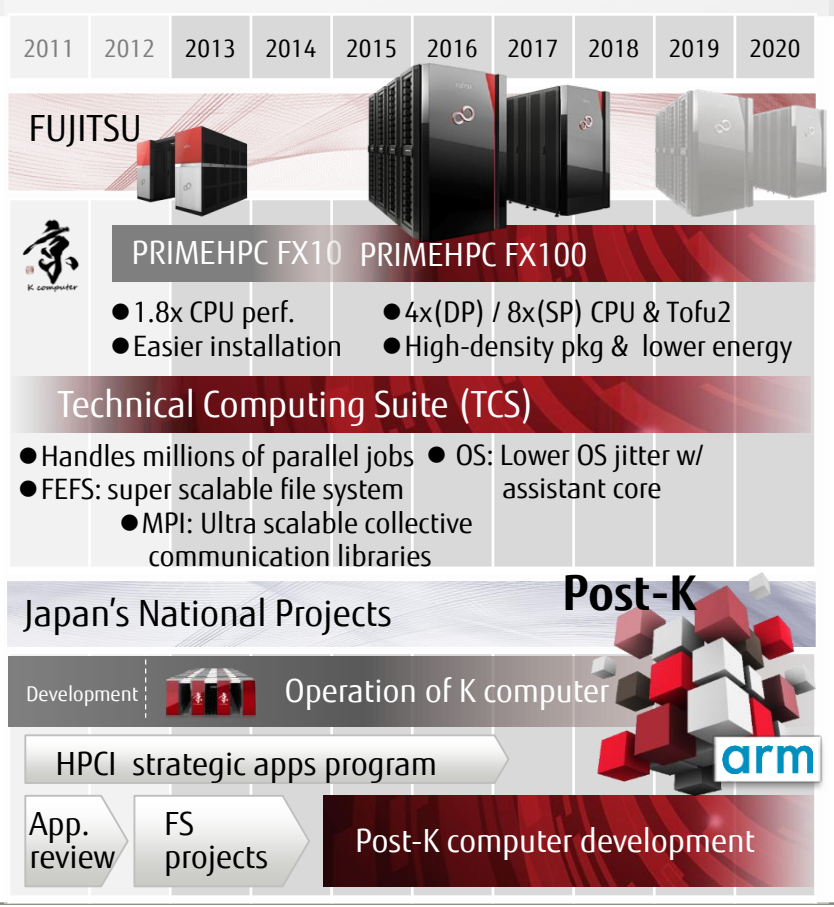
Agenda

FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

K computer, PRIMEHPC, and Post-K

FUJITSU



- **K computer, PRIMEHPC FX in operation**
  - Many applications are running and being developed for science and various industries
  - CPUs and interconnect inherit the K computer architectural concept, featuring state-of-the-art technologies
  - System software TCS supports hardware with newly developed technologies
- **Post-K**
  - RIKEN and Fujitsu are working together for Post-K
  - OS is running and design verification is proceeding as scheduled

K computer is passing a baton on to the Post-K

FUJITSU

- Stable & uninterrupted shared use of the K computer from Sep. 2012 is scheduled to end on Aug. 16, 2019
- Now, the name of the post-K is being called until April 8, 2019 at 5 pm

Call for proposals for the name of the post-K

The RIKEN Center for Computational Science (R-CCS) is calling for proposals for the name of the successor to the K computer (often referred to as the post-K computer), which has been under development with the target to start providing shared use service around the year 2021.

https://www.r-ccs.riken.jp/en/topics/naming.html

Shared use of the K computer will end in August

The K computer has been used by numerous academic and industrial users since it was made available for shared use in September 2012. Now, the service is scheduled to end on August 16, 2019. After necessary preparation, the K computer will be shut down by the end of August.

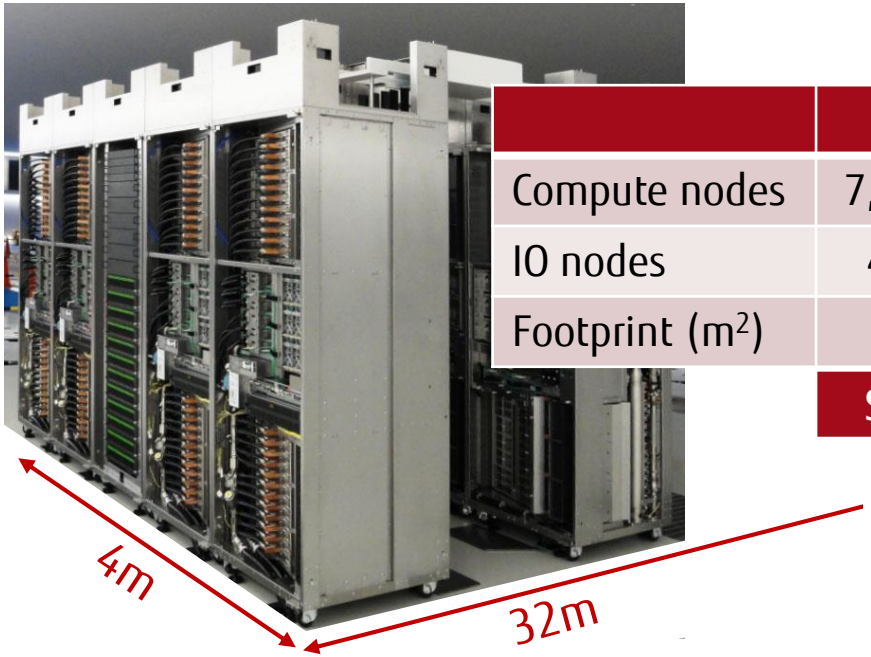
The development of the post-K computer, slated to succeed the K computer, is proceeding well towards the goal of starting shared use around 2021. All involved are working tirelessly to meet the expectation, so please stay tuned!

https://www.r-ccs.riken.jp/en/topics/20190131.html

1 Peta FLOPS by K computer & Post-K

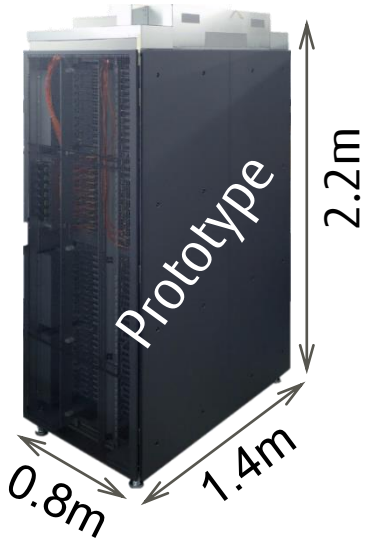
FUJITSU

- K computer
  - 80x compute racks & 20x disk racks
- Post-K
  - 1x rack w/ SSDs



	K computer	Post-K
Compute nodes	7,680(=96x80)	384
IO nodes	4,80(=6x80)	
Footprint (m^2)	128(=4x32)	1.1
	SPARC Linux	Arm Linux

More applications as well as system software in collaboration with Open Source Community



Agenda

FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

Post-K goals and approaches

FUJITSU

- Post-K goals
  - High application performance and good power efficiency
  - Good usability and better accessibility for users
  - Keeping application compatibility while advancing from predecessors
- Our approaches
  - Develops a high-performance and scalable custom CPU

    【Performance】 Wider SIMD & mathematical acc. primitives, high memory BW

    【Scalability】 Scalable interconnect “Tofu”

    【Power efficiency】 The best device tech, power control functions, optimal resources
  - Adopts Arm standard architecture for binary compatibility
  - Achieves ideal performance balance and supports advanced features

Agenda

FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

Fujitsu's new Arm CPU: A64FX

- A64FX approach
- Specs and technologies
- Tofu interconnect D
- Power management
- A64FX summary

A64FX: Approach

FUJITSU

- High performance Arm CPU
  - Develops original CPU core supporting Arm SVE (Scalable Vector Extension)
- Targeting high performance servers
  - Floating point calculations
  - High memory bandwidth
  - Low power consumption
  - Scalable performance and configuration
- Extension for emerging applications
  - Half precision (FP16) & INT16/8 partial dot product support

High BW/Calc. is the key for Real apps.

A64FX: Specs & technologies

FUJITSU

- CPU generations and key parameters

CPU	SPARC64 VIIIfx	SPARC64 IXfx	SPARC64 XIfx	A64FX
1st system w/ CPU	K	FX10	FX100	Post-K
Si tech. (nm)	45	40	20	7
Core perf. (GFLOPS)	16	14.8	34	57+
# of cores	8	16	32	48
Chip perf. (TFLOPS)	0.13	0.24	1.1	2.7+
Memory BW (GB/s)	64	85	480	1024
B/F (Bytes/FLOP)	0.5	0.4	0.4	0.4

SVE

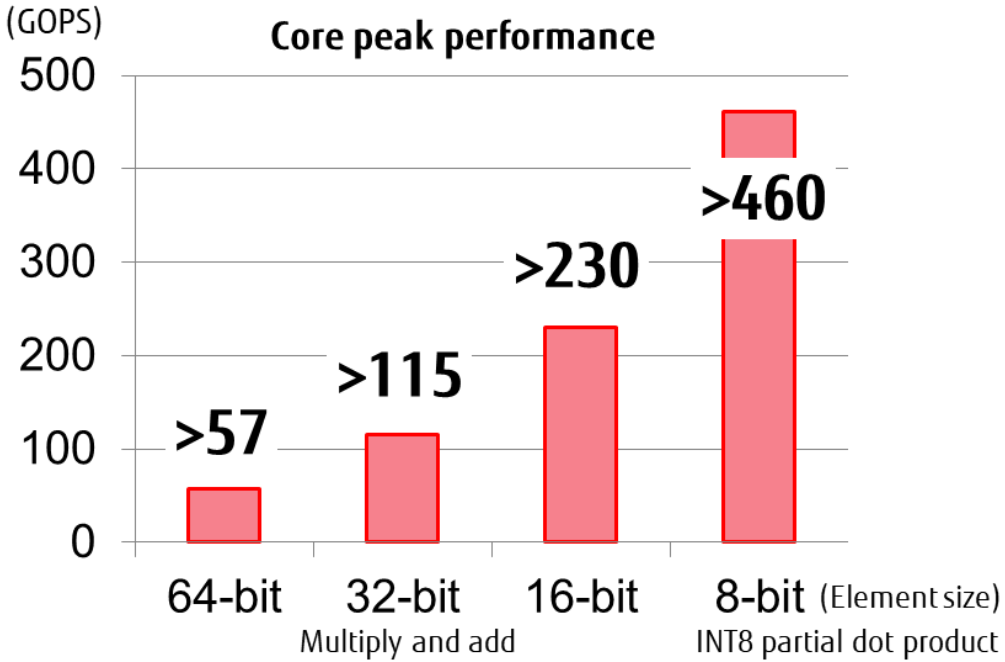
CMG

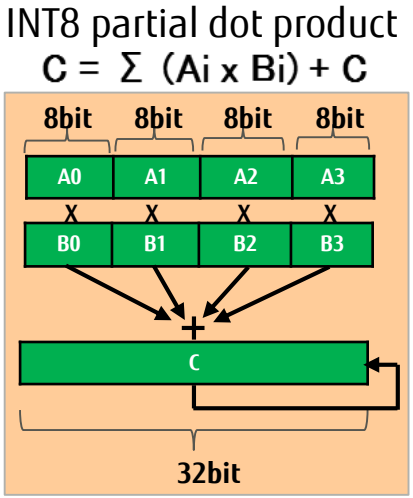
HBM

- Combination of technologies
  - SVE increases core performance
  - CMG is a scalable architecture to increase # of cores
  - HBM enables high memory bandwidth

A64FX technologies: Core performance

- High calc. throughput of Fujitsu's original CPU core w/ SVE
  - 512-bit wide SIMD x 2 pipelines and new integer functions

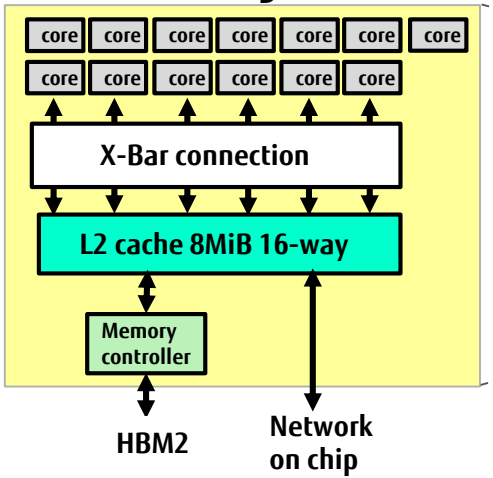




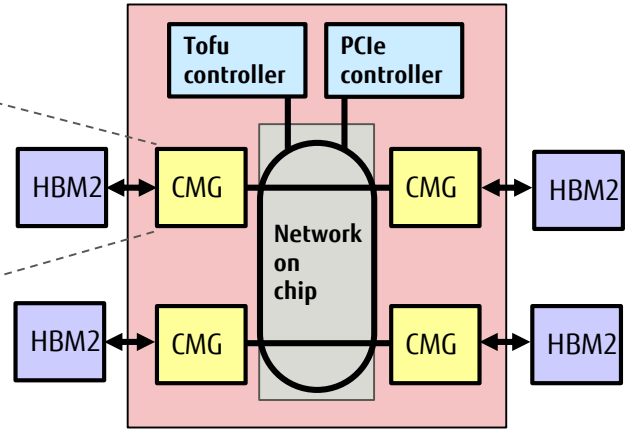
A64FX technologies: Scalable architecture

- Core Memory Group (CMG)
  - 12 compute cores for computing and an assistant core for OS daemon, I/O, etc.
  - Shared L2 cache
  - Dedicated memory controller
- Four CMGs maintain cache coherence w/ on-chip directory
  - Threads binding within a CMG allows linear speed up of cores' performance

CMG configuration

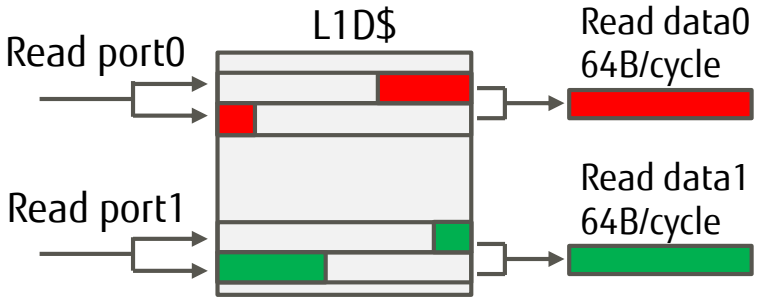


A64FX chip configuration

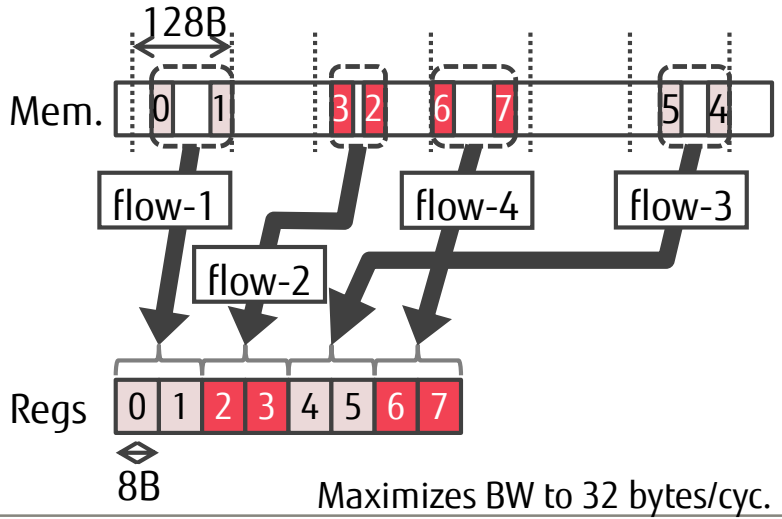


A64FX: L1D cache uncompromised BW

- 128B/cycle sustained BW even for unaligned SIMD load



- "Combined Gather" doubles gather (indirect) load's data throughput, when target elements are within a "128-byte aligned block" for a pair of two regs, even & odd

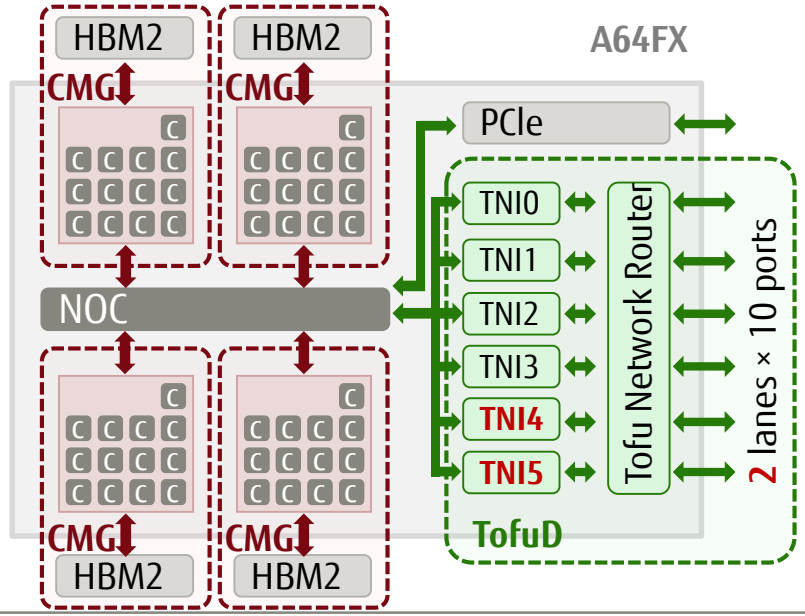


Maximizes BW to 32 bytes/cyc.

A64FX: Tofu interconnect D

- Integrated w/ rich resources
  - Increased TNIs achieves higher injection BW & flexible comm. patterns
  - Increased barrier resources allow flexible collective comm. algorithms
- Memory bypassing achieves low latency
  - Direct descriptor & cache injection

	TofuD spec
Port bandwidth	6.8 GB/s
Injection bandwidth	40.8 GB/s
	Measured
Put throughput	6.35 GB/s
Ping-pong latency	0.49~0.54 μs



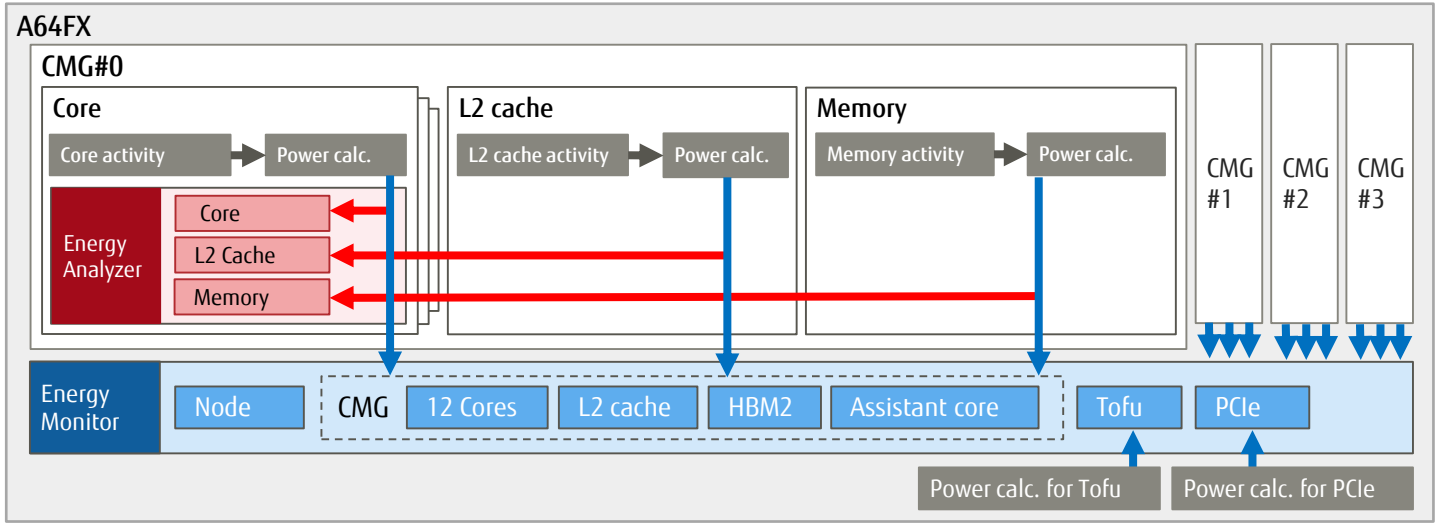
A64FX: Power monitor and analyzer

- Energy monitor (per chip)
  - Node power via Power API*1 (~msec)
  - Averaged power of a node, CMG (cores, L2 cache, memory) etc.

*1: Sandia National Laboratory

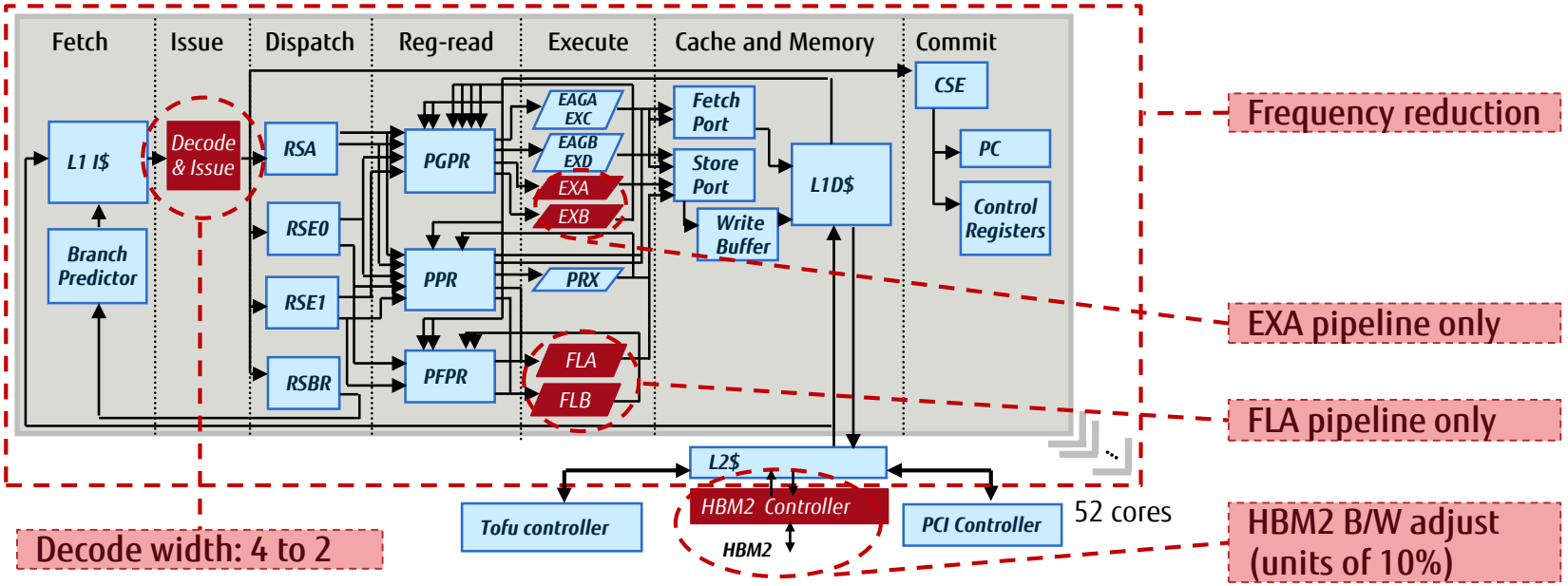
- Energy analyzer (per core)
  - Power profiler via PAPI*2 (~nsec)
  - Fine grained power analysis of a core, L2 cache, and memory

*2: Performance Application Programming Interface



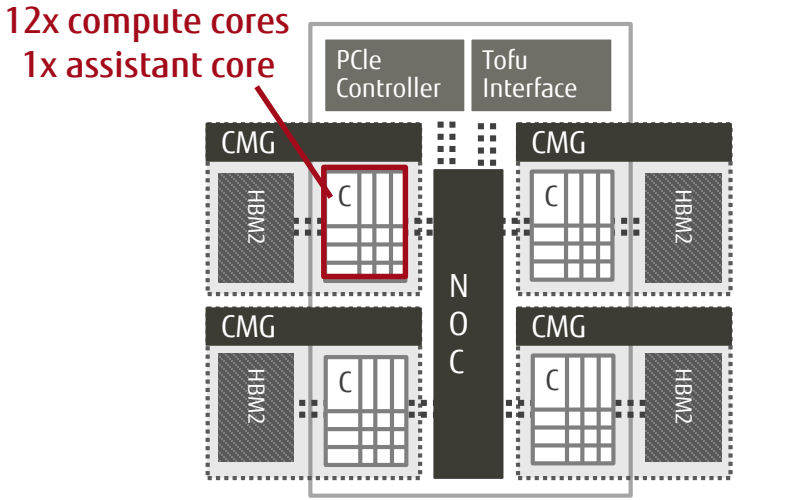
A64FX: Power Knobs to reduce power consumption

- "Power knob" limits units' activity via user APIs
- Performance/W can be optimized by utilizing Power knobs, Energy monitor & analyzer



A64FX: Summary

- Arm SVE, high performance and high efficiency
  - DP performance 2.7+ TFLOPS, >90%@DGEMM
  - Memory BW 1024 GB/s, >80%@STREAM Triad



CMG : Core Memory Group NOC : Network on Chip

	A64FX
ISA (Base, extension)	Armv8.2-A, SVE
Process technology	7 nm
Peak DP performance	2.7+ TFLOPS
SIMD width	512-bit
# of cores	48 + 4
Memory capacity	32 GiB (HBM2 x4)
Memory peak bandwidth	1024 GB/s
PCIe	Gen3 16 lanes
High speed interconnect	TofuD integrated

Agenda

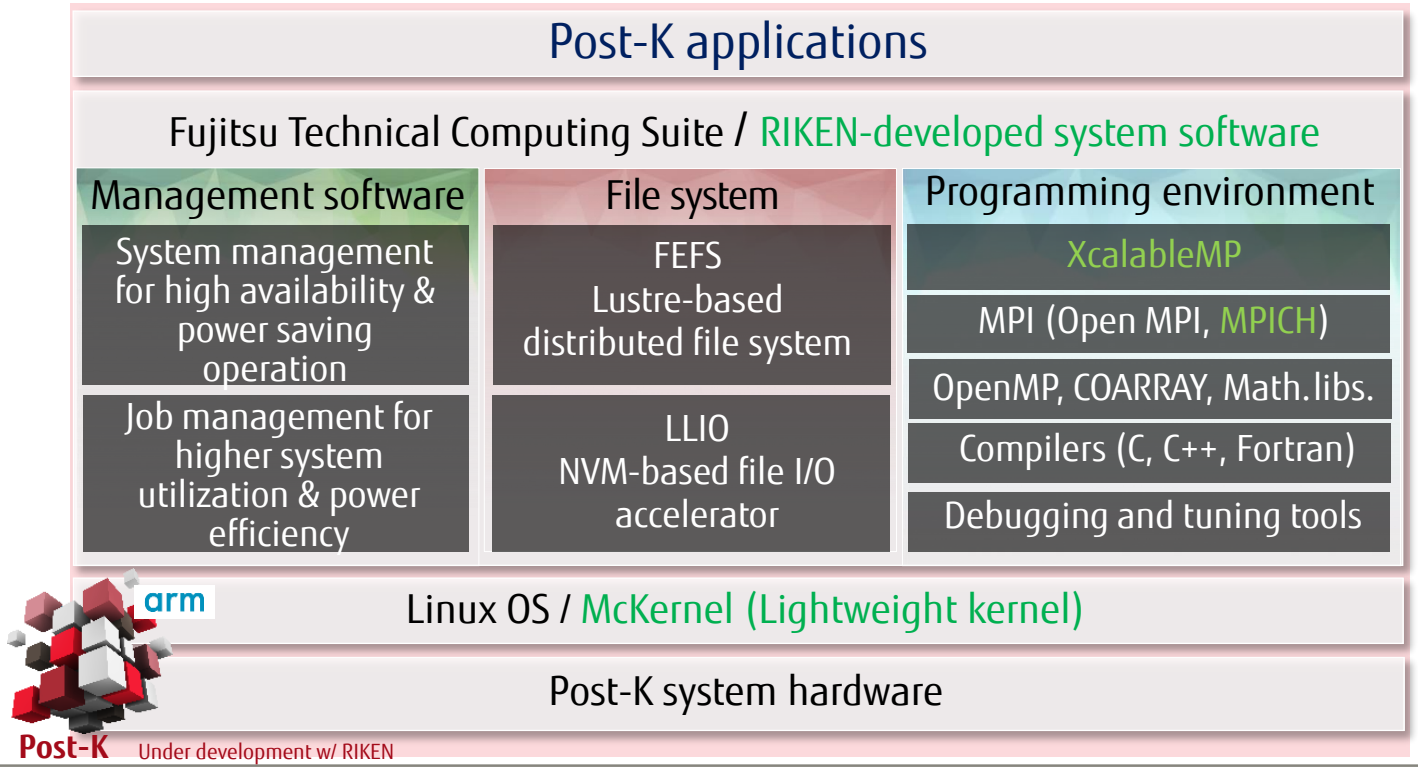
FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

Post-K system software

FUJITSU

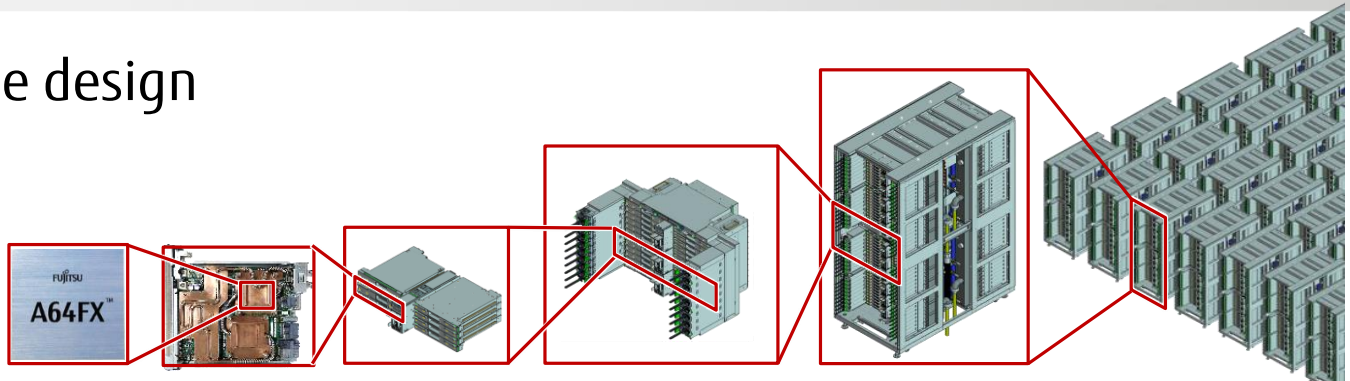
- RIKEN and Fujitsu are developing a software stack for Post-K



Post-K system configuration

FUJITSU

- Scalable design

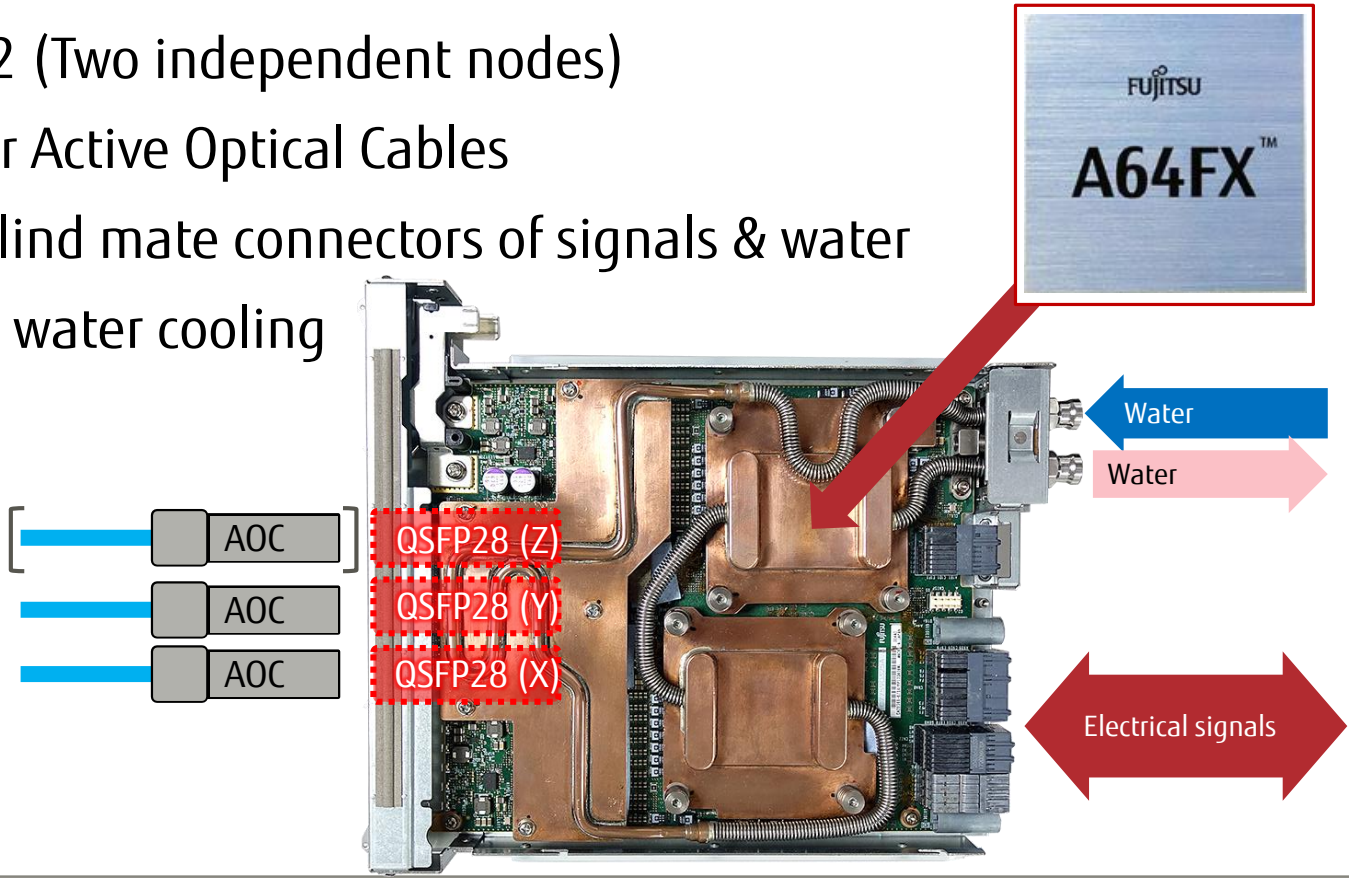


CPU CMU BoB Shelf Rack System

Unit	# of nodes	Description
CPU	1	Single socket node with HBM2 & Tofu interconnect D
CMU	2	CPU Memory Unit: 2x CPU
BoB	16	Bunch of Blades: 8x CMU
Shelf	48	3x BoB
Rack	384	8x Shelf
System	150k+	As a Post-K system

CMU: CPU Memory Unit

- A64FX CPU x2 (Two independent nodes)
- QSFP28 x3 for Active Optical Cables
- Single-side blind mate connectors of signals & water
- ~100% direct water cooling



Agenda

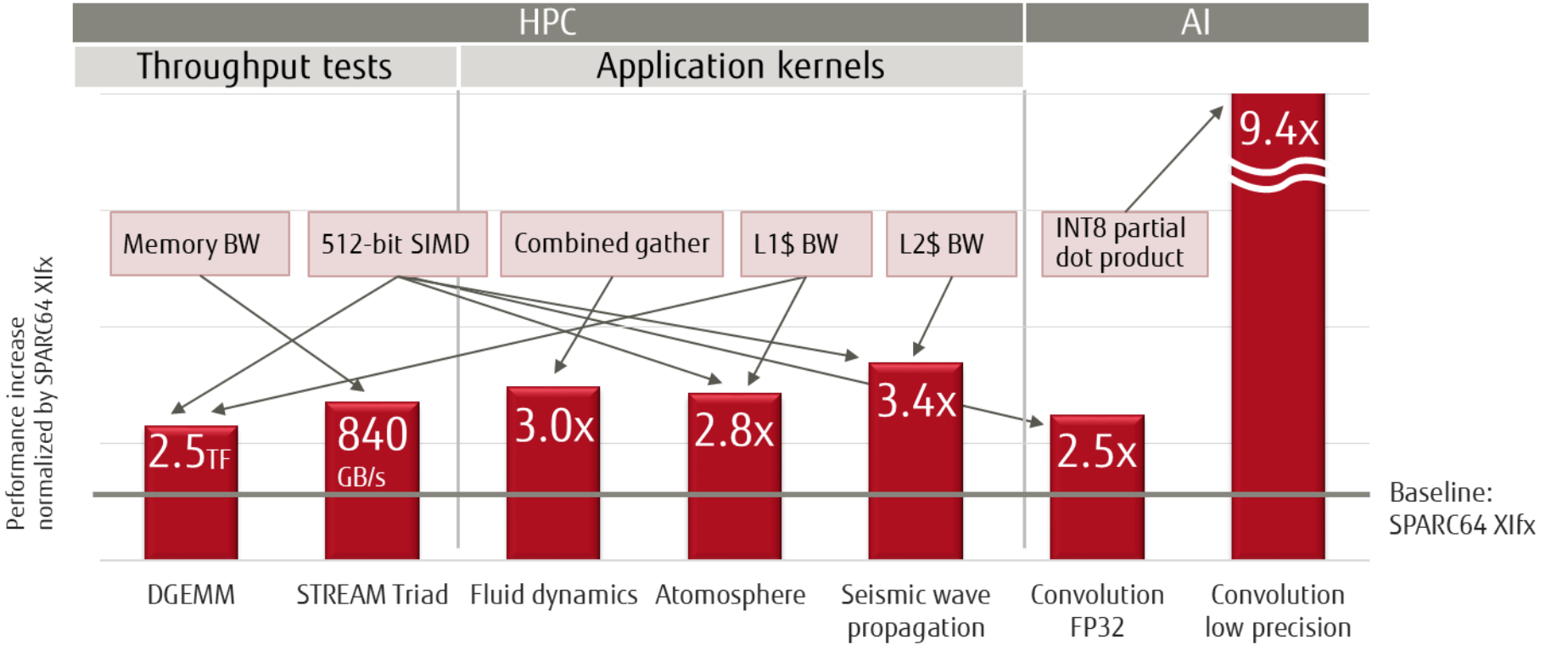
FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

Preliminary performance evaluation results

FUJITSU

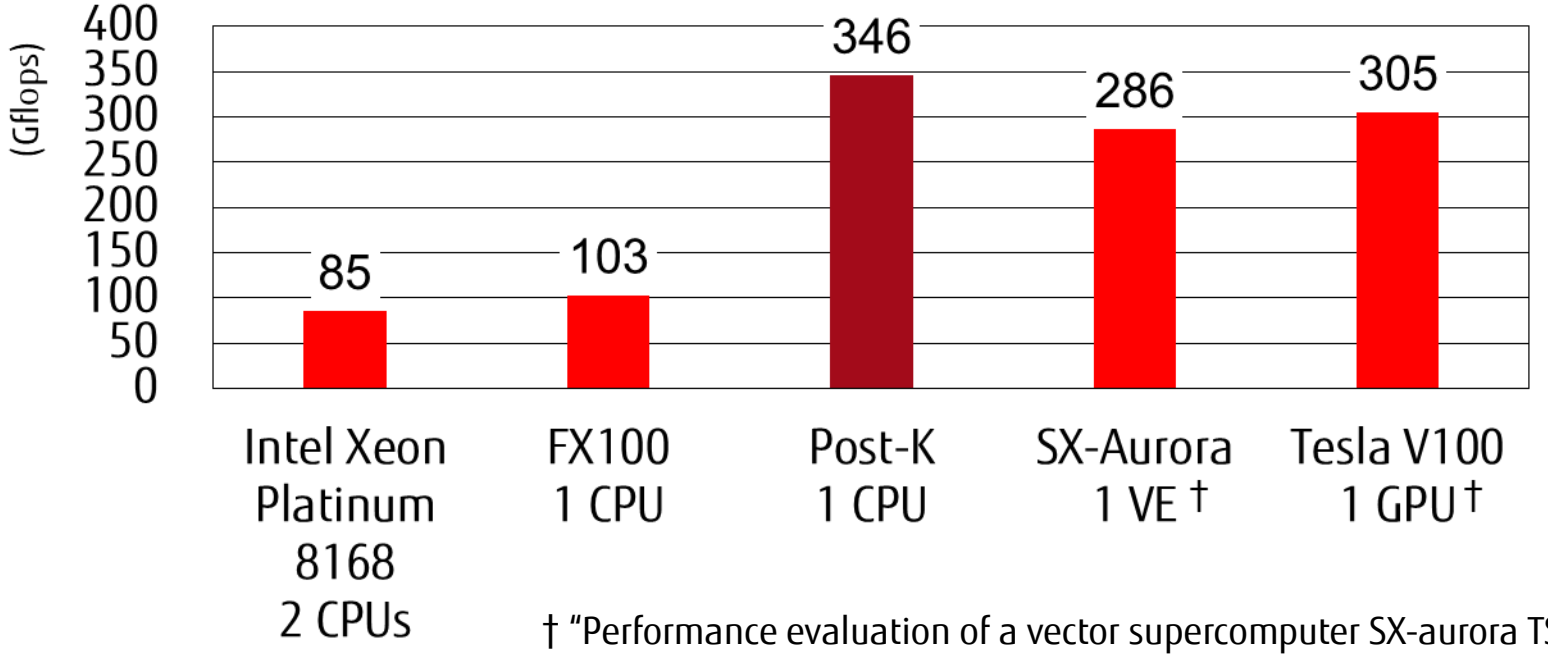
- Over 2.5x faster in HPC & AI benchmarks than SPARC64 XIfx



Post-K performance evaluation

FUJITSU

- Himeno Benchmark (Fortran90)



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

Agenda

FUJITSU

- Background
  - K computer and Post-K
  - Post-K goals and approaches
- Fujitsu's new Arm CPU A64FX
- Post-K system
  - System software
  - Configuration
- Performance discussion
- Summary and development status

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	ongoing
ABINIT	Fortran	Modified	Modified	Modified	Modified
CP2K	Fortran	Ok in as is	Issues found	Issues found	ongoing
NEST*	C++	Ok in as is	Modified	Modified	Modified
BLAST*	C++	Ok in as is	Modified	Modified	Modified

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	ongoing
ABINIT	Fortran	Modified	Modified	Modified	Modified
CP2K	Fortran	Ok in as is	Issues found	Issues found	ongoing
NEST*	C++	Ok in as is	Modified	Modified	Modified
BLAST*	C++	Ok in as is	Modified	Modified	Modified

Twelve primary OSS applications are listed and being tested in the Users Group for each compilers, collaboratively w/ Arm

Summary of Post-K

FUJITSU

- High application performance with superior power efficiency
  - High memory bandwidth & wider SIMD
- Arm SVE support
  - Scalable vector extension, state-of-the-art Arm instruction set architecture
  - Binary compatibility advances and expands the power of ecosystem by incorporating HPC technologies and applications
- Optimizing for new supercomputer standards
  - Low power consumption in many aspects
  - Being focusing on the existing applications and emerging applications

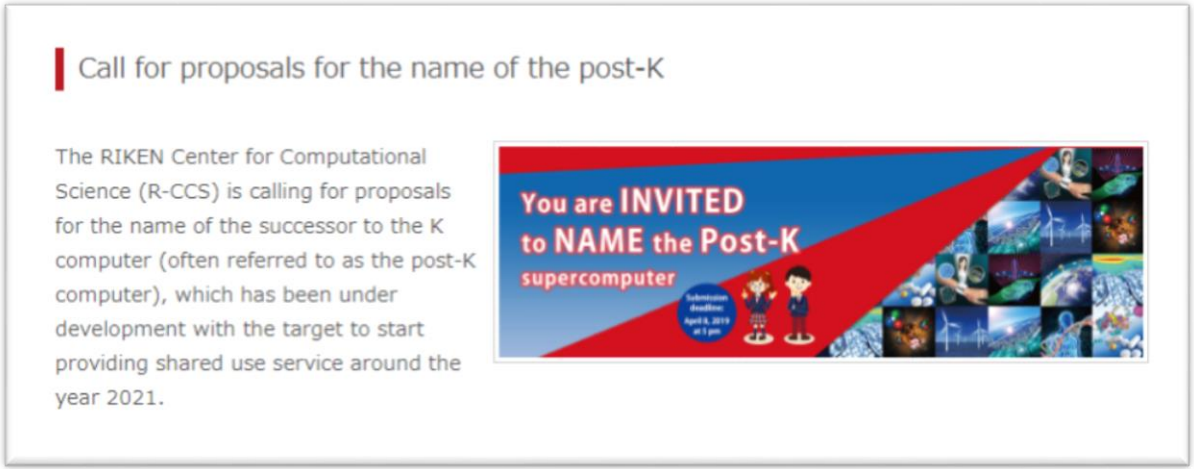
Post-K is being built based on the strong and proven microarchitecture and system software. Arm and its ecosystem boost Post-K's value even further

Submit today!

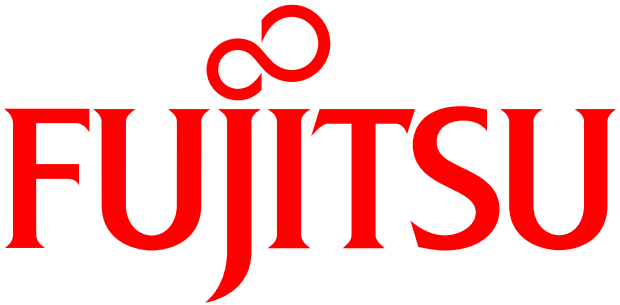
until April 8, 2019 at 5 pm

FUJITSU

- Requirements for the post-K's name are:
  - The name should preferably express the idea that RIKEN is a world-class research institute operating a state-of-the–art supercomputer.
  - The name should be attractive not only to Japanese speakers but to people around the world.



https://www.r-ccs.riken.jp/en/topics/naming.html



shaping tomorrow with you