FUJITSU

shaping tomorrow with you

Supercomputer Fugaku

Toshiyuki Shimizu

FUJITSU LIMITED

Feb. 18th, 2020

Outline

FUJITSU

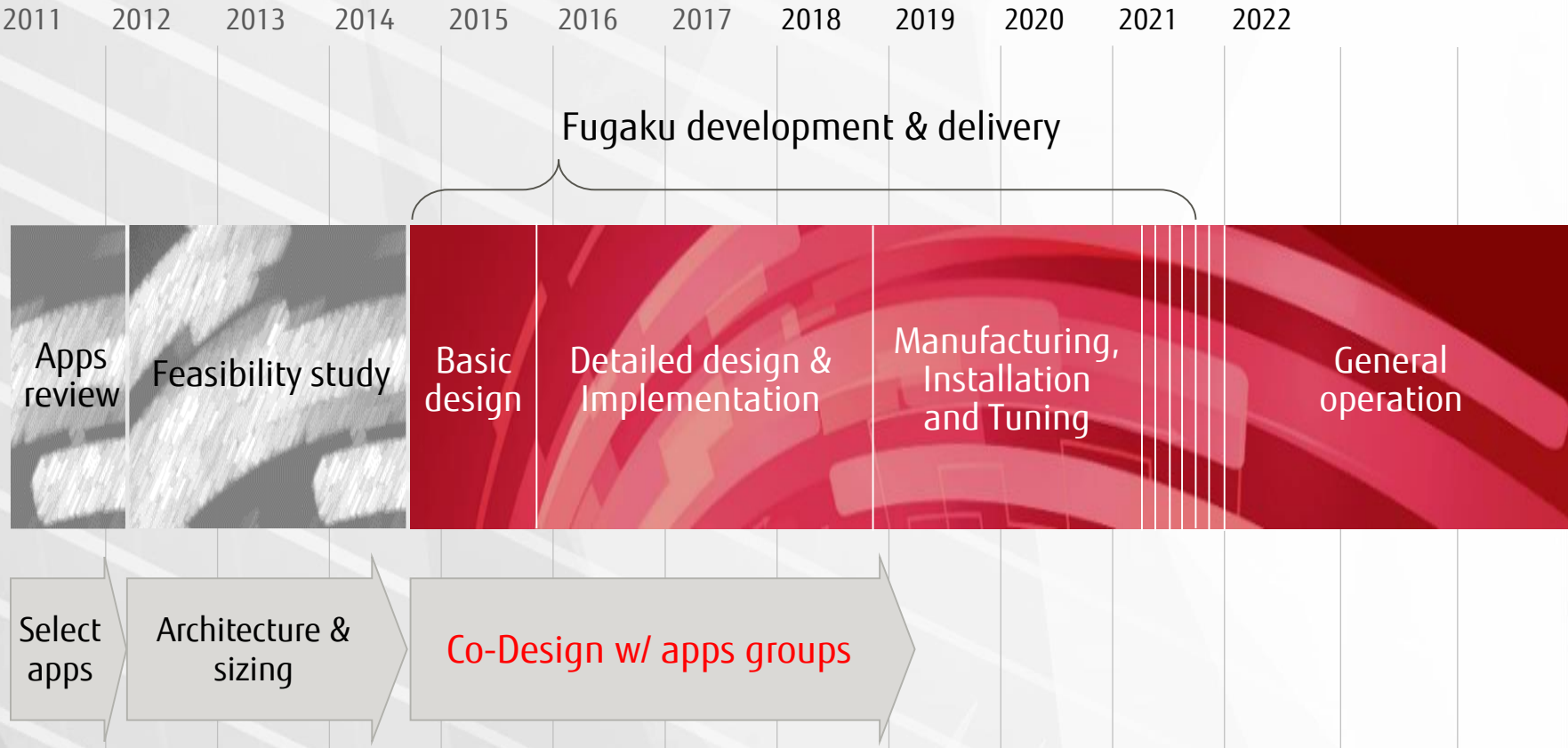
- Fugaku project overview
- Co-design
  - Approach
  - Design results
- Performance & energy consumption evaluation
  - Green500
  - OSS apps
  - Fugaku priority issues
- Summary

Supercomputer “Fugaku”, formerly known as Post-K

FUJITSU

Focus	Approach
 Application performance	Co-design w/ application developers and Fujitsu-designed CPU core w/ high memory bandwidth utilizing HBM2
 Power efficiency	Leading-edge Si-technology, Fujitsu's proven low power & high performance logic design, and power-controlling knobs
 Usability	Arm®v8-A ISA with Scalable Vector Extension (“SVE”), and Arm standard Linux

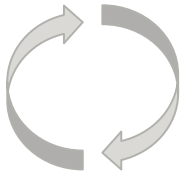
Fugaku project schedule



Fugaku co-design

FUJITSU

- Co-design goals
  - Obtain the best performance, 100x apps performance than K computer, within power budget, 30-40MW
    - Design applications, compilers, libraries, and hardware
- Approach
  - Estimate perf & power using apps info, performance counts of Fujitsu FX100, and cycle base simulator
    - Computation time: brief & precise estimation
    - Communication time: bandwidth and latency for communication w/ some attributes for communication patterns
    - I/O time:
  - Then, optimize apps/compilers etc. and resolve bottlenecks

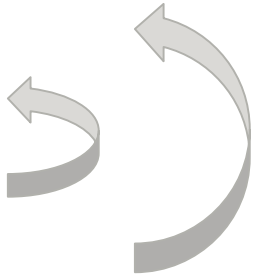


- Estimation of performance and power
  - Precise performance estimation for primary kernels
    - Make & run Fugaku objects on the Fugaku cycle base simulator
  - Brief performance estimation for other sections
    - Replace performance counts of FX100 w/ Fugaku params: # of inst. commit/cycle, wait cycles of barrier, inst. fetch, branch, fp exec, data load/store wait cycles of L1D/L2, etc.
  - Power estimation
    - DGEMM execution toggles on the emulator + estimation of memory and interconnect considering utilization + loss of convertors

Co-design iterations around year 2015

FUJITSU

- Apply each of application kernels
  - Define/refine a set of architecture parameters
  - Implement/tune the kernel under the architecture parameters
  - Evaluate execution time using the estimation tools
  - Identify hardware bottlenecks and explore design space



- Examples of architecture parameters
  - Frequency, # of arch regs, SIMD width, cache structure & size, # of cores...
  - Memory, interconnect parameters
  - Implementation of instructions: i.e. Combined gather...

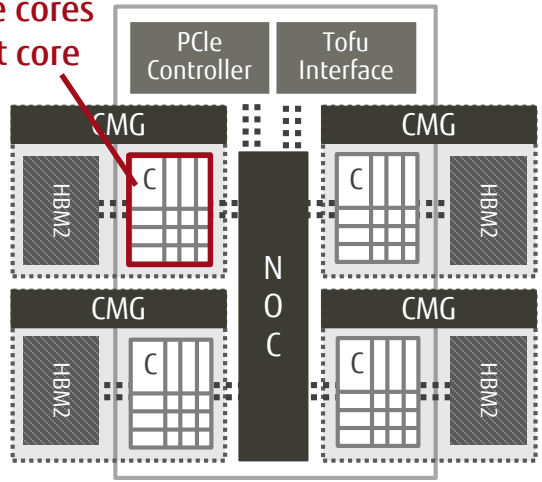
A64FX CPU

FUJITSU

- Arm SVE, high performance and high efficiency
  - DP performance 2.7+ TFLOPS, >90%@DGEMM, (CPU freq=1.8/2.0/2.2 GHz)
  - Memory BW 1024 GB/s, >80%@STREAM Triad

FUGAKU operating cond.

12x compute cores
1x assistant core



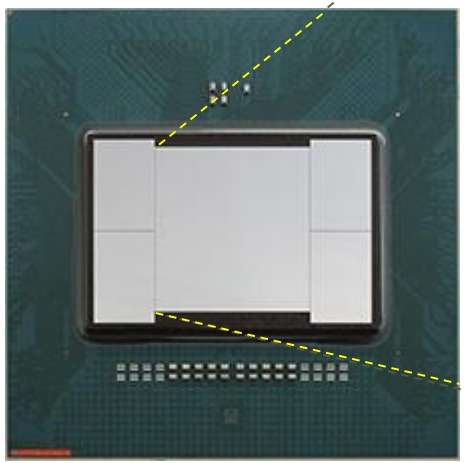
CMG : Core Memory Group NOC : Network on Chip

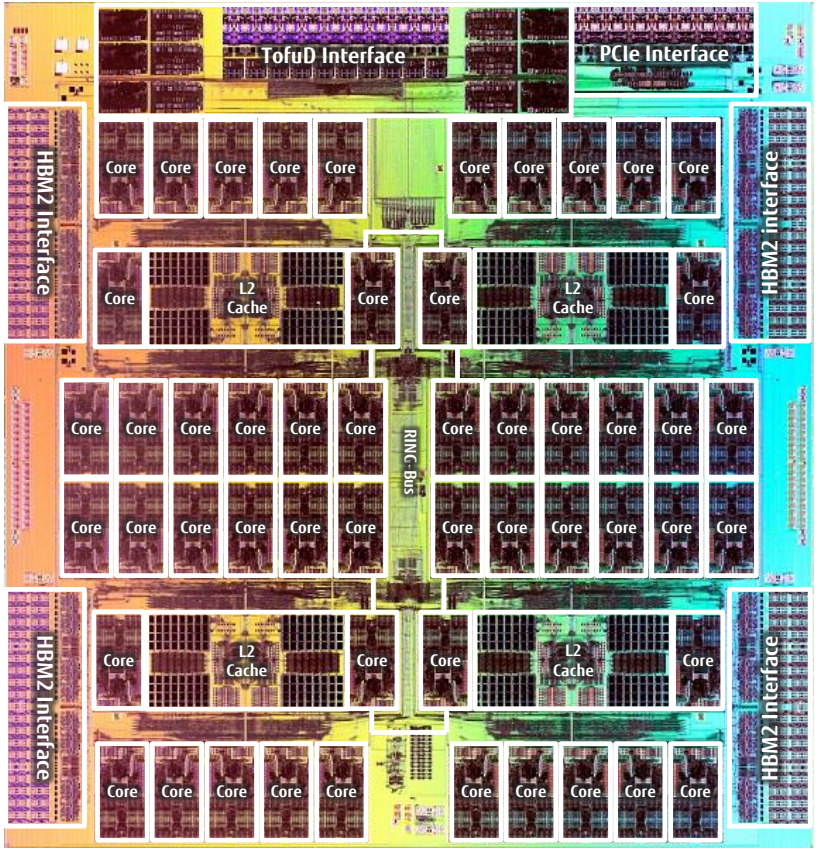
	A64FX
ISA (Base, extension)	Armv8.2-A, SVE
Peak DP performance	2.7+ TFLOPS
SIMD width	512-bit
# of cores	48 + 4
Memory capacity	32 GiB (HBM2 x4)
Memory peak bandwidth	1024 GB/s
PCIe	Gen3 16 lanes
High speed interconnect	TofuD integrated

A64FX leading-edge Si-technology

FUJITSU

- TSMC 7nm FinFET & CoWoS
  - Broadcom SerDes, HBM I/O, and SRAMs
  - 8.786 billion transistors
  - 594 signal pins

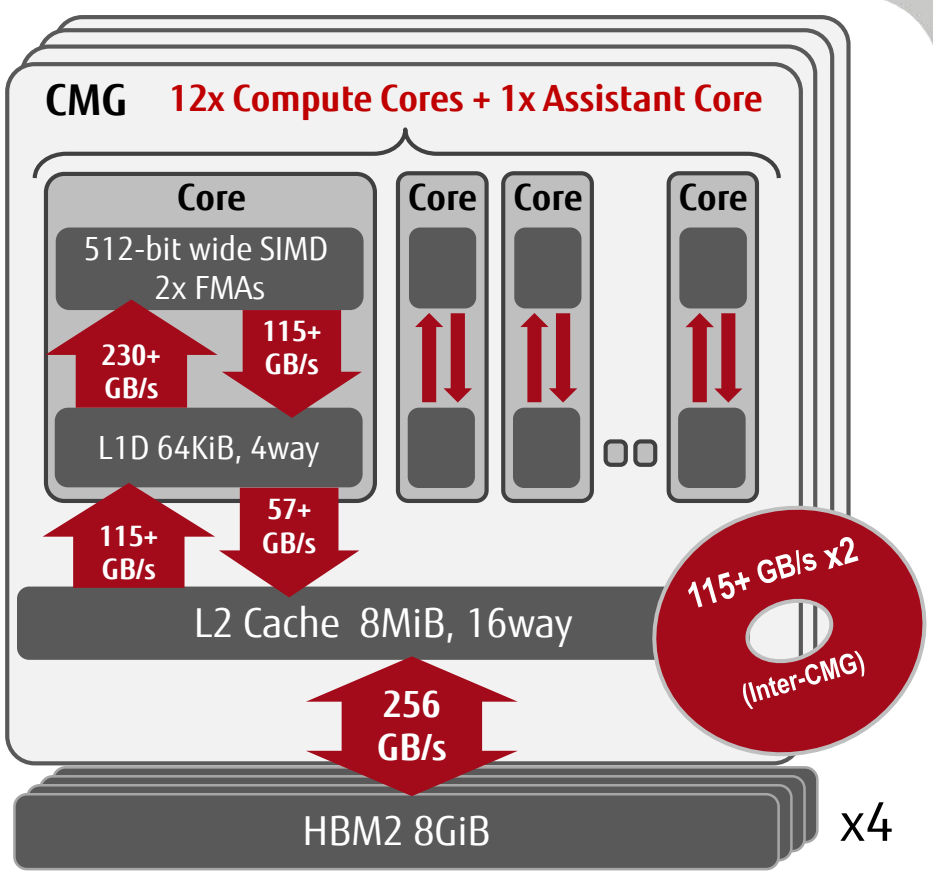




Fujitsu-designed CPU core w/ High Memory Bandwidth

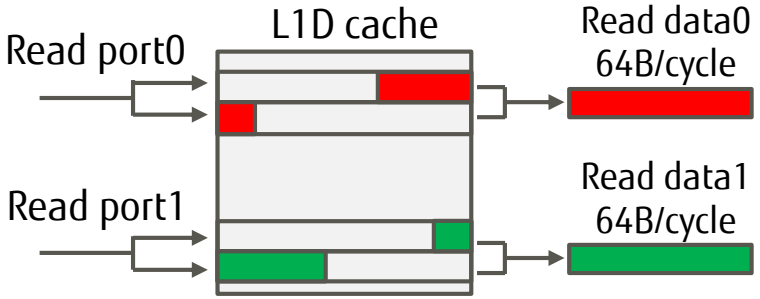
- A64FX out-of-order controls in cores, caches, and memories achieve superior throughput

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



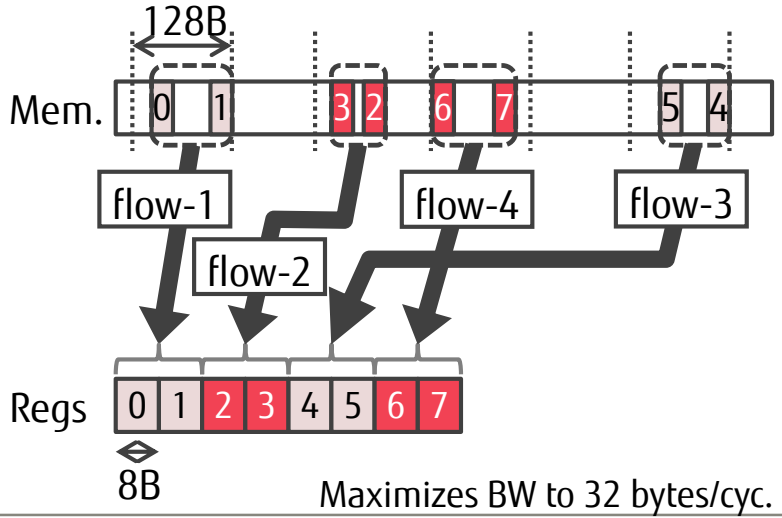
A64FX optimized load efficiency for apps performance

- 128 bytes/cycle sustained bandwidth 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

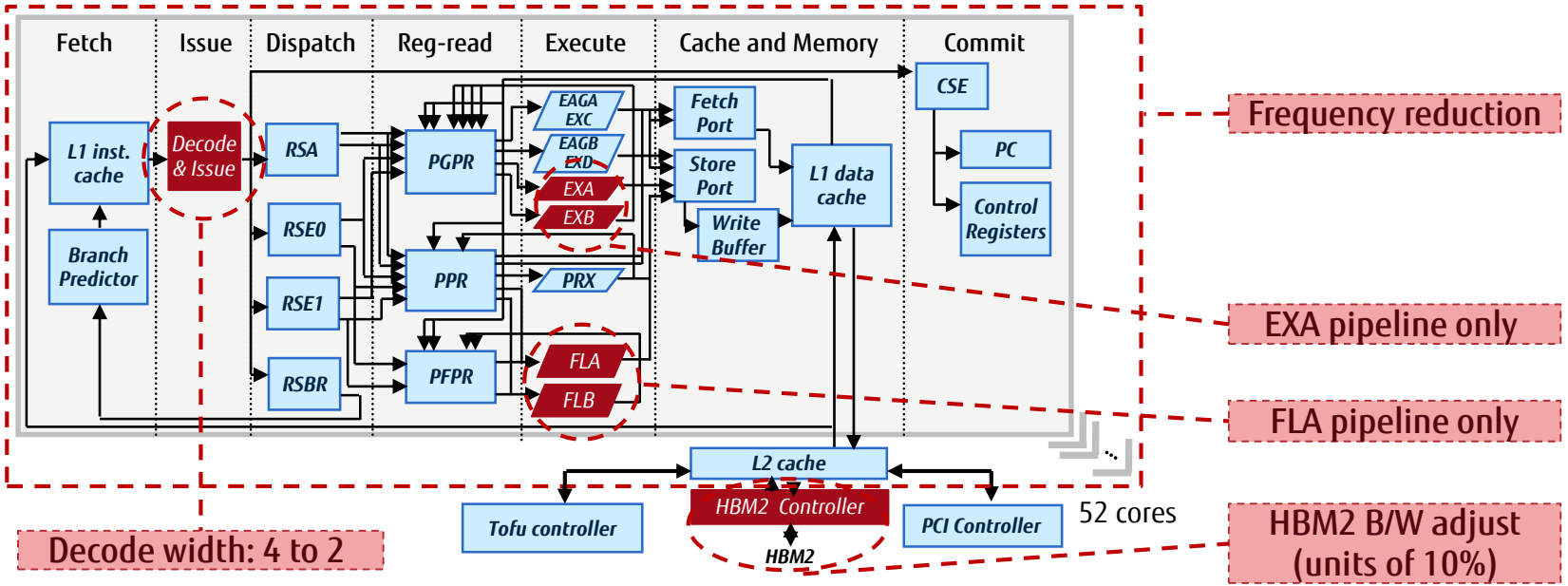
  Suggested through Co-design work w/ app teams



Maximizes BW to 32 bytes/cyc.

A64FX Power Knobs to reduce power consumption

- "Power knob" limits units' activities via user APIs
- Performance/Watt can be optimized by utilizing Power knobs



Fugaku system software developed with RIKEN

FUJITSU

Fugaku applications

Fujitsu Technical Computing Suite / RIKEN-developed system software

Management software	File system	Programming environment
System management for high availability & power saving operation	FEFS Lustre-based distributed file system	XcalableMP, FDPS
Job management for higher system utilization & power efficiency	LLIO NVM-based file I/O Accelerator for Fugaku	Open MPI, MPICH(PiP), DTF
		OpenMP, COARRAY, Math.libs.
		Compilers, AI frameworks
		Debug & tuning tools, Containers

Tensorflow
Chainer
Pytorch

Docker
KVM
Singularity

Multi-Kernel System: RHEL8 and light-weight kernel (IHK/McKernel)

Fugaku system hardware

Fugaku developed w/ RIKEN

FDPS: Framework for Developing Particle Simulator
PiP: Process-in-Process
DTF: Data Transfer Framework
IHK: Interface for Heterogeneous Kernels

Outline

FUJITSU

- Fugaku project overview
- Co-design
  - Approach
  - Design results
- Performance & energy consumption evaluation
  - Green500
  - OSS apps
  - Fugaku priority issues
- Summary

Green500, Nov. 2019

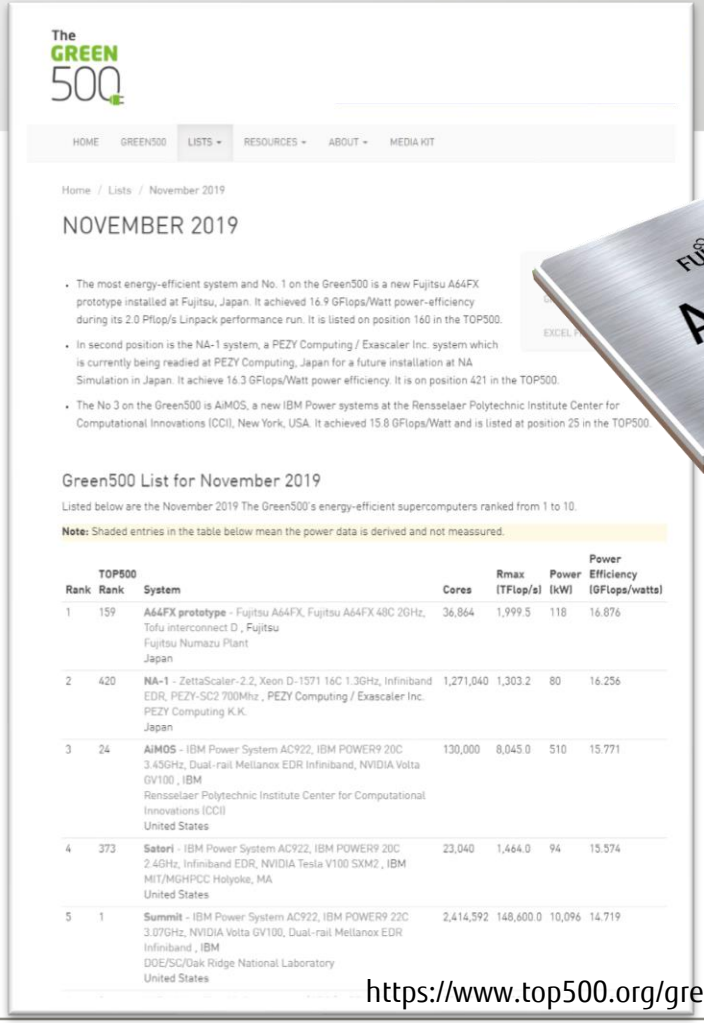
A64FX prototype –

Fujitsu A64FX 48C 2GHz

ranked #1 on the list

768x general purpose A64FX CPU w/o accelerators

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



The GREEN 500

HOME | GREEN500 | LISTS | RESOURCES | ABOUT | MEDIA KIT

Home / Lists / November 2019

NOVEMBER 2019

- The most energy-efficient system and No. 1 on the Green500 is a new Fujitsu A64FX prototype installed at Fujitsu, Japan. It achieved 16.9 GFlops/Watt power-efficiency during its 2.0 Pflop/s Linpack performance run. It is listed on position 160 in the TOP500.
- In second position is the NA-1 system, a PEZY Computing / Exascaler Inc. system which is currently being readied at PEZY Computing, Japan for a future installation at NA Simulation in Japan. It achieve 16.3 GFlops/Watt power efficiency. It is on position 421 in the TOP500.
- The No 3 on the Green500 is AiMOS, a new IBM Power systems at the Rensselaer Polytechnic Institute Center for Computational Innovations (CCI), New York, USA. It achieved 15.8 GFlops/Watt and is listed at position 25 in the TOP500.

Green500 List for November 2019

Listed below are the November 2019 The Green500's energy-efficient supercomputers ranked from 1 to 10.

Note: Shaded entries in the table below mean the power data is derived and not meassured.

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

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

How we achieved

FUJITSU

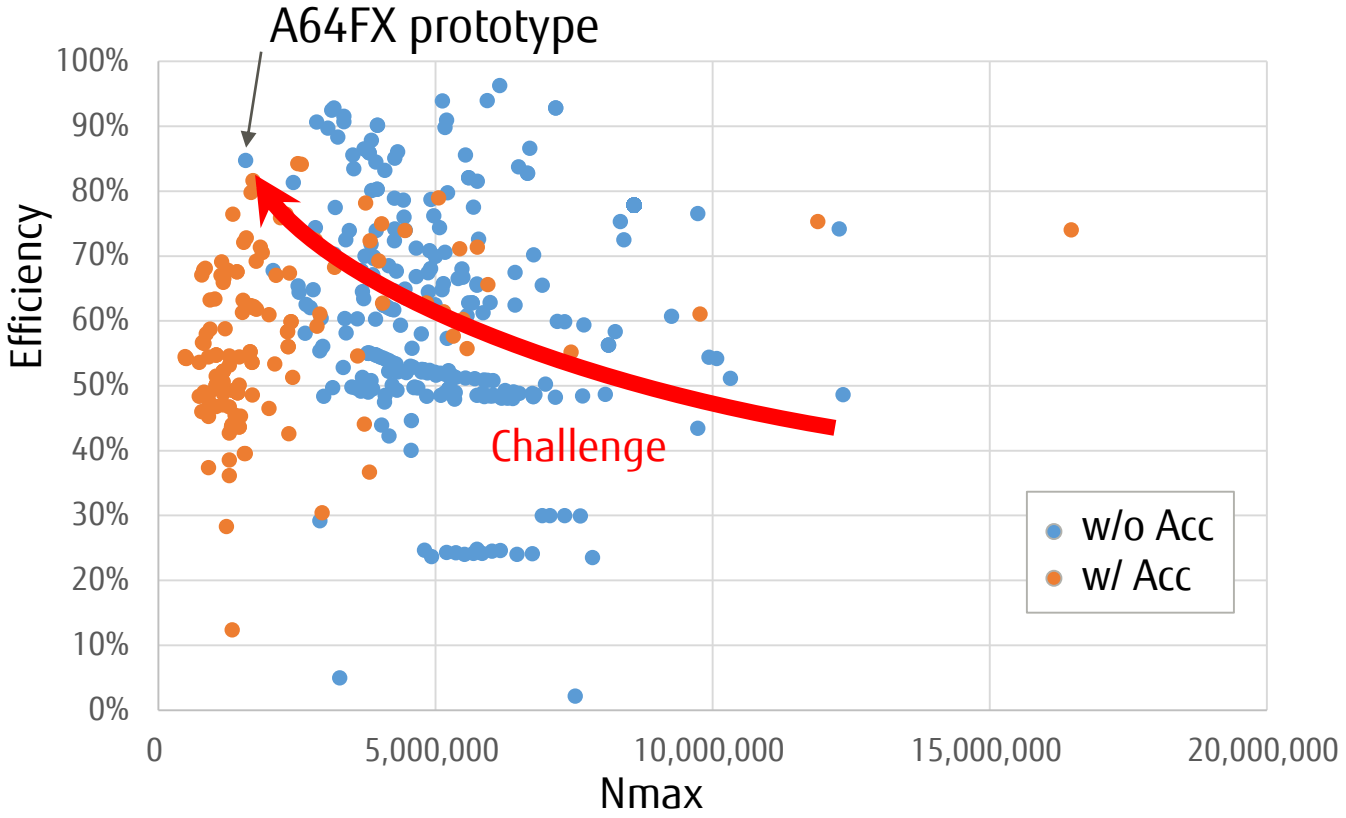
- Key for GF/W is {energy efficient HW} x {parallel/exec efficiency}
- A64FX is designed for energy efficient
  - Fujitsu's proven CPU microarchitecture & 7nm FinFET
  - SoC design: Tofu interconnect D integrated
  - CoWoS: 4x HBM2 for main memory integrated
- Superior parallel/exec efficiency
  - Math. libraries are tuned for application efficiency
  - Comm. libs are also tuned utilizing long experience of Tofu @ K computer
  - Performance tuning is efficiently done utilizing rich performance analyzer/monitor

+ Concerted efforts of co-design

SC19 TOP500 calculation efficiency

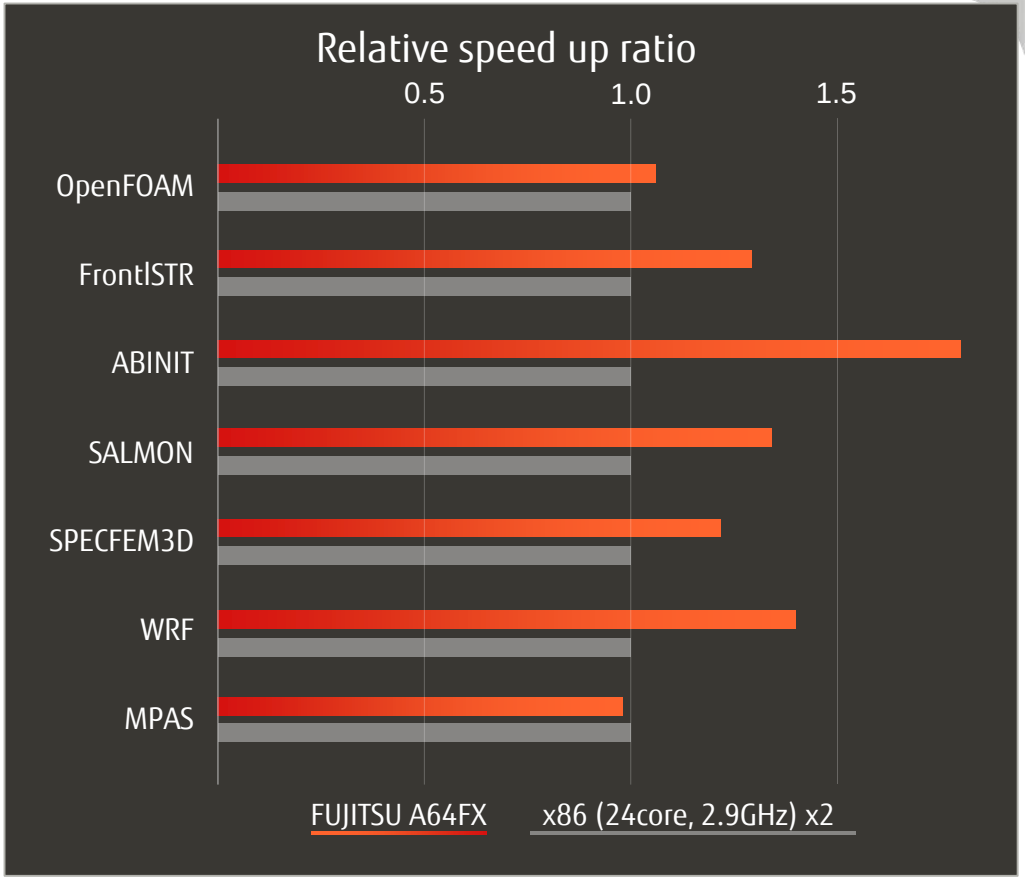
FUJITSU

- A64FX superior efficiency 84.75% with small Nmax
- Results of:
  - Optimized communication and math. libs
  - Optimization of overlapped communication



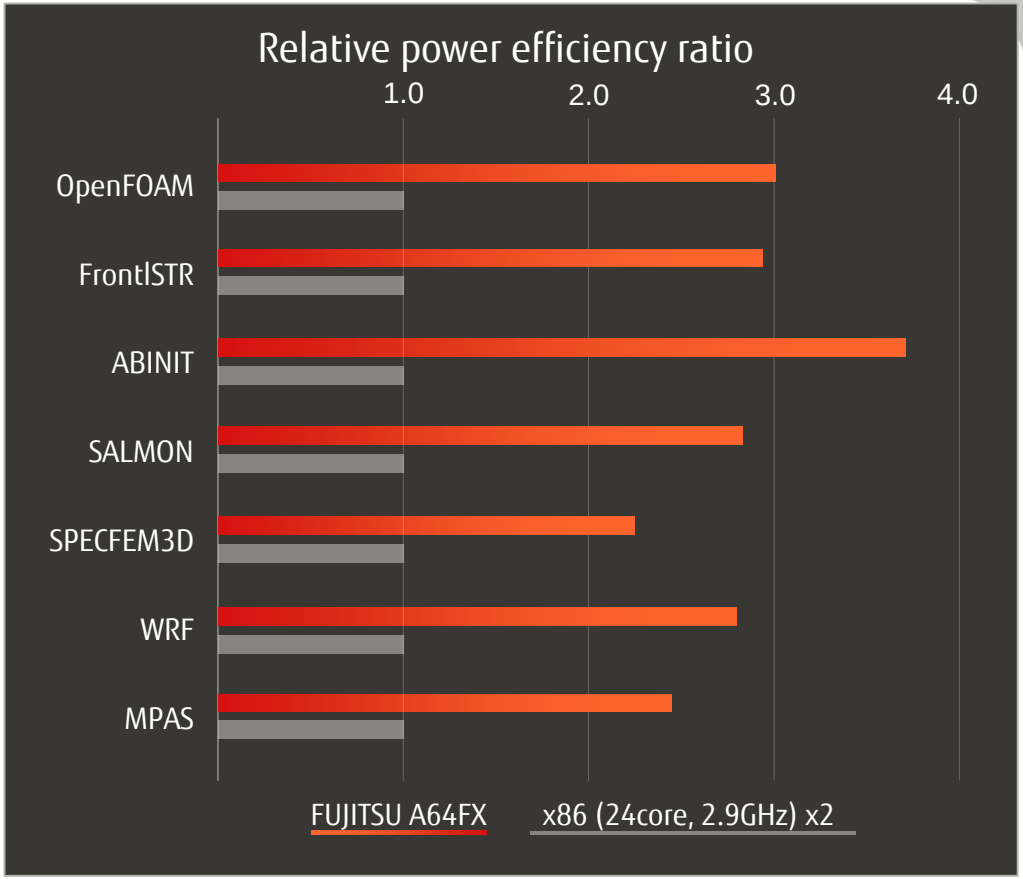
A64FX CPU performance evaluation for real apps

- Open source software, Real apps on an A64FX @ 2.2GHz
- Up to 1.8x faster over the latest x86 processor (24core, 2.9GHz) x2
- High memory BW and long SIMD length of A64FX work effectively with these applications



A64FX CPU power efficiency for real apps

- Performance / Energy consumption on an A64FX @ 2.2GHz
- Up to 3.7x more efficient over the latest x86 processor (24core, 2.9GHz) x2
- High efficiency is achieved by energy-conscious design and implementation



Fugaku priority issues and performance prediction

FUJITSU

- 100x app performance and power budget requirements are met
- Some apps utilize power knob to reduce power consumption and achieve high energy efficiency
- Measuring and optimizing using real HW

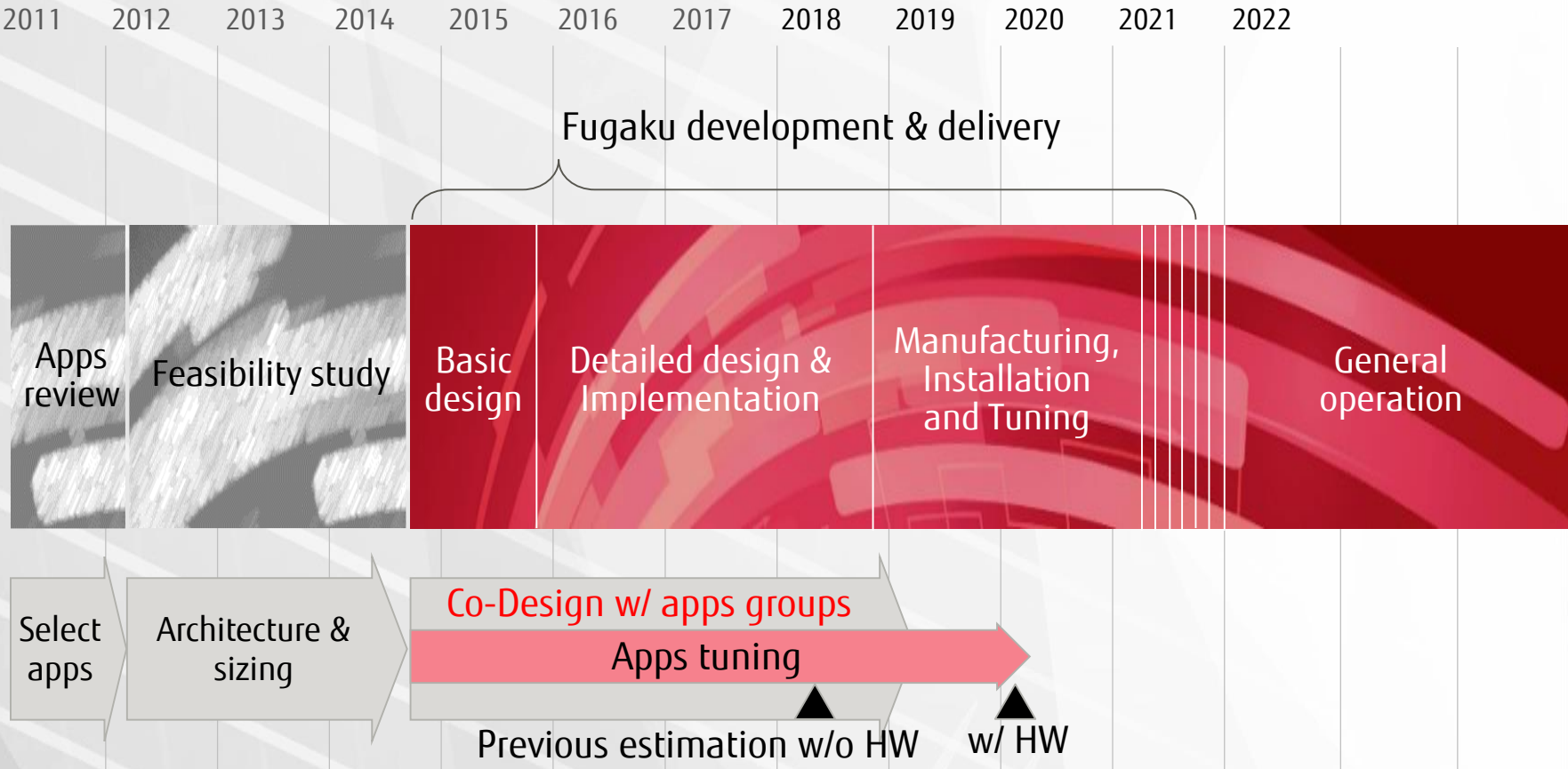
https://postk-web.r-ccs.riken.jp/perf.html

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	125x +	GENESIS	MD for proteins
	2. Personalized and preventive medicine using big data	8x +	Genomon	Genome processing (Genome alignment)
Disaster prevention and Environment	3. Integrated simulation systems induced by earthquake and tsunami	45x +	GAMERA	Earthquake simulator (FEM in unstructured & structured grid)
	4. Meteorological and global environmental prediction using big data	120x +	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	40x +	NTChem	Molecular electronic simulation (structure calculation)
	6. Accelerated development of innovative clean energy systems	35x +	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	30x +	RSDFT	Ab-initio simulation (density functional theory)
	8. Development of innovative design and production processes	25x +	FFB	Large Eddy Simulation (unstructured grid)
Basic science	9. Elucidation of the fundamental laws and evolution of the universe	25x +	LQCD	Lattice QCD simulation (structured grid Monte Carlo)

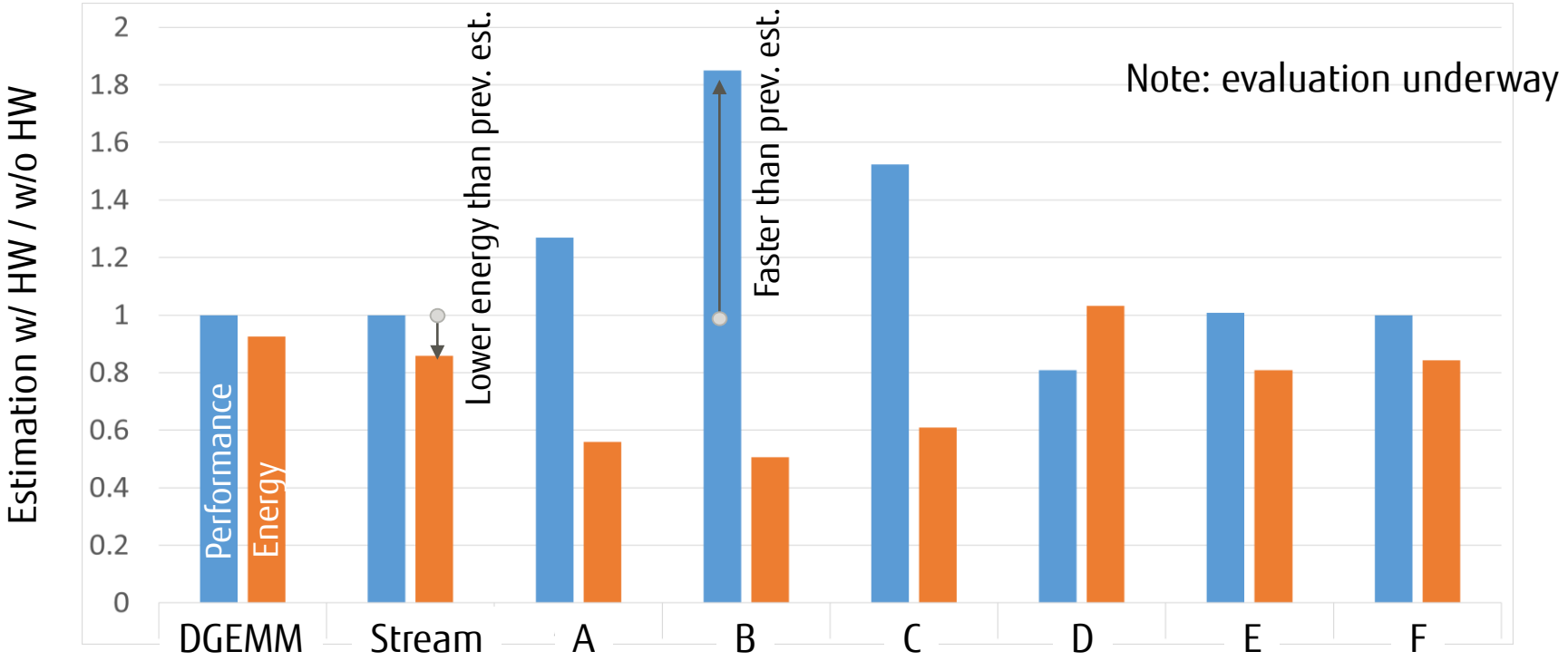
Fugaku project schedule



Current status normalized by the estimation w/o HW

FUJITSU

- Performance and energy (=power*elapse) of apps



Current status normalized by the estimation w/o HW

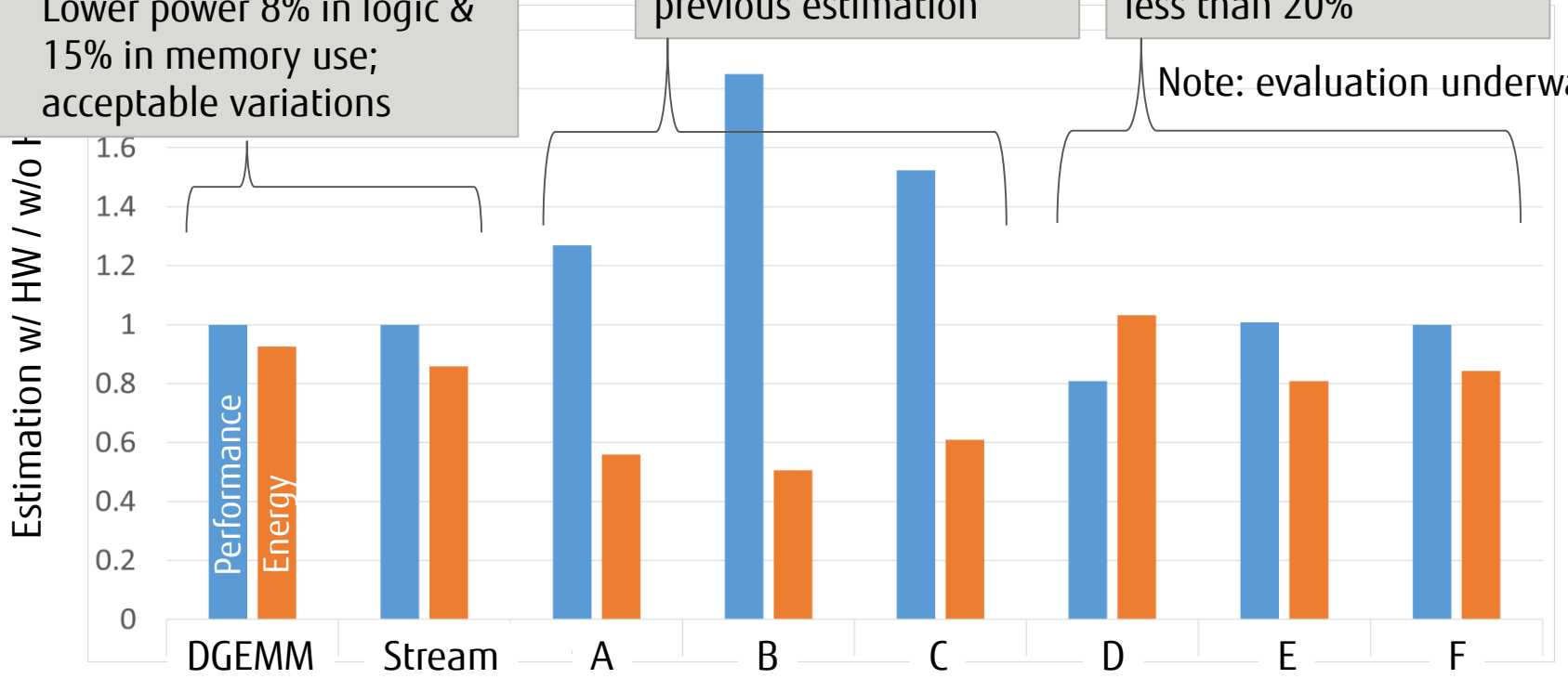
FUJITSU

- Performances are match due to precise simulation
- Lower power 8% in logic & 15% in memory use; acceptable variations

Computations were reduced after the previous estimation

Good estimation and optimization: errors are less than 20%

Note: evaluation underway



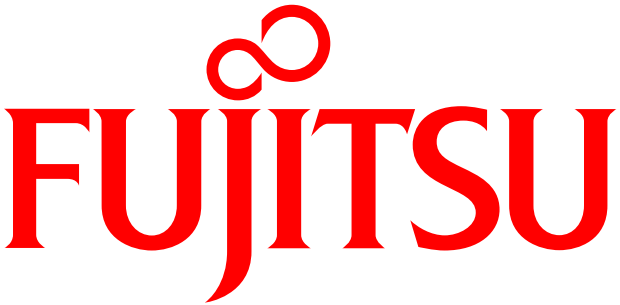
Summary

- "Fugaku" is co-designed and runs apps at high performance w/ optimal power consumption
- Arm HPC ecosystem and apps portfolio are growing by efforts of communities and selections of HW

https://www.riken.jp/pr/news/2019/20190827_1/

- Fujitsu Supercomputer PRIMEHPC FX1000 & FX700 based on Fugaku tech.
- Cray CS500 using Fujitsu A64FX Arm CPU

Fujitsu PRIMEHPC FX1000



shaping tomorrow with you