Fugaku: the First 'Exascale' Machine

Satoshi Matsuoka, Director R-CCS / Prof. Tokyo Inst. Tech. The 6th ENES HPC Workshop (virtual) 26 May 2020

Ye,





The "Fugaku" 富岳 "Exascale" Supercomputer for Society 5.0 High-Peak Mt. Fujj representing _arge the ideal computing Scale Applicatior (Capability cceleration **Broad Base --- Applicability & Capacity** Broad Applications: Simulation, Data Science, AI, ... Broad User Bae: Academia, Industry, Cloud Startups, ... For Society 5.0

Arm64fx & Fugaku 富岳 /Post-K are:



- Fujitsu-Riken design A64fx ARM v8.2 (SVE), 48/52 core CPU
 - HPC Optimized: Extremely high package high memory BW (1TByte/s), on-die Tofu-D network BW (~400Gbps), high SVE FLOPS (~3Teraflops), various AI support (FP16, INT8, etc.)
 - Gen purpose CPU Linux, Windows (Word), other SCs/Clouds
 - Extremely power efficient > <u>10x power/perf efficiency for CFD</u>
 <u>benchmark</u> over current mainstream x86 CPU
- Largest and fastest supercomputer to be ever built circa 2020
 - > 150,000 nodes, superseding LLNL Sequoia
 - > 150 PetaByte/s memory BW
 - Tofu-D 6D Torus NW, 60 Petabps injection BW (10x global IDC traffic)
 - 25~30PB NVMe L1 storage

6

many endpoint 100Gbps I/O network into Lustre



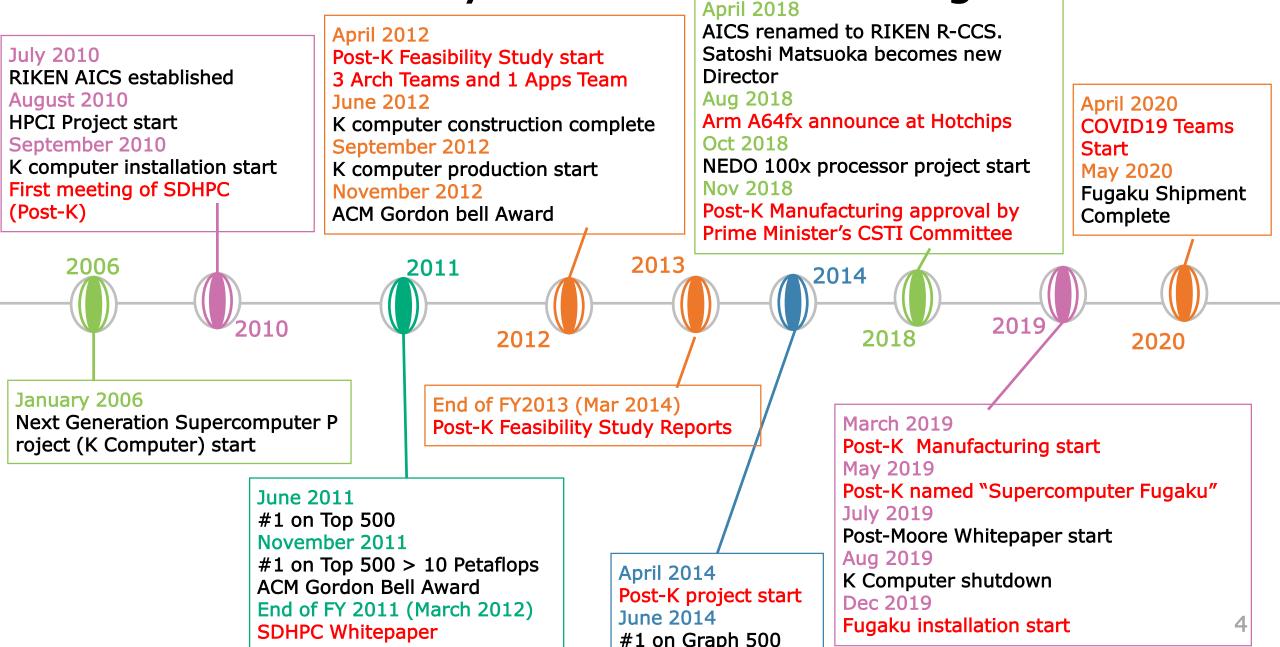
FUITSU

A64FX



Brief History of R-CCS towards Fugaku





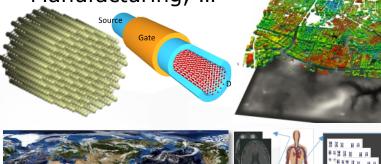
Co-Design Activities in Fugaku





Science by Computing

• 9 Priority App Areas: High Concern to General Public: Medical/Pharma, Environment/Disaster, Energy, Manufacturing, ...



6

Select representatives fr om 100s of applications signifying various compu tational characteristics Design systems with param eters that consider various application characteristics





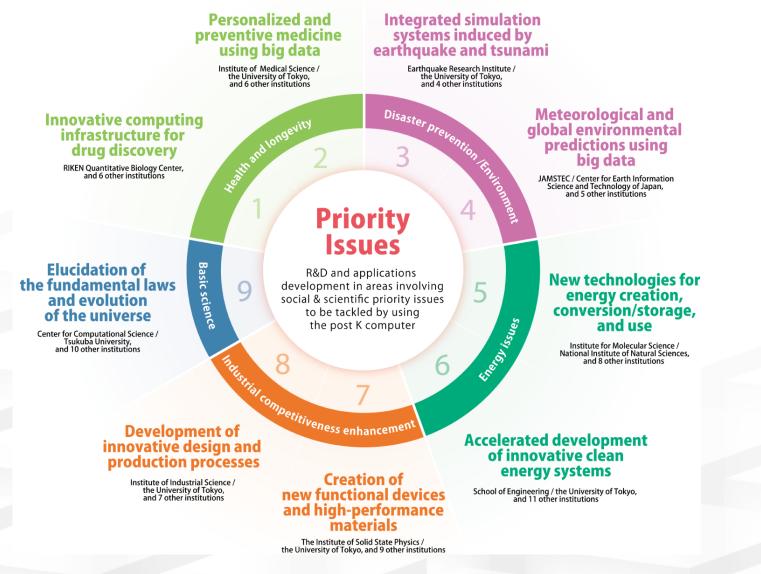


- Extremely tight collabrations between the Co-Design apps centers, Riken, and Fujitsu, etc.
- Chose 9 representative apps as "target application" scenario
- Achieve up to x100 speedup c.f. K-Computer
- Also ease-of-programming, broad SW ecosystem, very low power, ...

Research Subjects of the Post-K Computer



The post K computer will expand the fields pioneered by the K computer, and also challenge new areas.



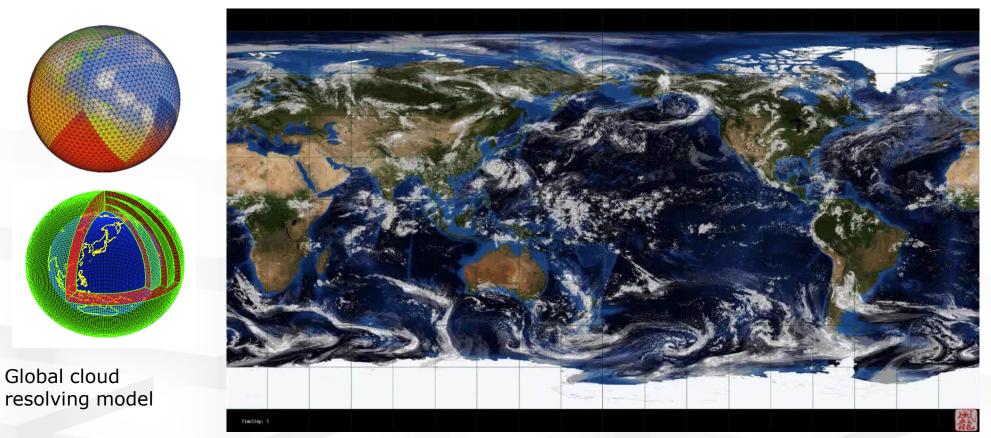


NICAM: Global Climate Simulation

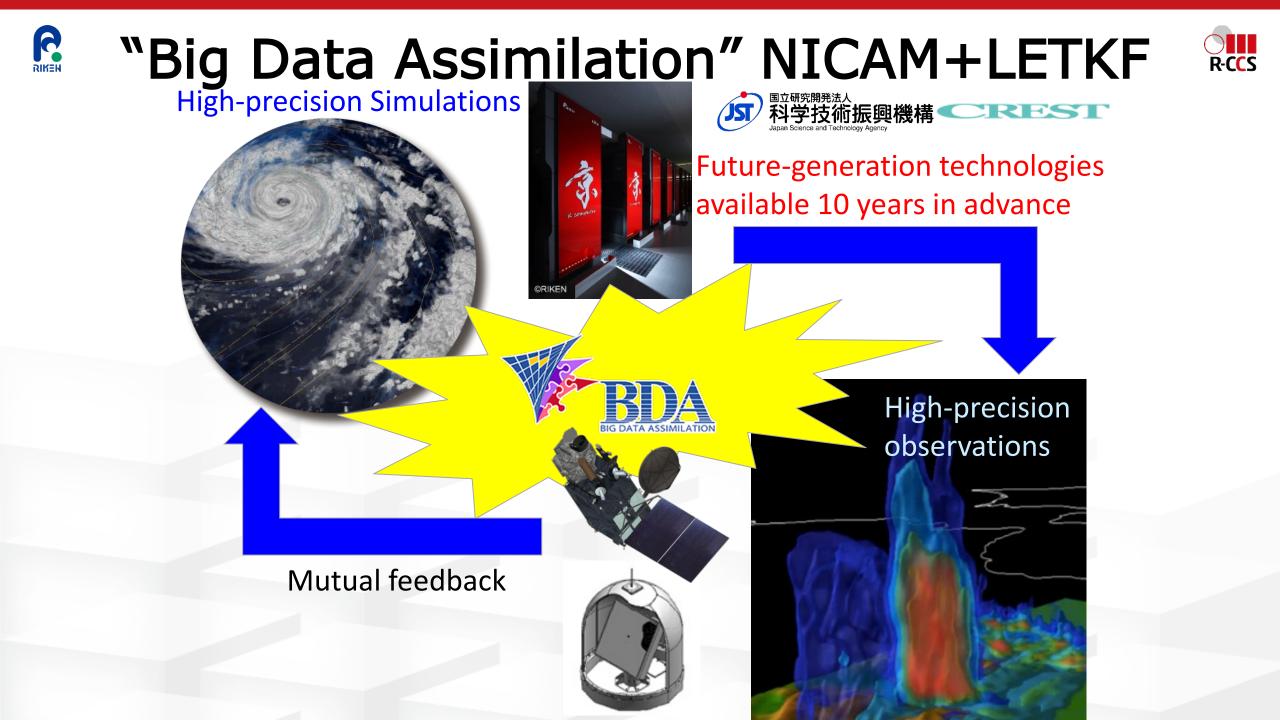


Global cloud resolving model with 0.87 km-mesh which allows resolution of cumulus clouds

Month-long forecasts of Madden-Julian oscillations in the tropics is realized.



Miyamoto et al (2013) , Geophys. Res. Lett., 40, 4922–4926, doi:10.1002/grl.50944.





Fugaku's FUjitsu A64fx Processor is...



Memor

Memor

PCle

Controller

Memory

Tofu

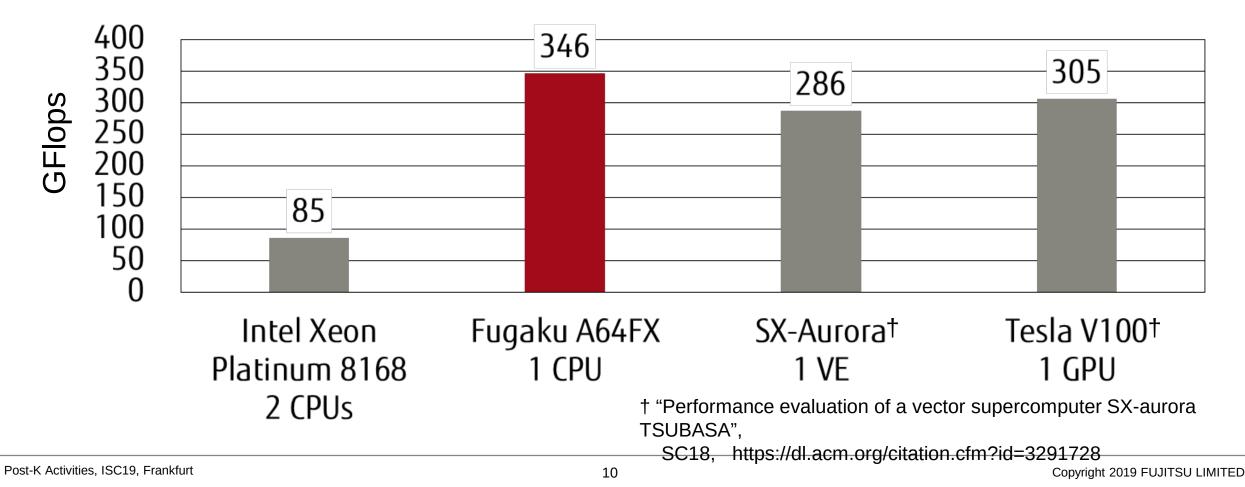
Interface

- an Many-Core ARM CPU...
 - 48 compute cores + 2 or 4 assistant (OS) core
 - Brand new core design
 - Near Xeon-Class Integer performance core
 - ARM V8 --- 64bit ARM ecosystem
 - Tofu-D + PCIe 3 external connection
- ...but also an accelerated GPU-like processor
 - SVE 512 bit x 2 vector extensions (ARM & Fujitur)
 - Integer (1, 2, 4, 8 bytes) + Float (16, 32, 64 bytes)
 - Cache + memory localization (sector cache)
 - HBM2 on package memory Massive Mem BW (Bytes/DPF ~0.4)
 - Streaming memory access, strided access, scatter/gather etc.
 - Intra-chip barrier synch. and other memory enhancing features
- GPU-like High performance in HPC especially CFD-- Weather & Climate (even with traditional Fortran code) + AI/Big Data

"Fugaku" CPU Performance Evaluation (2/3)

FUjitsu

- Himeno Benchmark (Fortran90)
 - Stencil calculation to solve Poisson's equation by Jacobi method



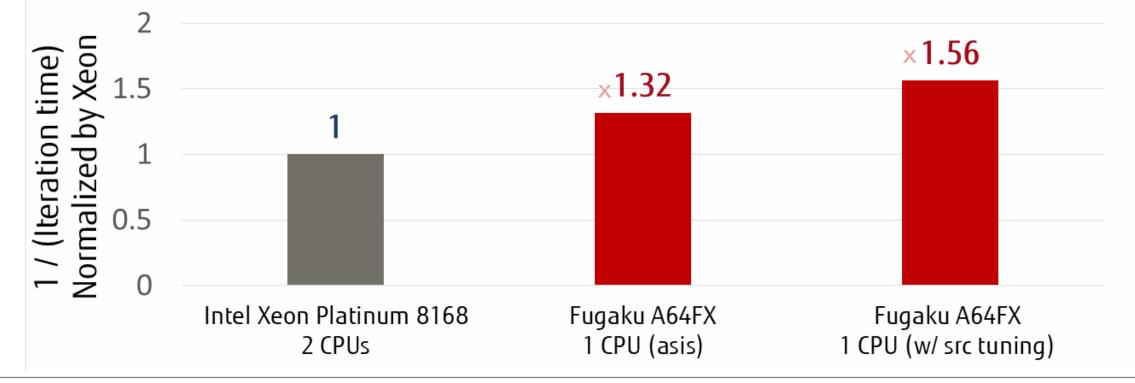
"Fugaku" CPU Performance Evaluation (3/3)

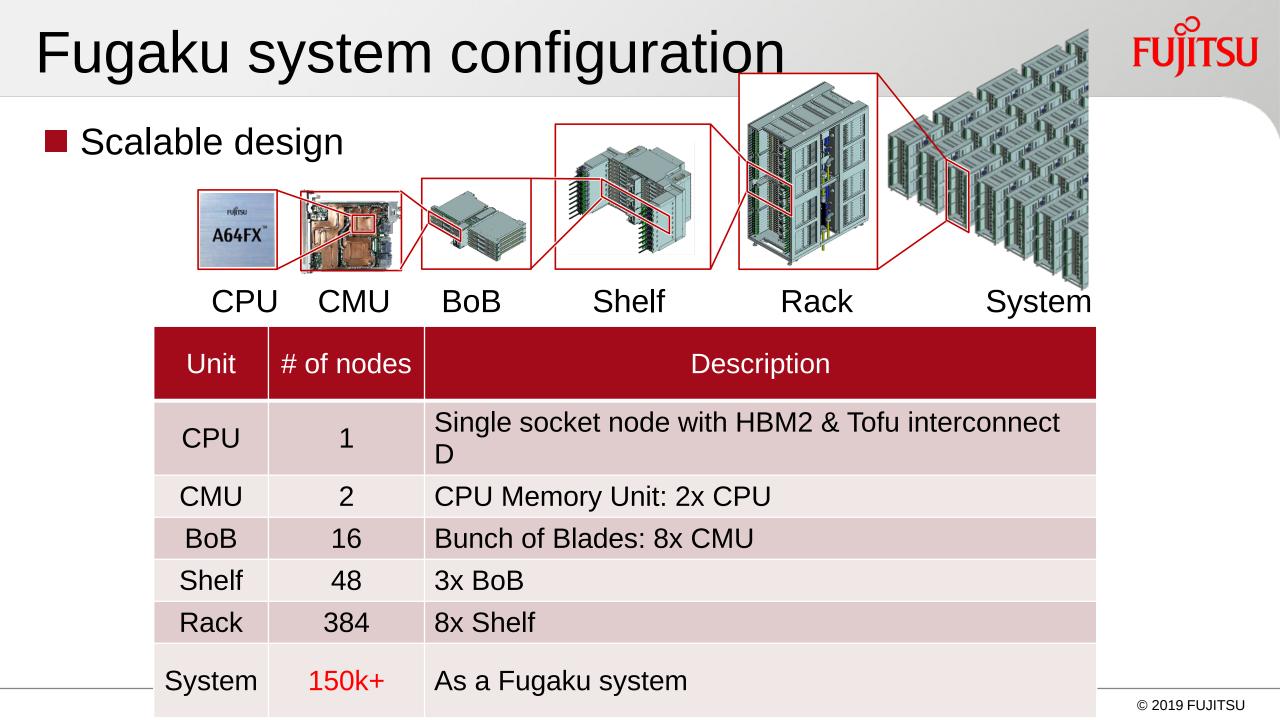


WRF: Weather Research and Forecasting model

- Vectorizing loops including IF-constructs is key optimization
- Source code tuning using directives promotes compiler optimizations

WRF v3.8.1 (48-hour, 12km, CONUS) on 48 cores





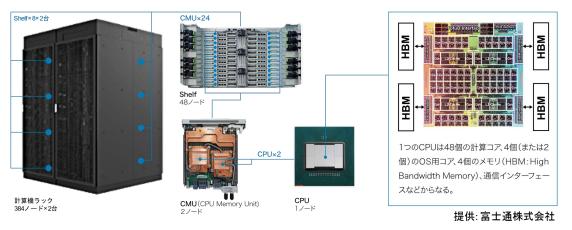


Fugaku Total System Config & Performance



• Total # Nodes: 158,976 nodes

- 384 nodes/rack x 396 (full) racks = 152,064 nodes
- 192 nodes/rack x 36 (half) racks = 6,912 nodes
 c.f. K Computer 88,128 nodes
- Theoretical Peak Compute Performances
 - Normal Mode (CPU Frequency 2GHz)
 - 64 bit Double Precision FP: 488 Petaflops
 - 32 bit Single Precision FP: 977 Petaflops
 - 16 bit Half Precision FP (AI training): 1.95 Exaflops
 - 8 bit Integer (AI Inference): 3.90 Exaops
 - Boost Mode (CPU Frequency 2.2GHz)
 - 64 bit Double Precision FP: 537 Petaflops
 - 32 bit Single Precision FP: 1.07 Exaflops
 - 16 bit Half Precision FP (AI training): 2.15 Exaflops
 - 8 bit Integer (Al Inference): 4.30 Exaops
- Theoretical Peak Memory Bandwidth: 163 Petabytes/s



<u>C.f. K Computer performance comparison (Boost)</u>

- 64 bit Double Precision FP: 48x
- 32 bit Single Precision: 95x
- 16 bit Half Precision (AI training): 190x
 - K Computer Theoretical Peak: 11.28 PF for all precisions
- 8 bit Integer (Al Inference): > 1,500x
 - K Computer Theoretical Peak: 2.82 Petaops (64 bits)
- Theoretical Peak Memory Bandwidth: 29x
 - K Computer Theoretical Peak: 5.64 Petabytes/s

Fugaku is a Year's worth of IT in Japan R-

6				
0				
7				
\leq	2	Ξ	=	
H-	•	r :	S	
		-	_	

	Smartphones		IDC Servers incl Clouds		Fugaku	K Computer	
Units	20 million (2/3 annual shipments in Japan)	=	300,000 (2/3 annual shipments in Japan)	11	1	30~100	
Power	10W×20 mil = 200MW		600-700W x 30K= 200MW (incl cooling)		30MW	15MW	
CPU ISA System SW	Arm iOS/ Android Linux		x86/Arm Linux (Red Hat etc.)/Win		Arm Linux (Red Hat etc.)	Sparc Proprietary Linux Low generality	
AI Acceleration	Custom ASIC Inference Only		Gen. Purpos Accelerator e.g. GPU		Gen. CPU SVE instructions	None	

RIKEN

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



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 TOF
- The No 3 on the Green500 is AiMOS, a new IBM Power systems at the Rensselaer Polytechnic Ins Computational Innovations (CCI), New York, USA. It achieved 15.8 GFlops/Watt and is listed at post

Green500 List for November 2019

Listed below are the November 2019 The Green500's energy-efficient supercomputers ranked fro

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



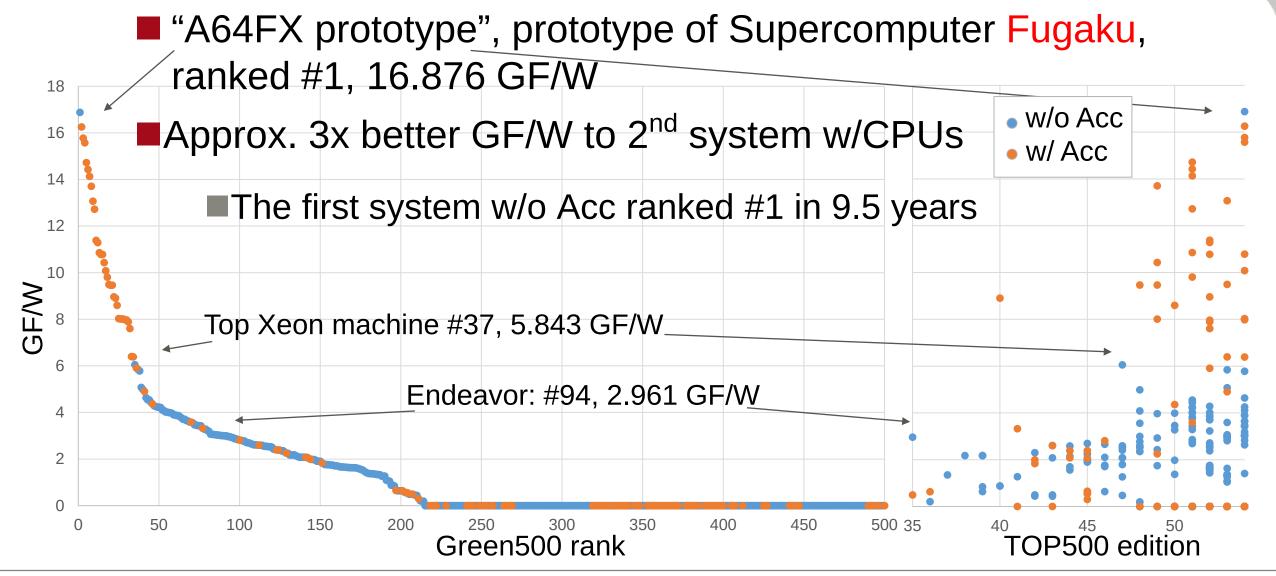


京

富岳省エネ世界一

SC19 Green500 ranking and 1st appeared TOP500 edition





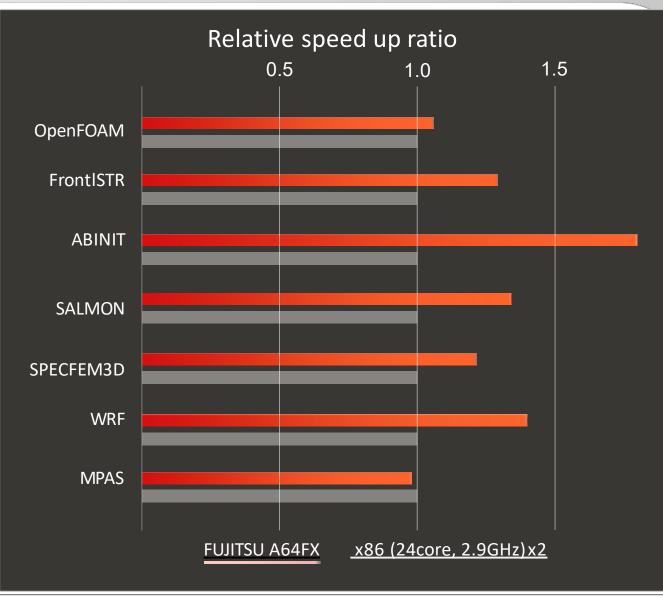
A64FX CPU performance evaluation for real apps



Open source software, Real apps on an A64FX @ 2.2GHz

Up to 1.8x faster over the latest x86 processor (24core, 2.9GHz) x 2, or 3.6x per socket

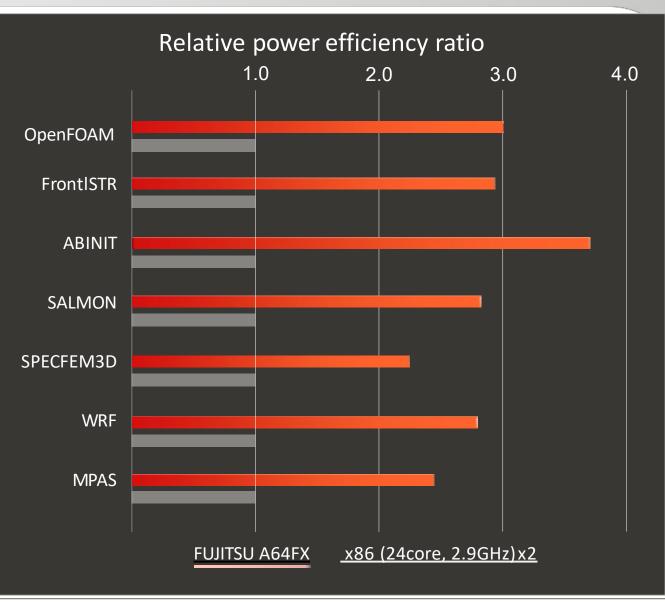
High memory B/W and long SIMD length of A64FX work effectively with these applications



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 Performance Estimate on 9 Co-Design Target Apps



Performance target goal				Priority Issue Area	Performance Speedup over K	Application	Brief description
 ✓ 100 times faster than K for some applications (tuning included) ✓ 30 to 40 MW power consumption □ Peak performance to be achieved 			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)
	PostK	К	Di: preve Envii	3. Integrated simulation systems induced by earthquake and tsunami	45x +	GAMERA	Earthquake simulator (FEM in unstructured & structured grid)
Peak DP (double precision)	>400+ Pflops (34x +)	11.3 Pflops	Disaster prevention and Environment	4. Meteorological and global environmental prediction using	120x +	NICAM+	Weather prediction system using Big data (structured grid stencil &
Peak SP (single precision)	>800+ Pflops (70x +)	11.3 Pflops		big data 5. New technologies for		LETKF	ensemble Kalman filter) Molecular electronic simulation
Peak HP	>1600+ Pflops		Energy issue	energy creation, conversion / storage, and use	40x +	NTChem	(structure calculation)
(half precision) Total memory bandwidth	(141x +) >150+ PB/sec (29x +)	5,184TB/sec	issue	6. Accelerated development of innovative clean energy systems	35x +	Adventure	Computational Mechanics System for Large Scale Analysis and Design (unstructured grid)
Geometric Mean of Performance Speedup of the 9 Target Applications over the K-Computer			Industrial competitiveness enhancement	7. Creation of new functional devices and high-performance materials	30x +	RSDFT	Ab-initio simulation (density functional theory)
			trial iveness ement	8. Development of innovative design and production processes	25x +	FFB	Large Eddy Simulation (unstructured grid)
	<u>> 37x+</u>	As of 2019/05/14	Basic science	9. Elucidation of the fundamental laws and evolution of the universe	25x +	LQCD	Lattice QCD simulation (structured grid Monte Carlo)

RIKEN



Fugaku / Fujitsu FX1000 System Software Stack



Fugaku Al RIKEN: Chainer, PyTo	(DL4Fugaku) rch, TensorFlow, DNI	NL		e Data Ai ie Flink, I				~3000 Apps sup- ported by Spack
Math Libraries Fujitsu: BLAS, LAPACK, ScaLAPACK, SSL II RIKEN: EigenEXA, KMATH_FFT3D, Batched BLAS,,,,		BLAS,,,,	Cloud Software Stack OpenStack, Kubernetis, NEWT					Open Source Management Tool
Compiler an Fortran, C/C++, Ope (Multiple Compilers s LLVM/CLANG, PGI,	suppoted: Fujitsu, Ar	B	5	and Manager System ical File Syste		Object S3 Com		Spack
	l Debugging Tools iler, Debugger, GUI			t Enterpris				Most applications will work
High-level Prog. Lang. XMP	Domain Spec. Lang. FDPS	Communication Fujitsu MPI RIKEN MPI		ile I/O DTF		zation & C M, Singula		with simple recompile from x86/RHEL environment.
Process/Thread PIP		Communicat ofu, LLC	tion	File I/O for Hierarchical Storage Lustre/LLIO		orage	LLNL Spack automates this.	
Red Hat E	nterprise Linux Ke	rnel+ option	nal light-	weight ker	mel (M	cKernel)		



The HPE/Cray CS500 - Fujitsu A64FX Arm Server



- Cray Fujitsu Technology Agreement
 Supported in Cray CS500 infrastructure
- Full Cray Programming Environment
- Leadership performance for many memory intensive HPC applications, e.g., weather
- GA in mid'2020
- A number of adoptions US: Stony Brook, DoE Labs, etc. Multiple yet-to-be-named EU centers









- Pipelined manufacturing, installation, and bringup, first rack shipped on Dec 3 2019.
- All racks on the floor May 13, 2020(!)
- 2020 early users, incl. COVID-19 apps running already
- Open to international users through HPCI, general allocation April 2021 (application starting Sept. 2020) (does not need to involve a Japanese PI)
- Also some internal test nodes (Apr 2020) and allocations (Apr. 2021) are available for R-CCS



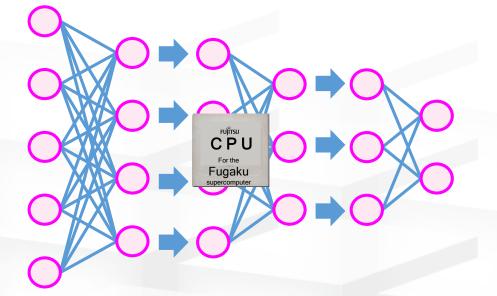
Massive Scale Deep Learning on Fugaku



Fugaku Processor is AI-DL ready

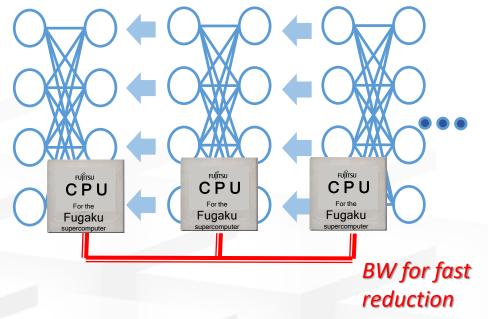
- High perf FP16&Int8
- High mem BW for convolution
- Built-in scalable Tofu network

High Performance DNN Convolution



Unprecedened DL scalability

High Performance and Ultra-Scalable Network for massive scaling model & data parallelism



Low Precision ALU + High Memory Bandw idth + Advanced Combining of Convolutio n Algorithms (FFT+Winograd+GEMM)

Ultra Scalability of Data/Model Parallelism Scalability shown to 20,000 nodes HPL-AI to beyond exaflops 24



Large Scale Public Al Infrastructures in Japan

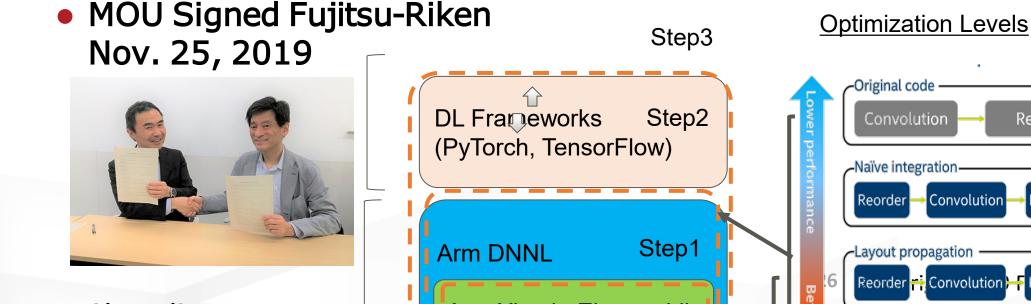


1									R-CC.
		Deployed	Purpose	AI Processor	Inference Peak Perf.	Training Peak Perf.	HPL-AI Perf	Top500 Perf/Rank	Green500 Perf/Rank
	Tokyo Tech. TSUBAME3	July 2017	HPC + Al Public	NVIDIA P100 x 2160	45.8 PF (FP16)	22.9 PF / 45.8PF (FP32/FP16)		8.125 PF #22	13.704 GF/W #8
	U-Tokyo Reedbush-H/L	Apr. 2018 (update)	HPC + Al Public	NVIDIA P100 x 496	10.71 PF (FP16)	5.36 PF / 10.71PF (FP32/FP16)		(Unranked)	(unranked)
	U-Kyushu ITO-B	Oct. 2017	HPC + Al Public	NVIDIA P100 x 512	11.1 PF (FP16)	5.53 PF/11.1 PF (FP32/FP16)		(Unranked)	(Unranked)
it	AIST-AIRC AICC	Oct. 2017	Al Lab Only	NVIDIA P100 x 400	8.64 PF (FP16)	4.32 PF / 8.64PF (FP32/FP16)		(Unranked)	(Unranked)
	Riken-AIP Raiden	Apr. 2018 (update)	AI Lab Only	NVIDIA V100 x 432	54.0 PF (FP16)	6.40 PF/54.0 PF (FP32/FP16)		1.213 PF #462	(Unranked)
	AIST-AIRC ABCI	Aug. 2018	Al Public	NVIDIA V100 x 4352	544.0 PF (FP16)	65.3 PF/544.0 PF (FP32/FP16)		19.88 PF #8	14.423 GF/W #6
	NICT & Sakura Internet	Summer 2019	Al Lab Only	NVIDIA V100 x 1700	~210 PF (FP16)	~26 PF/~210 PF (FP32/FP16)		4.128 #51 3.712 #58	(Unranked)
	C.f. US ORNL Summit	Summer 2018	HPC + Al Public	NVIDIA V100 x 27,000	3,375 PF (FP16)	405 PF/3,375 PF (FP32/FP16)	445 PF (FP16)	143.5 PF #1	14.719 GF/W #5
	Riken R-CCS Fugaku	2018 2020 ~2021	HPC + AI Public	Fujitsu A64fx > x 150,000	> 4000 PO (Int8)	>1000PF/>2000PF (FP32/FP16)	<pre>>1000PF (FP16)</pre>	*1 > 400PF #1 (2020?)	#5 16.876 GF/W #1 (proto)

(equiv. ~100K GPUs)



Fujitsu-Riken-Arm joint effort on AI framework development on SVE/A64FX

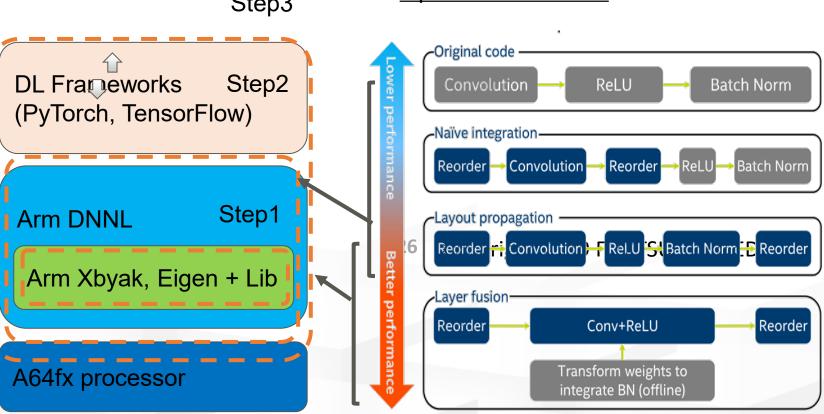


Also w/Arm

R

אוא=א

- 1st release May 2020
 - First ver. optimized for inference



Next ver. training Exaops of sim, data, and AI on Fugaku and Cloud



Cloud Service Providers Partnership

https://www.r-ccs.riken.jp/library/topics/200213.html (in Japanese)



Action Items

R

RIKEN

- Cool Project name and logo!
- Trial methods to provide computing resources of Fugaku to end-users via service providers
- Evaluate the effectiveness of the methods quantitatively as possible and organize the issues

COVID19 Fugaku Early Production & HPCI



- Supercomputer Fugaku used to help fight against COVID-19
 - https://www.r-ccs.riken.jp/en/topics/fugaku-coronavirus.html
 - Production a year ahead of schedule
 - Max availability in April~May: 72 racks (80 Petaflops), 1/6 full
 - More to be added after June

6

- Public call by MEXT and Riken fast track, bypasses normal allocation procedure, priority allocation to massive resource w/extensive R-CCS support
- Must work closely w/Riken R-CCS and other COVID19 groups
- HPCI Tier-2 COVID19 resource allocations
 - https://www.hpci-office.jp/pages/e_hpci_covid19
 - Most HPCI tier-2 resources, more variety (e.g., GPUs), less capacity & less restrictions c.f. Fugaku
 - Goes thru accelerated peer-review process by RIST
- Both are available to international research groups









Prediction and Countermeasure for Virus Droplet Infection under the Indoor Environment

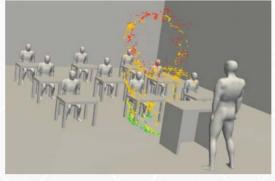
RIKEN R-CCS Makoto TSUBOKURA

Outline of the Research:

Virus droplet infection caused by sneezing, coughing, or talking is strongly influenced by the flow, temperature and humidity of the air around an infected person and potential victims. Especially in the case of the new coronavirus, possibility of aerosol infection by atomized droplets is suggested in addition to the usual droplet infection. Because smaller aerosol particles drift in the air for a longer time, it is imperative to predict the scattering route and to estimate how surrounding airflow affects the infection so that the risk of droplet infection can be properly assessed, and effective measures to reduce infection can be proposed. In this project, massively parallel coupling simulation of virus droplet scattering, with airflow and heat transfer under

the indoor environment such as inside a commuter train, offices, classrooms, and hospital rooms will be conducted. By taking into account the characteristics of the virus, its infection risk of virus droplets is assessed under various conditions. Then countermeasures to reduce the risk are proposed from a viewpoint of controlling the air flow.

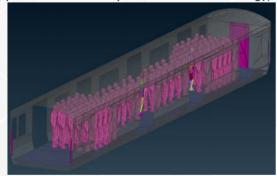
This project is a collaboration with RIKEN, Kyoto Institute of Technology, Kobe University, Osaka University, Toyohashi University of Technology, and Kajima Corporation. Complex Unified Simulation framework called CUBE, developed at RIKEN R-CCS and implemented on the supercomputer Fugaku, is mainly used, which will be the world-largest and highly accurate virus droplet simulation ever conducted.



An Example of virus droplet simulation in a classroom (By prof. Yamakawa of Kyoto Institute of Technology)

Expected Achievements:

The risk of droplet infection under the indoor environment is quantitatively evaluated, and specific countermeasures to reduce the infection risk is proposed in terms of effective ways of opening/closing windows, use of air conditioning, and placement of partitions. In addition, by creating animation of the droplet scattering and its spreading speed in the rooms from the simulation results, people can visually understand the risk of droplet infection and its countermeasures. These outputs from the simulation can protect the living and working environment from virus droplet infection, and contribute to earlier recovery of the socio-economic activities.



Simulation model of a cabin of a commuter train







R-CCS

Simulation analysis of pandemic phenomena

RIKEN Nobuyasu Ito

Research content:

Social and economic impact is increasing globally, and Japan is now at critical bifurcation point. And challenges to make its visualization and "big data" mining have started. In this project, making the most of the "Fugaku" and other supercomputers, estimations of possible future of our social and economic activities, and policy options to control and resolve the situation. For the purpose, simulations of disease propagation and economic activities, and SNS text mining are applied together with the National Institute of Advanced Industrial Science and Technology,Kyoto University, Tokyo Institute of Technology, the University of Hyogo, the University of Ryukyus and the University of Tsukuba.

1.0-0.80.8-0.6

0.6-0.4

0 2-0 0

Expected results:

- Candidates of policy options to control and resolve the disease propagation and its social and economic effects are visible.
- Dynamic control of the situation together with localized policy will be clear.
- Not only in case of disease propagation, policy options in cases of large scale disasters and accidents will also be guided.

Preliminary simulation result of economic damage after lockdown of Tokyo area: the left figure shows the first days and the right 14th day by Dr. Inoue of the University of Hyogo. In this project, not only the case of lockdown, but also partia restrictions in various areas are searched using the "Fugaku" supercomputer.









Exploring new drug candidates for COVID-19 by "Fugaku"

RIKEN / Kyoto University Yasushi OKUNO, Prof. PhD.

Research content:

Currently, clinical trials are underway in Japan and overseas to confirm the effects of existing drugs on COVID-19. Some reports have shown that the drug has shown efficacy through these clinical trials, but the number of cases has been small, and no effective therapeutic drug has yet been identified. Furthermore, due to the small number of drugs being tested, it is possible that none of the drugs have a definite effect.

Therefore, in this study, we performe molecular dynamics calculations using "Fugaku" to search and identify therapeutic drug candidates showing high affinity for the target proteins of COVID-19 from approximately 2,000 existing drugs that are not limited to existing antiviral drugs targeted in clinical trials.

Expected results:

- New therapeutic drug candidates other than those currently undergoing clinical trials can be discovered.
- Combination effects of multiple drugs can be estimated
- The molecular action mechanism of existing drugs currently undergoing clinical trials will be elucidated. In addition, these findings provide a clear direction for developing new drugs that go beyond the existing drugs.

