EU-JAPAN DIGITAL WEEK 2025

31 MARCH – 7 APRIL, 2025

TOKYO, JAPAN

THE EU-JAPAN DIGITAL WEEK IS ORGANISED AS PART OF THE EU-JAPAN DIGITAL PARTNERSHIP

SESSION 1: INNOVATIONS IN JAPAN'S CLOUD–EDGE LANDSCAPE

Fugaku and Beyond:
Cloud-Enabled Supercomputer bridging Europe and Japan

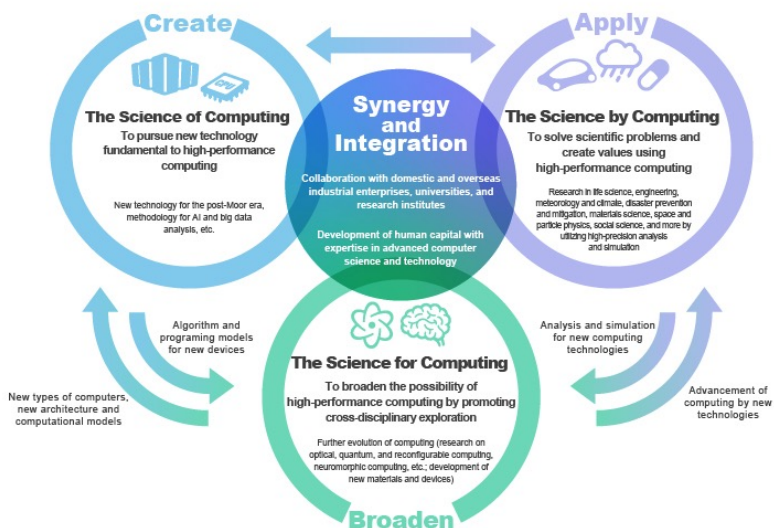"Smart Connectivity and Computing" Workshop, EU-JAPAN DIGITAL WEEK 2025
March 31, 2025

Kento Sato, RIKEN R-CCS

- **RIKEN R-CCS is one of 13 RIKEN strategic research and infrastructure centers**
- **Mission in R-CCS:**
  - Conduct the state-of-the-art research that integrates "simulation," "big data analysis," and "AI" through high performance computing (HPC) to solve scientific and social issues and revolutionize society
- **Three R&D thrusts in R-CCS:**
  - Cultivate the future through "Science of computing, by computing, and for computing"
  - Integrate Computational Science, Data Science, and AI
  - Aim to Solve Issues for Advanced Modern Society

**"The Science of Computing, by Computing, and for Computing"**

Striving for excellence in science and becoming the cornerstone of Society 5.0

Create

The Science of Computing
To pursue new technology fundamental to high-performance computing

New technology for the post-Moor era, methodology for AI and big data analysis, etc.

Synergy and Integration

Collaboration with domestic and overseas industrial enterprises, universities, and research institutes

Development of human capital with expertise in advanced computer science and technology

Apply

The Science by Computing
To solve scientific problems and create values using high-performance computing

Research in life science, engineering, meteorology and climate, disaster prevention and mitigation, materials science, space and particle physics, social science, and more by utilizing high-precision analysis and simulation

Algorithm and programing models for new devices

Analysis and simulation for new computing technologies

New types of computers, new architecture and computational models

Advancement of computing by new technologies

The Science for Computing
To broaden the possibility of high-performance computing by promoting cross-disciplinary exploration

Further evolution of computing (research on optical, quantum, and reconfigurable computing, neuromorphic computing, etc.; development of new materials and devices)

Broaden

Tokyo

Kobe

「富岳」 Supercomputer Fugaku

Peak Performance	Normal Mode: 2.0 GHz	488 PFLOPS (FP64)
		977 PFLOPS (FP32)
		1.95 EFLOPS (FP16)
		3.90 EOPS (INT8)
	Boost Mode: 2.2 GHz	537 PFLOPS (FP64)
		1.07 EFLOPS (FP32)
		2.15 EFLOPS (FP16)
		4.30 EOPS (INT8)
Total Memory		4.85 PiB
Total Memory Bandwidth		163 PB/s

Fugaku 3D Vritual Tour

Help | Terms Matterport
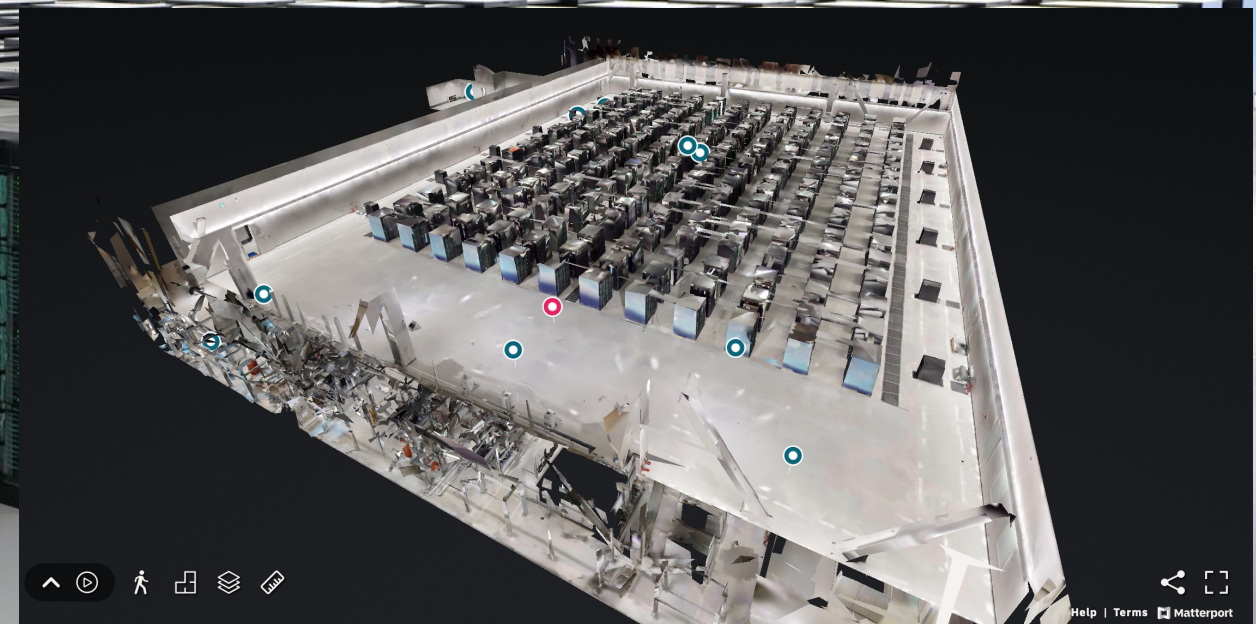
Major achievements of Fugaku

#1 in major benchmark rankings:
TOP500 and HPL-AI (Jun.2020-Nov.2021),
Graph500 and HPCG (Jun.2020-)

#1 in MLPerf HPC(Nov.2021-)

ACM Gordon Bell Special Prize for HPC based
COVID-19 research(Nov.2021), also 2022
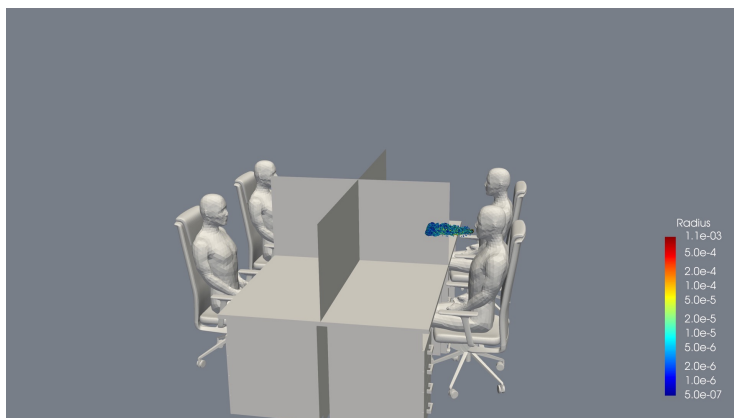
Weather forecasting trial for "guerrilla downpour"
in TOKYO2020 Olympic/Paralympic games
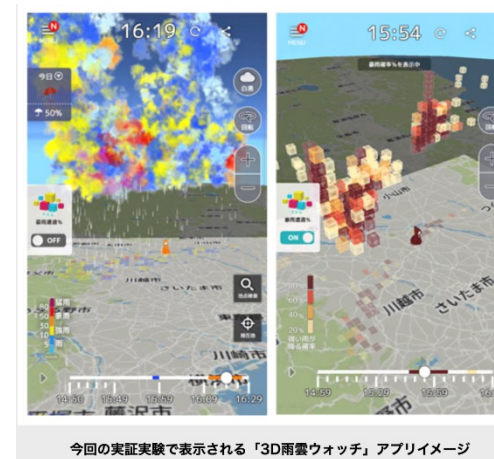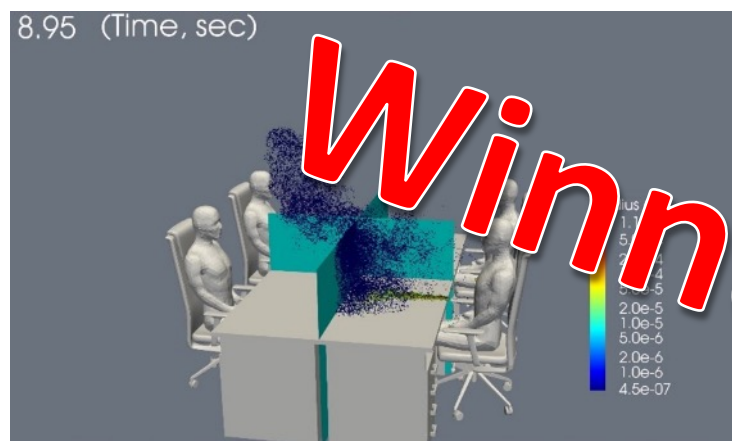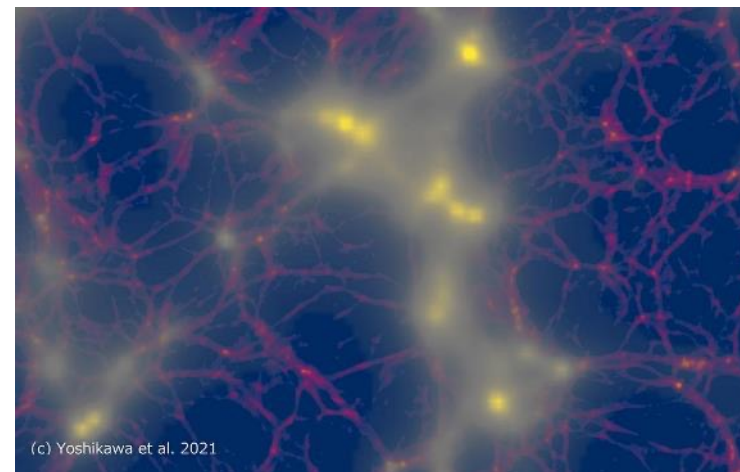
2021 ACM Gordon Bell Prize Finalists

The purpose of the award is to track the progress over time of parallel computing, with particular emphasis on rewarding innovation in applying high performance computing to applications in science, engineering, and large-scale data analytics.

3 of 6 research results presented as finalists for the 2022 Gordon Bell Prize use Fugaku.
One of them done by Earthquake Research Institute, The University of Tokyo,
Prof. Tsuyoshi Ichimura（R-CCS Visiting Researcher）

Gordon Bell Prize winner is presented at "SC" in November annually.

Winner!

"Pushing the Frontier in Laser-Based Electron Accelerators Design with Ground-Breaking Mesh-Refined Particle-In-Cell Simulations on Pre-Exascale Supercomputers"
International team； *(CEA),* *(Lawrence Berkeley National Laboratory), (Arm), (ATOS), (Laboratoire d'Optique Appliquée, ENSTA Paris), (GENCI) and (RIKEN, R-CCS)*

Extreme-scale many-against-many protein similarity search(Lawrence Berkeley National Laboratory)
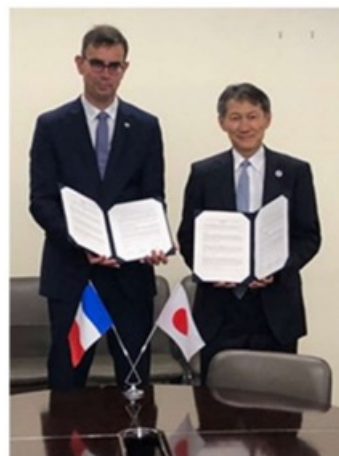
History of CEA-RIKEN collaborations:
RIKEN and CEA to cooperate in high performance computing and computational science [1, 2]

- **Both Japan and France have been carrying out national projects for the development of next-generation supercomputers used for simulations in various areas of science**
  - Under this framework, RIKEN and the French Alternative Energies and Atomic Energy Commission (CEA) have concluded an agreement to strengthen the partnership in the areas of high performance computing and computational science
  - On January 11, 2017, HPC collaboration agreement was signed by RIKEN and CEA in Tokyo
    - In June 2019, an amendment was then to formalize the addition of new collaboration topics.
  - On January 11, 2022, we extended the agreement for collaborative research for a new five-year period

Signature of the collaboration agreement Tokyo, January 2017

Signature of the 1st amendment Tokyo June 2019
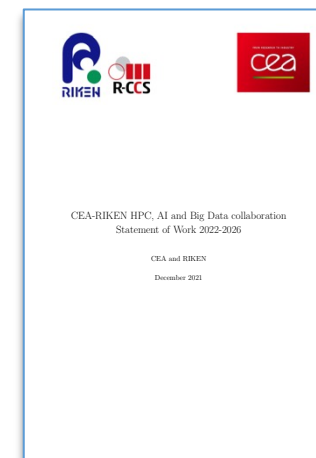
New SoW in Jan. 2022

[1] https://www.riken.jp/en/news_pubs/news/2017/20170118_1/
[2] https://www.r-ccs.riken.jp/wp/wp-content/uploads/2021/02/CEA-RIKEN-HPC.pdf

History of CEA-RIKEN collaborations:
RIKEN and CEA to cooperate in high performance computing and computational science [1, 2]

- **Under the agreement, RIKEN and CEA have been collaborating following areas:**
  - Training and Skills
  - System Management
  - Programming and Tools
  - Quantum Computing
  - AI and Big Data
  - Benchmarking new Architectures
  - Applications Development for Scientific and Societal Challenges

CEA-RIKEN Workshop 2023 in Paris
March 16-17, 2023

CEA-RIKEN Workshop 2023 in Kobe
December 7-8, 2023

CEA-RIKEN Workshop 2024 in Paris
May 21-22, 2024

CEA-RIKEN Workshop 2024 in Kobe
December 11-12, 2024

[1] https://www.riken.jp/en/news_pubs/news/2017/20170118_1/
[2] https://www.r-ccs.riken.jp/wp/wp-content/uploads/2021/02/CEA-RIKEN-HPC.pdf

MISSION:
GENERAL INFORMATION SCIENTIFIC AREAS CONSORTIUM

- **Support the implementation of the Digital Partnership in order to strengthen cooperation between EU-Japan**
  - Advance on the facilitation of mutual access of EuroHPC JU*1 and Japanese supercomputing resources for European and Japanese researchers.

- **Address the priority scientific domains in HPC**
  - Biomedical, materials science, weather and climate modeling, performance measuring, test and optimization for different architectures

- **Promote the exchange of researchers and engineers between Japan and the EU, and elaborate a roadmap for future actions to enhance cooperation**

- **Advise and promote a sustainable collaboration between Europe and Japan for the HPC community**

*1: EuroHPC: The European High Performance Computing Joint Undertaking

# GENERAL INFORMATION

- **The project focuses on:**
  - High-Performance Computing and AI
  - Climate Modeling
  - Materials Science
  - Biomedical Science

- **HANAMI in a nutshell:**
  - Started in March 2024
  - About 60 researchers involved
  - 1 Open Symposium per year
  - Travel grants for HANAMI researchers
  - 6 European Centres of Excellence (CoE) for HPC applications[*1] on board (MaX, EoCoE, TREX,RAISE, EsiWACE and PerMedCoE)

*1: CoE promote the use of upcoming exascale and extreme performance computing capabilities and scale up existing computing codes towards exascale performance.

Scientific Areas

Climate and Weather Modeling

Materials Science

Biomedical Sciences
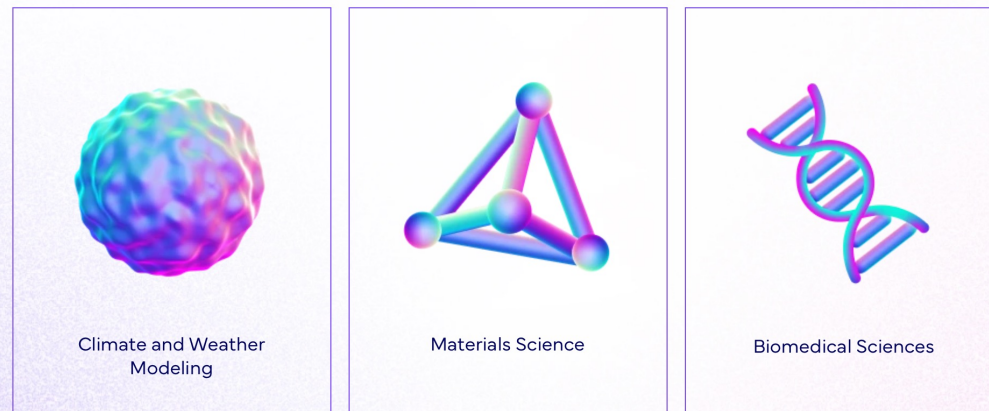
CONSORTIUM

- **HANAMI is gathering 15 research organizations in Europe and 11 in Japan**
  - HANAMI promotes scientific projects involving both Europe and Japanese institutes, and will assist the researchers to access supercomputers in both Japan and Europe

15 Organizations from Europe


11 Organizations from Japan

Many other 2-way EU/R-CCS collaborations with effective MOU or joint research agreement (as of March 31, 2025)

RIKEN R-CCS has been collaborating with 115 organizations
64 (domestic), 10 (US), 16 (EU), 25 (others) + many others without MOU/agreement

- Centre national de la recherche scientifique (CNRS), France
- Commissariat à l'énergie atomique et aux énergies alternatives (CEA), France
- IMT Atlantique, France
- Universite de Lille 1 Sciences et Technologies, France
- University of Vienna, Austria
- University of Warsaw, Porland
- Jagiellonian University, Porland
- Ludwig-Maximilians-University Munich, Germany
- Heidelberg Institute for Theoretical Studies, Germany

- INESC TEC, Portugal
- Barcelona Supercomputing Center, Spain
- Royal Institute of Technology (KTH), Sweden
- Friedrich-Alexander-Universitat Erlangen-Nurnberg (FAU), Germany
- J. Heyrovsky Institute of Physical Chemistry, Czech Academy of Sciences, Czech
- Faculty of Natural Sciences, Comenius University in Bratislava, Slovakia
- Donostia International Physics Center, University of the Basque, Spain
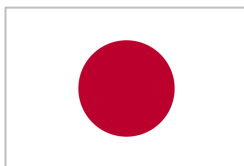
Virtual Fugaku:
MoU Between AWS & R-CCS (Effective: January 2023)

- **RIKEN R-CCS has been advancing usability of Fugaku to facilitate the collaboration**
  → **Expanding the Scientific Platforms of Fugaku to the AWS Cloud**

- **Compatible ISA in Arm eco-system**
  - Fugaku (by Fujitsu/RIKEN): A64FX HPC (Arm+SVE CPU), 2018
  - AWS (by Amazon): AWS Graviton3/3E (Arm+SVE CPU), 2022
  - → Applications are compatible on both Fugaku and AWS

A64FX   Graviton3

- **Virtual Fugaku platform**
  - Common interface provided on Fugaku and AWS
  - Users can enjoy computational power of both Fugaku and AWS through the same interface
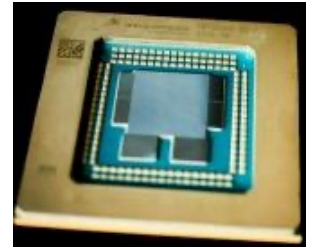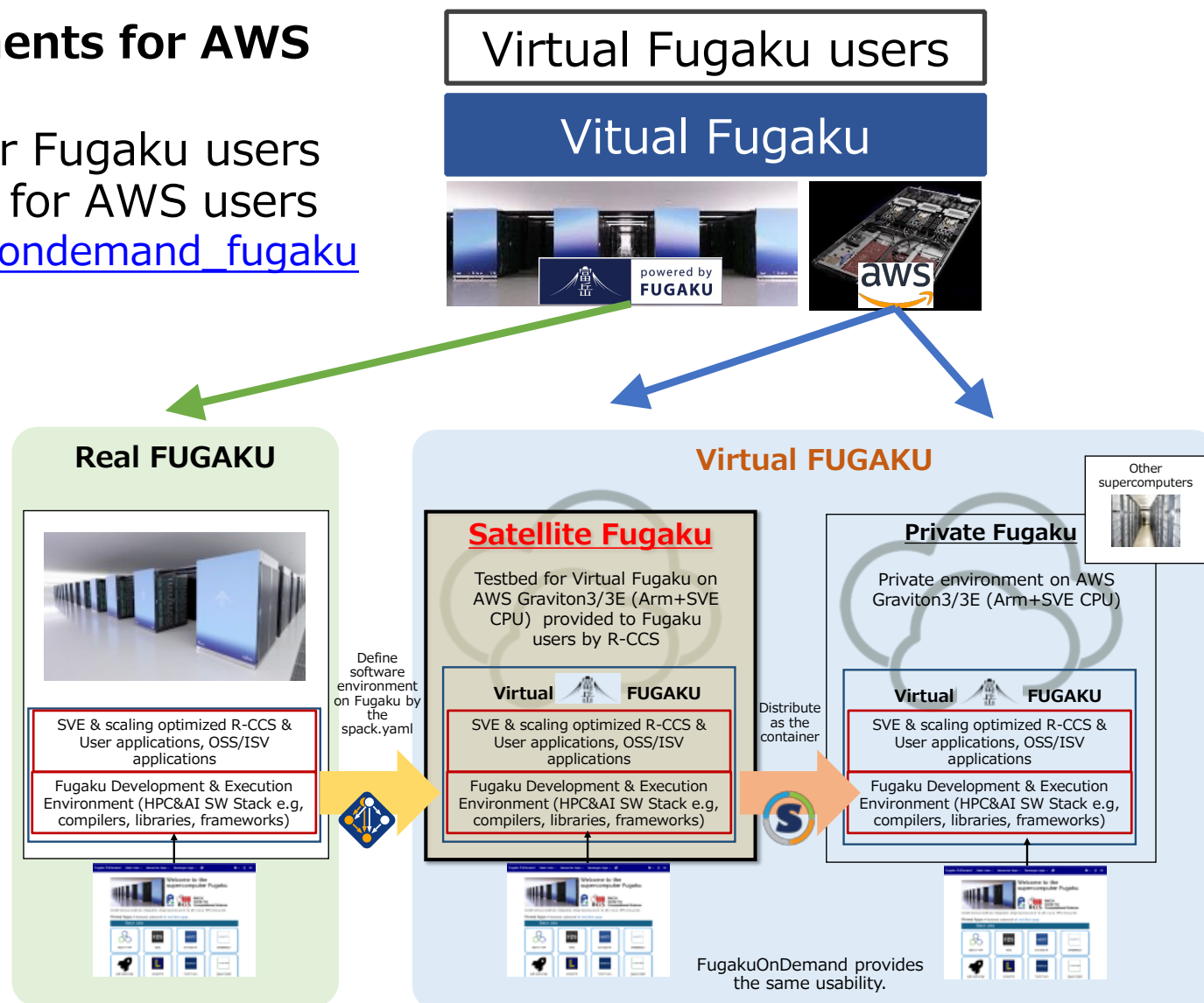
Open OnDeman

Virtual Fugaku users

Virtual Fugaku

# Virtual Fugaku enables gradual software/application dev.

- **Virtual Fugaku provides two environments for AWS Graviton CPUs:**
  - Satellite Fugaku: A test environment for Fugaku users
  - Private Fugaku: A Singularity container for AWS users
   - Github: https://github.com/RIKEN-RCCS/ondemand_fugaku

- **Both environments share the same software configuration**
  - Defined and containerized by Spack*
  - *Spack: a package mgmt. tool
- **Enables gradual software/application development**
  - E.g., Companies develop methods using massive Fugaku resource, conduct production run on AWS or private cloud
  → Allows immediate propagation of latest research results onto production

# Virtual Fugaku provides web portal powered by Open OnDemand

- **Open OnDemand is web portal for HPC clusters**
  - Developed mainly by the Ohio State University
  - Open-source software used by over 250 institutions worldwide
- **Use HPC systems from user's web browser**
  - No other software installation required
  - File upload/download, job submission, use of applications/services and monitoring by browser
- **Multi-level applications/services provision**
  - PaaS: GUI desktop, CLI terminal
  - SaaS: Many scientific applications, data services, develop env., visualization tools and others
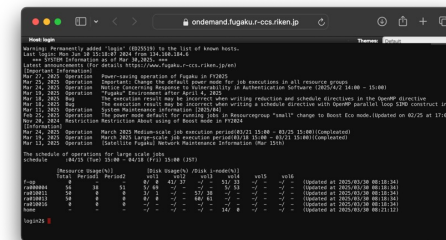

Virtual Fugaku users

Virtual Fugaku (By Open OnDemand)
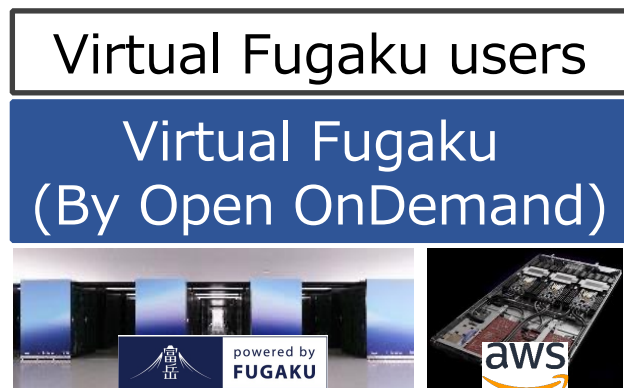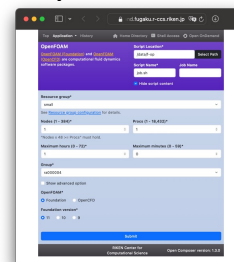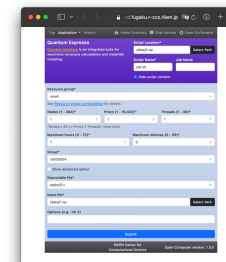
Open OnDemand
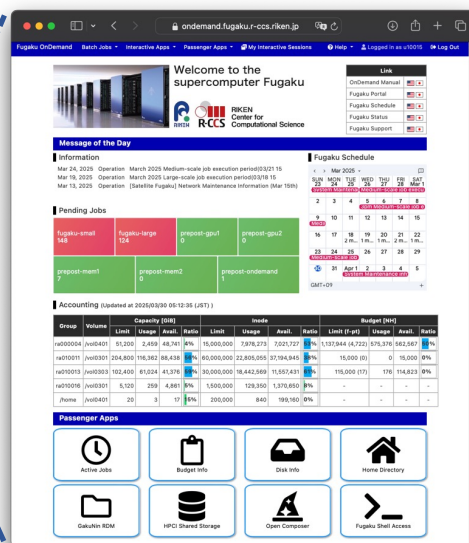
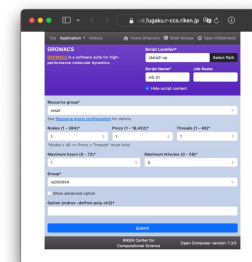GUI desktop env.
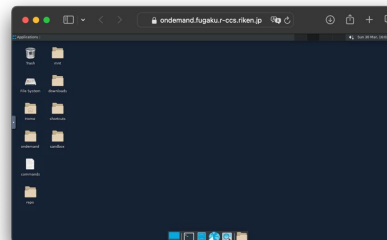
CLI terminal

OpenFOAM

Quantum Espresso

GROMACS

Supported Applications/Serivices in Virtual Fugaku (Release: Aug. 8, 2024))

- **Virtual Fugaku supports a wide variety of applications and servicies**
  - Interactive applications

Category	Application
Development	Remote Desktop, JupyterLab, MATLAB, RStudio, VSCode
Profiler	NVIDIA Visual Profiler, NVIDIA Nsight Compute*, NVIDIA Nsight Systems, Vampir
Viewer	AVS/Express, C-Tools, GaussView, ImageJ, OVITO, Paraview, PyMOL, SALMON view, Smokeview, VESTA, VMD, VisIt, XCrySDen
Workflow	WHEEL

  - Batch-based applications

Category	Application
Climate	SCALE
Computer Aided Engineering	FDS, FrontFlow (blue/X), FrontISTR, OpenFOAM (Foundation/OpenCFD)
Condensed Matter Physics	ALAMODE, AkaiKKR, HΦ, mVMC, OpenMX, PHASE/0, Quantum Espresso, SALMON
Molecular Dynamics	GENESIS, GROMACS, LAMMPS, MODYLAS
Quantum Chemistry	ABINIT-MP, Gaussian, NTChem, SMASH
Quantum Simulation	braket

RIKEN R-CCS keeps upgrading the virtual Fugaku platform to advance usability of supercomputers and the clouds and facilitate international collaborations

FugakuNEXT

FugakuNEXT is the project name for Japan's next-generation supercomputer as the successor to the current Fugaku system.

Overview of Feasibility Study (FS) in RIKEN System Research Team
(Team leader: Masaaki Kondo, RIKEN/Keio Univ.)

Project Overview

- **Aiming to design a versatile computing infrastructure**
  - Efficiently support wide varieties of computational requirements
  - E.g.) Scientific simulation, AI and Big data processing at scale
- **Conducting a holistic investigation on architecture, system software and library technologies through co-design with applications.**

Research DX platform by digital-twins

Higher performance

Wider application area

Research Groups

Architecture Group (Group Leader: Kentaro Sano, RIKEN)

- Investigating possible technologies (such as 3D stacked mem, new accelerators, chip-to-chip direct optical link) and evaluating architectures and the entire system based on trends in semiconductor and packaging technologies
- Extrapolating future system performance based on performance analysis of benchmark sets provided by Application Research Group, and feeding back to next-generation application development
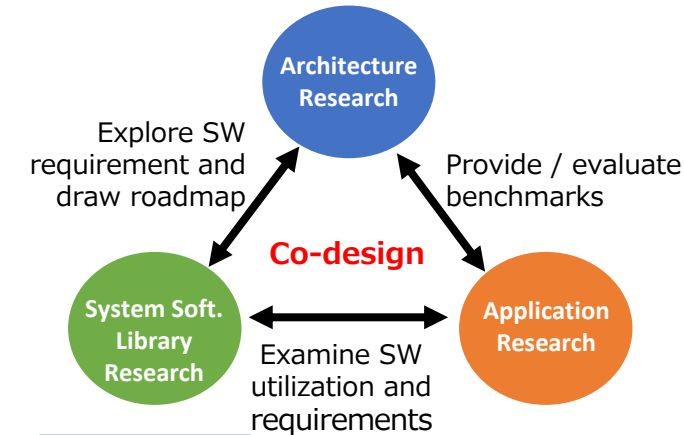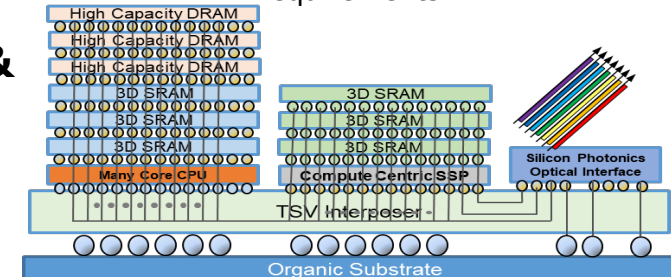
System Software and Library Group (Group Leader: Kento Sato, RIKEN)

- Making a roadmap for future system software development for the next system while considering large-scale scientific simulation and its integration with AI and big data technologies

### Applications Group (Group Leader: Takeshi Shimokawabe, Tokyo Univ. & Takeshi Iwashita, Hokkaido Univ.)

- Building a broad benchmark set to evaluate multiple architecture candidates while studying improvement of algorithms for candidates based on the results of architectural evaluations and conducting exploratory "what-if" performance analysis
- Investigating what classes of algorithms are expected to evolve significantly in near future

Architecture Research

Explore SW requirement and draw roadmap

Provide / evaluate benchmarks

Co-design

System Soft. Library Research

Application Research

Examine SW utilization and requirements

High Capacity DRAM
High Capacity DRAM
High Capacity DRAM
3D SRAM
3D SRAM
3D SRAM
Many Core CPU

3D SRAM
3D SRAM
3D SRAM
Compute Centric SSP

Silicon Photonics Optical Interface

TSV Interposer

Organic Substrate

Strawman processing element architecture

The research plans and the schedule are subjecct to change

Organization Chart of RIKEN FS team
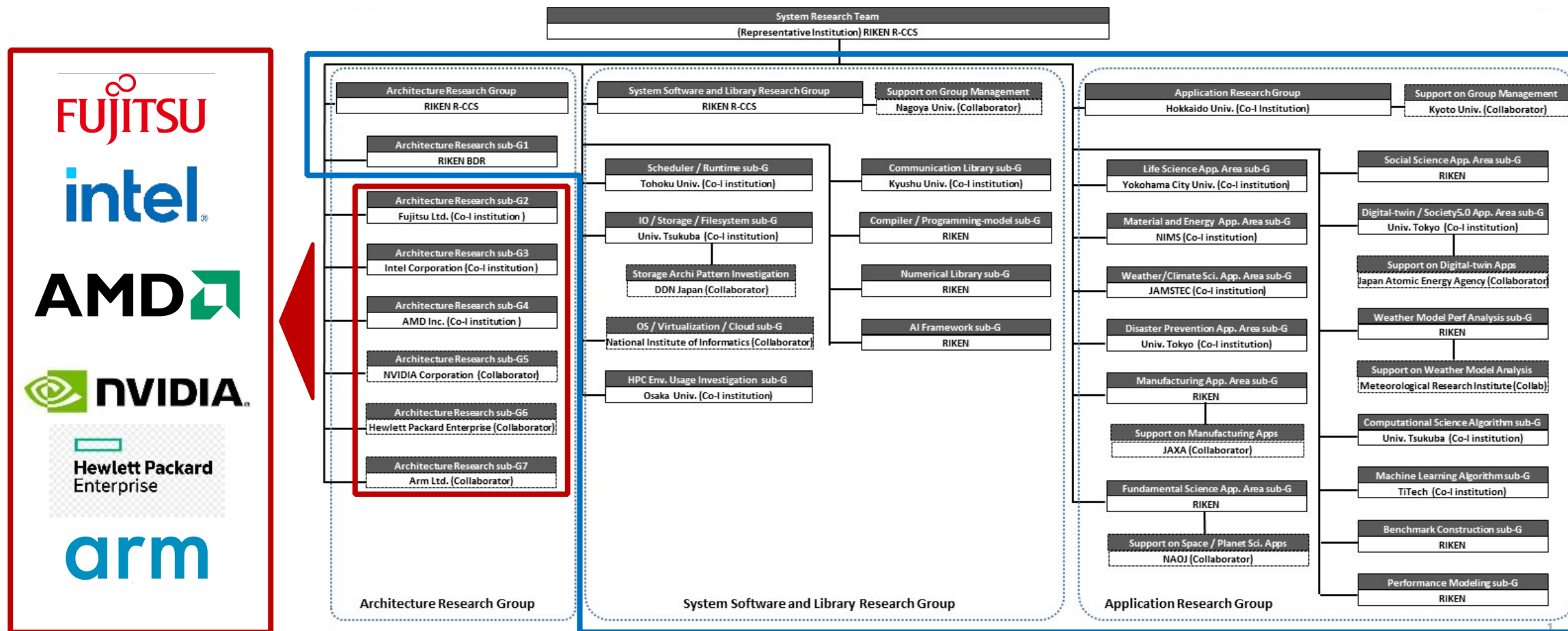
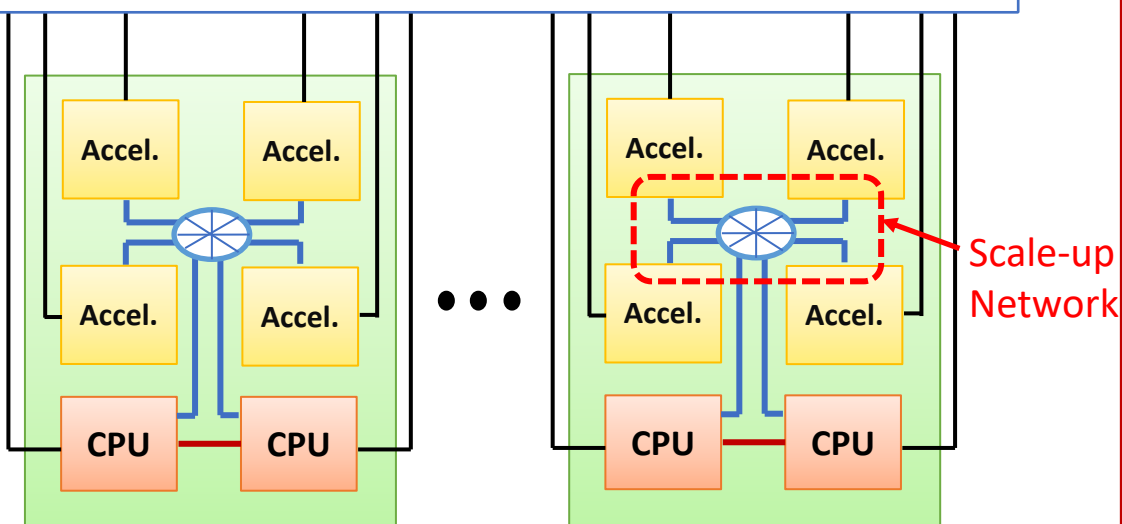**Collaborations with major chip vendors and system integrators**

ALL Japan alliance: Group and subgroups are organized by a wide rage of national labs, universityes in Japan

The research plans and the schedule are subjecct to change

Example Node Architecture for FugakuNEXT

System Architecture for FugakuNEXT

- **System network which is good for both strong/weak scaling**
  - Combination of scale-up/scale-out NW
- **Having more than 10K accelerator sockets in the system**
  - NW among accelerator sockets

- **Heterogeneous node architecture**
  - CPU + GPU architecture
  - Tentatively 2-CPU and 4-GPU configuration
  - High BW with advanced memory technology
- **Scale-up NW (intra-node socket NW)**
  - P2P or switched connection with, e.g., UALink
- **Scale-out NW (inter-node NW)**
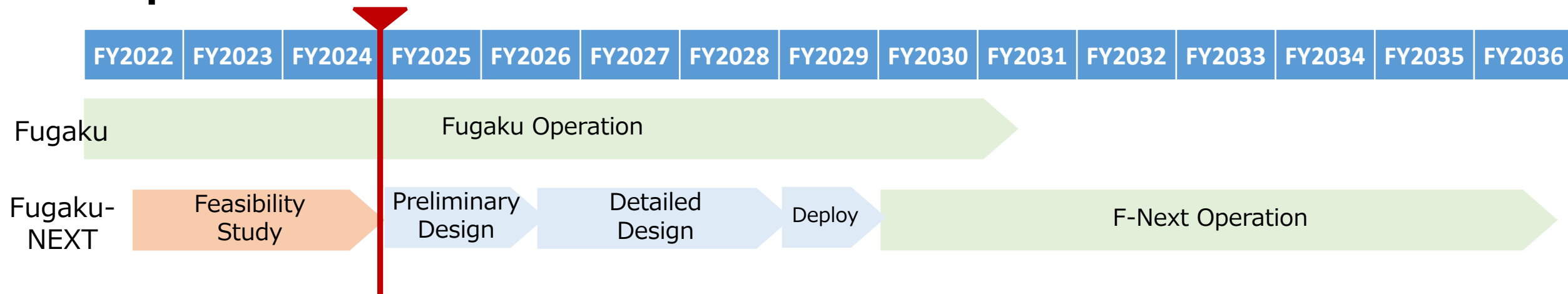 - Fat-tree topology with, e.g., Ultra-Ethernet

**System target:
More than 5-10x effective performance improvement in HPC applications and more than 50EFLOPS AI training performance (needs Zetta-scale low-precision arithmetic perf.)**
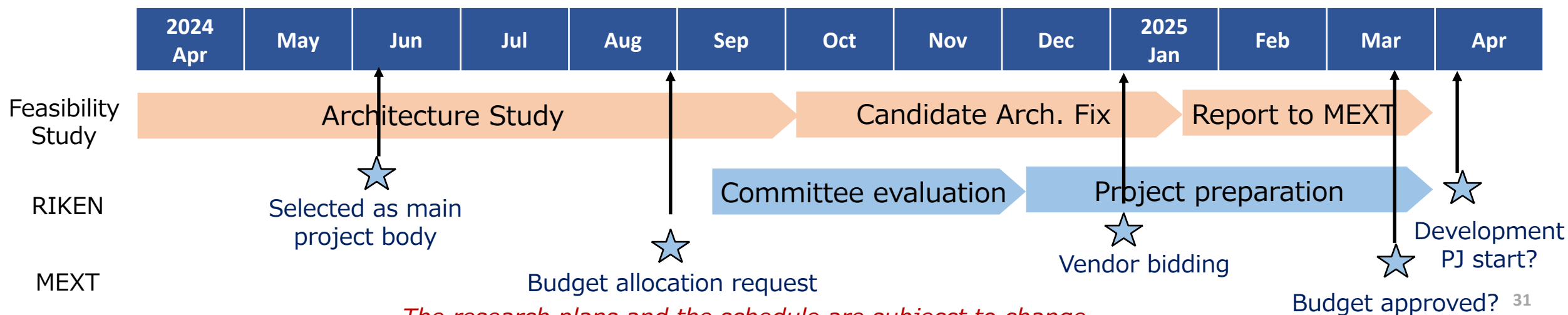
The research plans and the schedule are subjecct to change

System Performance Requirement in RFP

- **Performance requirement for FugakuNEXT entire system**

	CPU	GPU
Total Num. of Nodes	>= 3400 Nodes	
FP64 Vector FLOPS	>= 48PFLOPS	>= 3.0EFLOPS
FP16/BF16 AI FLOPS	>= 1.5EFLOPS	>= 150EFLOPS
FP8 AI FLOPS	>= 3.0ELOP	>= 300EFLOP
FP8 AI FLOPS (w/ sparsity)	—	>= 600EFLOPS
Memory Size	>= 10PiB	>= 10PiB
Memory Bandwidth	>= 7PB/s	>= 800PB/s
Total power consumption	< 40MW (compute node and storage)	

The research plans and the schedule are subjecct to change

Expected Timeline of FugakuNEXT R&D and Future Plan

Expected schedule

FY2022	FY2023	FY2024	FY2025	FY2026	FY2027	FY2028	FY2029	FY2030	FY2031	FY2032	FY2033	FY2034	FY2035	FY2036

Fugaku

Fugaku Operation

Fugaku-NEXT

Feasibility Study → Preliminary Design → Detailed Design → Deploy → F-Next Operation

What's going on in FY2024 for FugakuNEXT development

2024 Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	2025 Jan	Feb	Mar	Apr

Feasibility Study

Architecture Study → Candidate Arch. Fix → Report to MEXT

RIKEN

Committee evaluation → Project preparation

★ Selected as main project body

★ Vendor bidding

★ Development PJ start?

MEXT

★ Budget allocation request

★ Budget approved?

The research plans and the schedule are subjecct to change

31

Where will the FugakuNEXT supercomputer be installed ?

- **The FugakuNEXT system to be Installed next to RIKEN R-CCS**
  — Establishing a Cutting-Edge Hub for Computational and Computer Science Research (Announced on March 28, 2025)

RIKEN R-CCS, Kobe

FugakuNEXT

RIKEN R-CCS will drive cutting-edge supercomputing research in partnership with the global community

THANK YOU FOR YOUR ATTENTION!

The EU-Japan Digital Week is an initiative under the EU-Japan Digital Partnership and is supported by the following projects and organisations

ASPIRE 先端国際共同研究推進事業
Adopting Sustainable Partnerships for Innovative Research Ecosystem

BDV BIG DATA VALUE ASSOCIATION

DATA SPACES SUPPORT CENTRE

INPACE
Indo-Pacific-European Hub for Digital Partnerships

ENRICH GLOBAL
TAKING EUROPEAN INNOVATION GLOBAL

IOG 地経学研究所
Institute of Geoeconomics

公益財団法人 国際文化会館
International House of Japan

Japan Business Council in Europe

NEXUS FORUM.EU

東京大学
THE UNIVERSITY OF TOKYO

明治大学
MEIJI UNIVERSITY