Data Transfer Energy is all that matters in post-Exa machines

Satoshi Matsuoka, Director Riken R-CCS

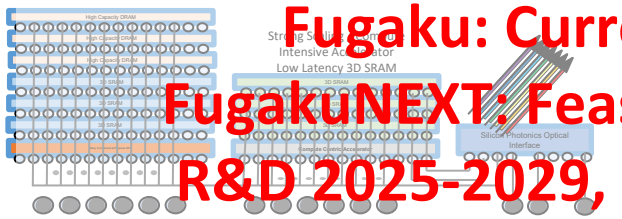
SC23 InPEX BOF

2023/11/15

RIKEN R-CCS

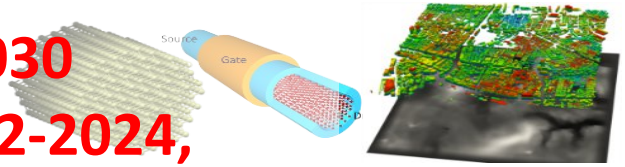
Riken-Intel Strategy for Innovation by Computing
Scientific Innovations are the 'Blue Ocean' in Computing

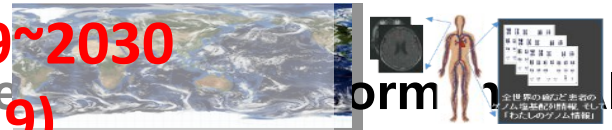
- Science **of** High Performance Computing (towards 'Zettascale')



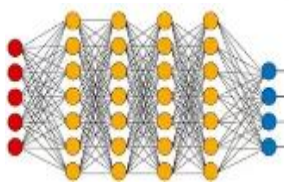
Fugaku: Current until 2029~2030
FugakuNEXT: Feasibility Study 2022-2024,
R&D 2025-2029, Deployment 2029~2030
$Billion R&D&D (FY 2022~2029)

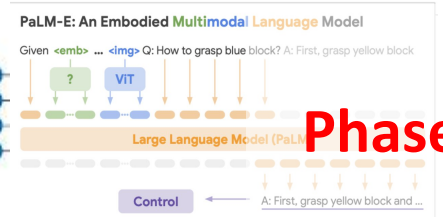
- Science **by** High Performance Computing





- Science **of** High Performance AI



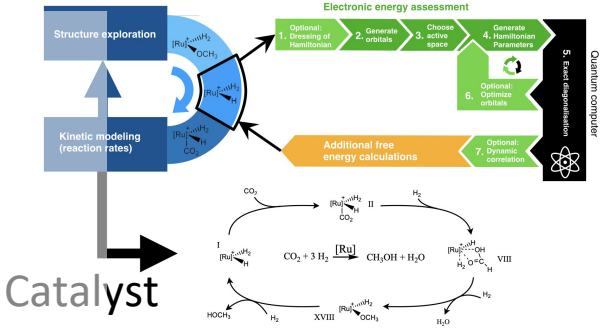


Riken AI for Science FY 2024~
Phased Infrastructure Deployment FY2024~2025
$X00 million (FY2024~2029)

- Science **by** Quantum-HPC Hybrid Computing

- Science **of** Quantum-HPC Hybrid Computing

Riken 'TRIP' Hybrid Quantum-HPC
Infrastructure Deployment FY2023~2025
$140 million+
(To be officially announced Nov. 1, 2023)



Overview of Study (System Research by RIKEN)

Project Overview

The next-generation computational infrastructure is expected to become a platform for realizing SDGs and Society 5.0 by **providing advanced digital twins** that will bring "Research DX" in the science. Aiming to realize a versatile computing infrastructure that can execute entire workflow by making full use of wide range of computational methods, simulation techniques, and BigData at scale, we conduct a holistic investigation on architecture, system software and library technologies through co-design with applications.

As a basic principle of system design, we **practice the "FLOPS to Byte" concept** from architecture development to algorithm or application design to streamline data transfer and computation under power constraints, while taking necessary computing accuracy into consideration. Under the ALL JAPAN team composition, we will investigate system configurations and elementary technologies which improve effective performance of the next-generation computing infrastructure.

Subject of Investigation

Research on Architecture

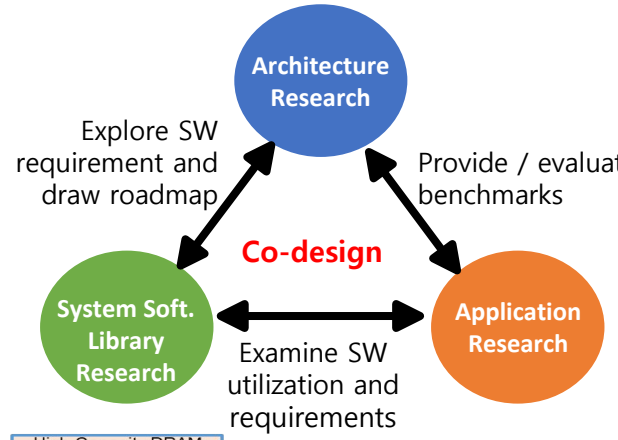
- Investigating technological possibilities (such as 3D stacked mem, accelerators, chip-to-chip direct optical link) and performance of the entire system or its components based on trends in semiconductor and packaging technologies
- Predicting future system performance based on performance analysis of benchmark sets provided by Application Research Group, and feeding back to next-generation application development

Research on System Software and Library

- Drawing roadmap for future system software development in Japan, specially considering data utilization enhancement, integration of AI technology with first-principles simulation, real-time data processing, and assurance of high security

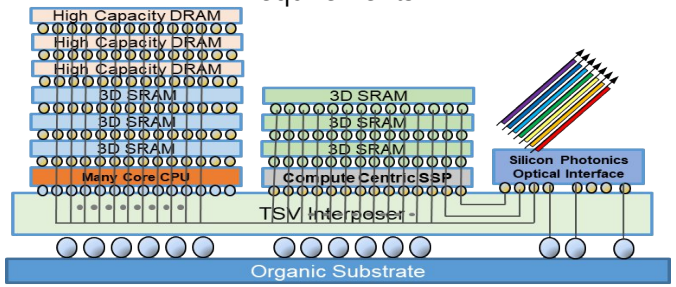
Research on Applications

- Building a broad benchmark set to evaluate multiple architecture choices while considering improvements in algorithms and parameters of application based on the results of architectural evaluations and exploratory "what-if" performance analysis
- Investigating what classes of algorithms are expected to evolve significantly for future systems



Investigation Schedule

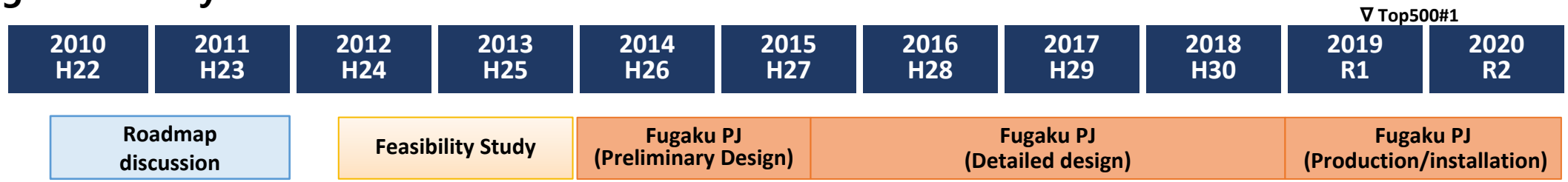
	2022 Q3	2022 Q4	2023 Q1	2023 Q2	2023 Q3	2023 Q4	2024 Q1
Architecture	Explore device/architecture technology			Performance estimation with benchmarks		Architecture study	
System Software	Examine existing SW and its utilization			Identify requirement of SW development		Draw roadmap	
Application	Examine existing apps and benchmark design			Perf. analysis by benchmark evaluation		Study algorithm improvement	



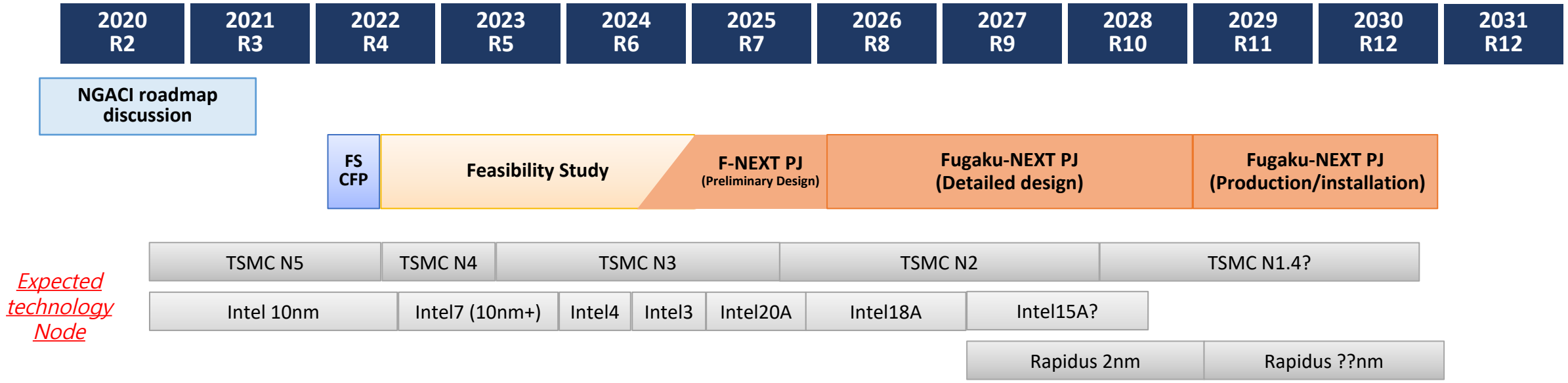
Strawman processing element architecture

Expected Timeline for Fugaku-NEXT Development

Fugaku History



Fugaku-NEXT Expected Schedule



Organization Chart of System Research by RIKEN

System Research Team
(Representative Institution) RIKEN R-CCS
【PI: M. Kondo, AD: S. Matsuoka(R-CCS)】

GL: Group Leader
AD: Advisor
SGL: Sub Group Leader

Architecture Research Group

Architecture Research Group
RIKEN R-CCS
【GL: Sano, Co-GL: Miwa (UEC), AD:Amano (Keio)】

- **Architecture Research sub-G1**
  RIKEN BDR
  【SGL: Taiji (RIKEN BDR】
- **Architecture Research sub-G2**
  Fujitsu Ltd. (Co-I institution)
  【SGL: Shinjo】
- **Architecture Research sub-G3**
  Intel Corporation (Co-I institution)
  【SGL:Yazawa】
- **Architecture Research sub-G4**
  AMD Inc. (Co-I institution)
  【SGL:Yoshida】
- **Architecture Research sub-G5**
  NVIDIA Corporation (Collaborator)
  【SGL:Wells】
- **Architecture Research sub-G6**
  Hewlett Packard Enterprise (Collaborator)
  【SGL:Negishi】
- **Architecture Research sub-G7**
  Arm Ltd. (Collaborator)
  【SGL:Lecomber】

System Software and Library Research Group

System Software and Library Research Group
RIKEN R-CCS
【GL: Sato,Co-GL:Katagiri (Nagoya-U), Sato (TUT), AD: Sato】

Support on Group Management
Nagoya Univ. (Collaborator)
【Delegate: Katagiri】

- **Scheduler / Runtime sub-G**
  Tohoku Univ. (Co-I institution)
  【SGL: Takizawa】
- **IO / Storage / Filesystem sub-G**
  Univ. Tsukuba (Co-I institution)
  【SGL: Tatebe】
  - **Storage Archi Pattern Investigation**
    DDN Japan (Collaborator)
    【Delegate: Hashizume】
- **OS / Virtualization / Cloud sub-G**
  National Institute of Informatics (Collaborator)
  【SGL: Takefusa】
- **HPC Env. Usage Investigation sub-G**
  Osaka Univ. (Co-I institution)
  【SGL: Date】
- **Communication Library sub-G**
  Kyushu Univ. (Co-I institution)
  【SGL: Nanri】
- **Compiler / Programming-model sub-G**
  RIKEN
  【SGL: Tsuji】
- **Numerical Library sub-G**
  RIKEN
  【SGL: Imamura】
- **AI Framework sub-G**
  RIKEN
  【SGL: Mohamed】

Application Research Group

Application Research Group
Hokkaido Univ. (Co-I Institution)
【GL: Iwashita, Co-GL :Takahashi (U. Tsukuba), Fukazawa (Kyoto U.), AD: Nakajima / Tomita (R-CCS)】

Support on Group Management
Kyoto Univ. (Collaborator)
【Delegate: Fukazawa】

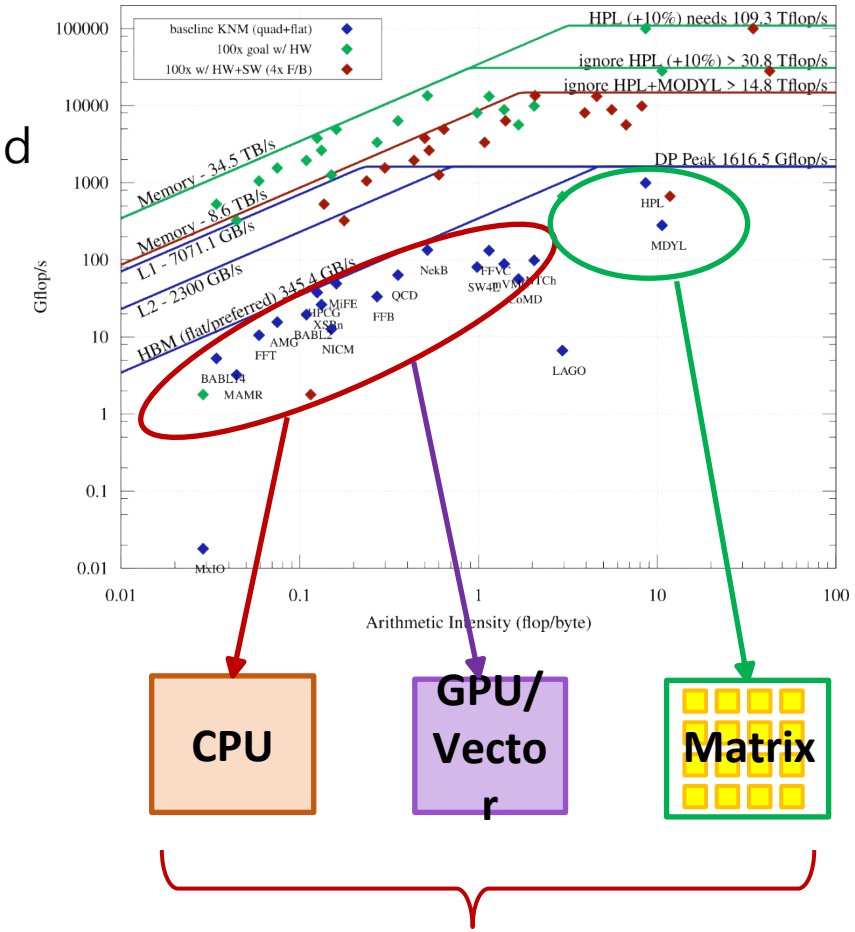
- **Life Science App. Area sub-G**
  Yokohama City Univ. (Co-I institution)
  【SGL: Terayama】
- **Material and Energy App. Area sub-G**
  NIMS (Co-I institution)
  【SGL: Yamaji,Co-SGL:Fukushima(UTokyo) 】
- **Weather/Climate Sci. App. Area sub-G**
  JAMSTEC (Co-I institution)
  【SGL: Kodama】
- **Disaster Prevention App. Area sub-G**
  Univ. Tokyo (Co-I institution)
  【SGL: Fujita】
- **Manufacturing App. Area sub-G**
  RIKEN
  【SGL: Onishi】
  - **Support on Manufacturing Apps**
    JAXA (Collaborator)
    【Delegate: TBA】
- **Fundamental Science App. Area sub-G**
  RIKEN
  【SGL: Aoki】
  - **Support on Space / Planet Sci. Apps**
    NAOJ (Collaborator)
    【Delegate: Takiwaki】
- **Social Science App. Area sub-G**
  RIKEN
  【SGL: Umemoto】
- **Digital-twin / Society5.0 App. Area sub-G**
  Univ. Tokyo (Co-I institution)
  【SGL: Shimokawabe】
  - **Support on Digital-twin Apps**
    Japan Atomic Energy Agency (Collaborator)
    【Delegate: Onodera】
- **Weather Model Perf Analysis sub-G**
  RIKEN
  【SGL: Kodama】
  - **Support on Weather Model Analysis**
    Meteorological Research Institute (Collab)
    【Delegate: Eito】
- **Computational Science Algorithm sub-G**
  Univ. Tsukuba (Co-I institution)
  【SGL: Takahashi】
- **Machine Learning Algorithm sub-G**
  TiTech (Co-I institution)
  【SGL: Yokota】
- **Benchmark Construction sub-G**
  RIKEN
  【SGL: Murai】
- **Performance Modeling sub-G**
  RIKEN
  【SGL: Domke】

Key Research Item for Node Architecture Selection

- **Needs for a power-efficient compute node**
  → Exploration of accelerators
  - Truly useful accelerator for HPC and AI workloads
  - HPC & AI Inference→Memory bound, AI Training→Compute bound
- **Characteristics of current processing element**
  - CPU: high generality, low-latency, low compute density
  - GPU (SP): vector processing, middle compute density
  - Matrix: dedicated for dense algebra, high compute density
    (ex. Tensor core, XMM, SME, AMX, TPU, CGRA, …)
- **What to study in node architecture exploration**
  - What and how to integrate them
  - Effective memory bandwidth + data movement with high programming productivity

Quantitative benchmarking analyses is necessary

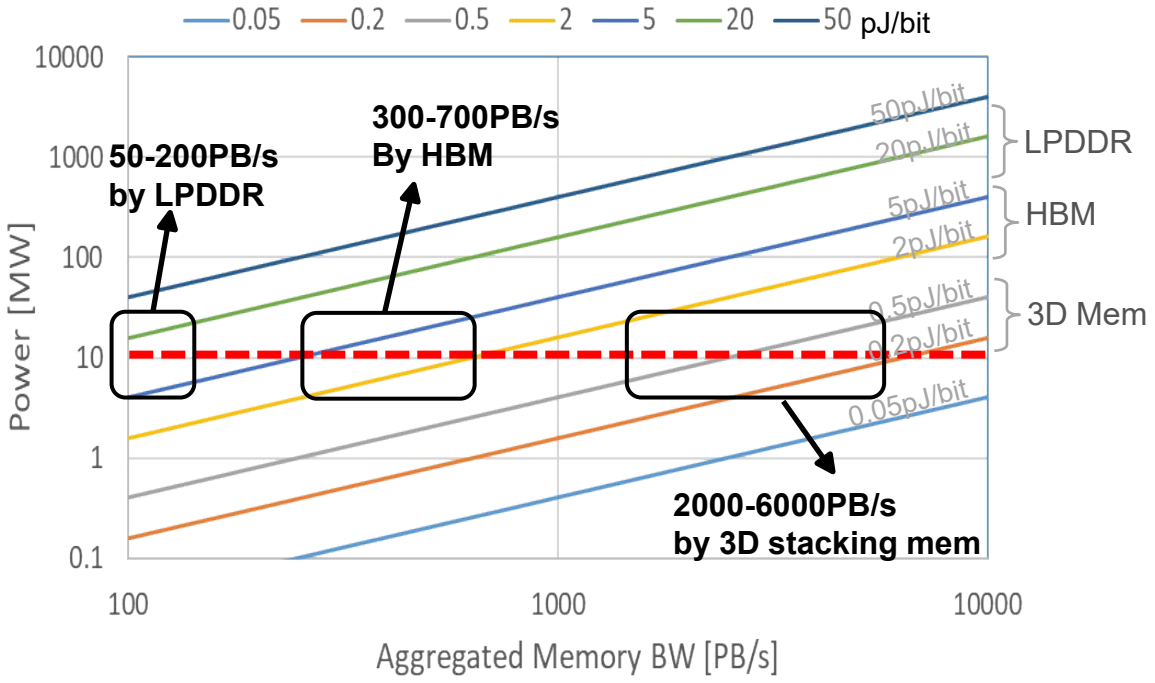
Roofline analysis on A64FX

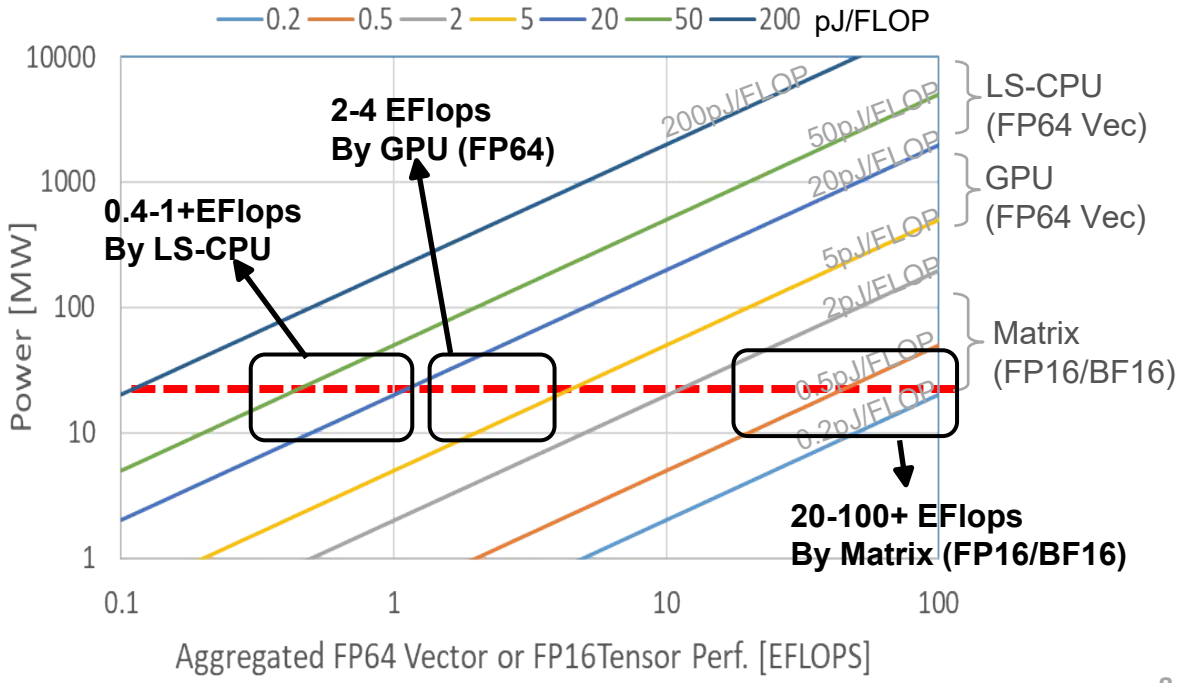


Need to find the optimal balance

Performance Projection in Power Constrained Scenarios

- **Estimated energy per operation on current and future technologies**
  - Based on historical trend obtained by publically available data
  - Not related to any partner vendors' perspective
- **Case for 30MW power budget (10MW for memory and 20MW for compute)**
  - Network is omitted for simplicity but it is very important
  - May not be realistic due to other constraint such as cost and thermal issues



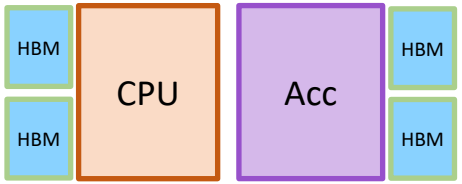
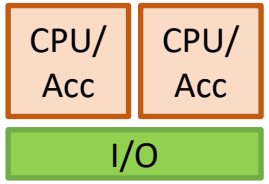
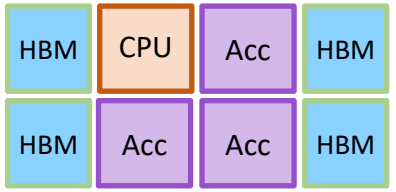
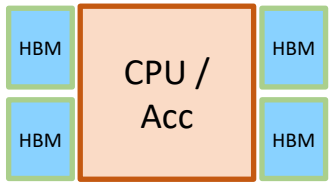
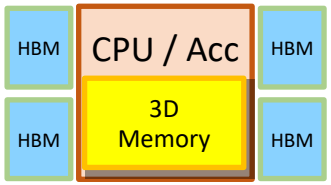
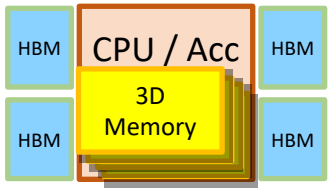
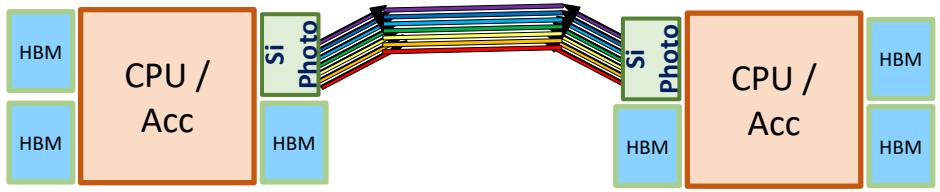


Implementation Approaches for Node Architectures

- **Candidates of packaging technologies**

Technical difficulty: Low ← → High

Power efficiency of data movement: Low ← → High

chip-to-chip connection (chiplets)	**Monolithic die (conventional)** 	**Chiplet-based (becoming main-stream)** 	**More aggressive chiplet-based (Future direction)** 
3D stacking approaches	**2.5D connection (conventional)** 	**3D - Hybrid Bonding (single chip stacked)** 	**3D implementation (multi chips stacked)** 
Optics	**AOC (conventional)** 	**Silicon-Photonics - chip-to-chip optical connection (various technology candidates incl. WDM)** 	

Next Steps in the Feasibility Study Project

- **Selecting architecture/system candidates for a next-generation system**
  - Accelerator, memory technology, photonics technology, and packaging
  - Consider effective accelerator architecture based on quantitative benchmarking analyses
  - Optimizing balance or fusion between HPC and AI performance

- **Creating R&D roadmap for system software**
  - Being strongly conscious of software ecosystem
  - Optimized workflow execution specially for HPC and AI cooperation

- **Application first system design**
  - Design a system target for science breakthrough NOT just for ranking such as Top500
  - Building benchmark framework for fair architectural comparison
  - Blushing up future science roadmap including roadmap on “AI for Science”

- **Collaborating operation technique and new computing-paradigm teams**
  - Data framework, realtimeness, carbon neutrality, ・・・
  - Extending computable areas by HPC-Quantum hybrid platforms