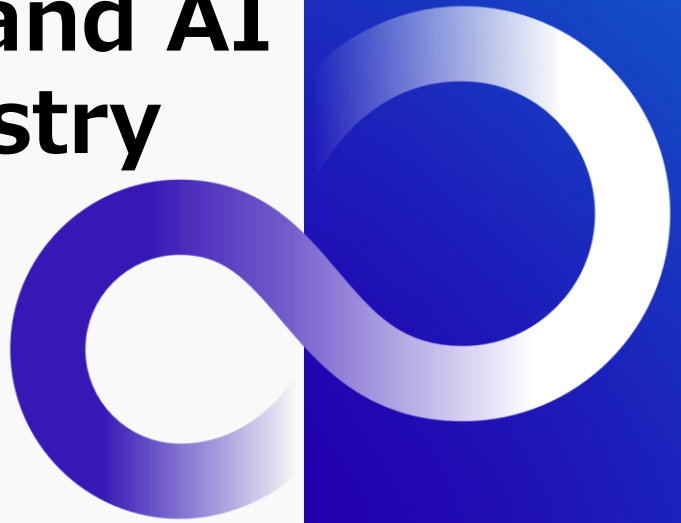
AI Computing Broker and AI based quantum chemistry solution

11 June 2025

Kohta Nakashima
Fujitsu Limited



Fujitsu's Research Strategy

FUJITSU

Sustainable society

Transforming into a regenerative enterprise

Materiality

Fujitsu Uvance
Service solutions
Fujitsu Kozuchi

Business Issues

Planet

People

Prosperity

Creating new value by combining technology areas centered on AI

Converging technologies

Combine digital technology with multidisciplinary knowledge to solve complex and diverse social challenges

Data & Security

Creating new digital economy with advanced technologies to both "Connect" and "Protect"

Computing

World-class computing technologies to innovate and accelerate AI processing

Amazing innovation powered by quantum computing

Network

Secure AI-enabled networks that operate with speed and energy efficiency from the edge to the core

AI

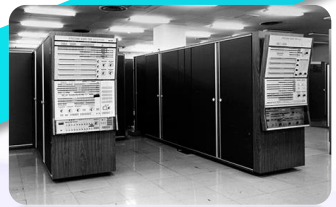
5 Key Technologies

Fujitsu's Computing Technologies

- **Fujitsu continues to provide computing technologies to solve social issues**

1977

FACOM230-75 APU
- ✓ Fujitsu's first supercomputer
- ✓ Just one year after CRAY-1



K computer
- ✓ No.1 in Top500 (2011.6~2011.12)
- ✓ # of CPU: 88,128

Supercomputer Fugaku
- ✓ No.1 in Top500 (2020.6~2021.11)
- ✓ # of CPU: 158,976

2027

FUJITSU-MONAKA
- ✓ Accelerating next-generation DCs
- ✓ High performance, low power

ABCI 1.0/2.0 Miyabi
- ✓ GPU-based supercomputer
- ✓ ABCI: 4,352 V100 + 960 A100
- ✓ Miyabi: 1,120 H100

Computing Technologies and innovation

FUJITSU

- ✓ Computing and AI lead future innovation
- ✓ Fujitsu continues to provide computing technologies

Innovation

Materials

Genomics

Generative AI

Middleware

Computing Workload Broker (CWB)

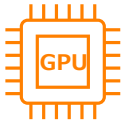
AI Computing Broker

Architecture

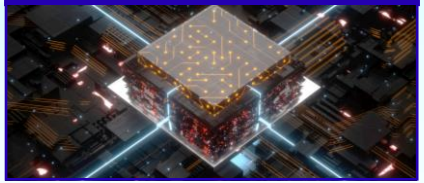
Quantum Computer

Photo Credit: RIKEN

GPU



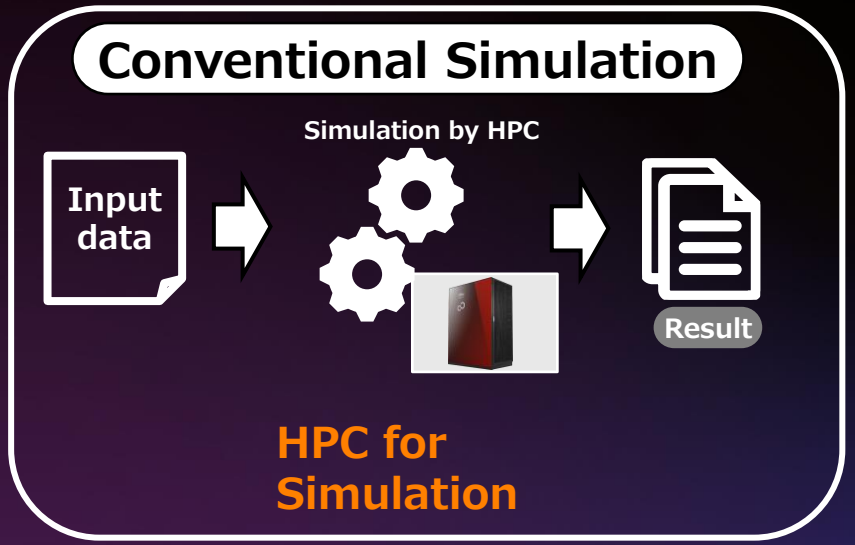
CPU (FUJITSU-MONAKA)

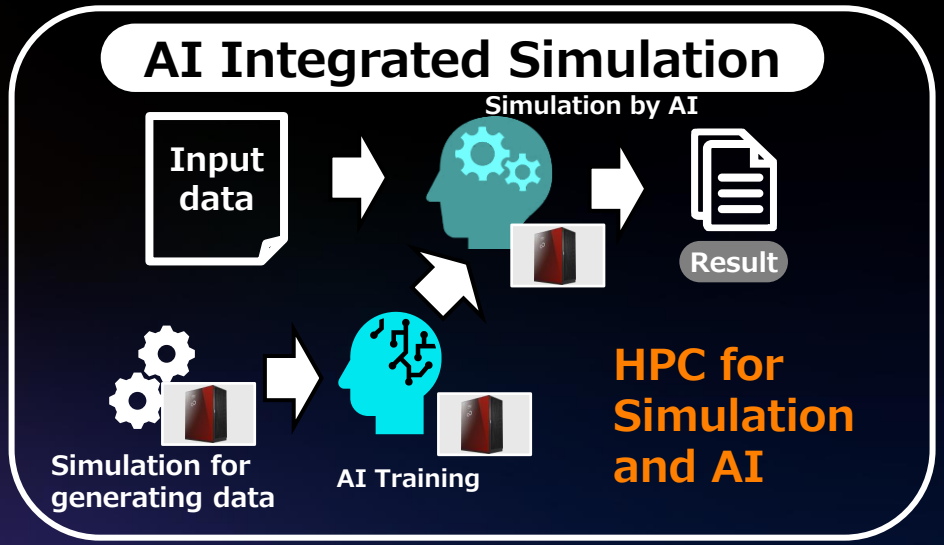


Fusion of HPC & AI: A New Computational Paradigm

FUJITSU

- **HPC systems are expected to accelerate not only traditional simulations but also AI workloads**
  - Integrating AI into simulations enables significant speedup
  - Achieving a seamless connection between HPC and AI has become a critical challenge





Material Design

FUJITSU

Molecular Dynamics Simulation:
For understanding dynamic characteristics of the target material

Classical MD:	Computation time ✓	Accuracy ✗
ab initio MD:	Computation time ✗	Accuracy ✓

→ The best of both worlds: **MD using Neural Network Potentials**

Conventional Material design — **ab initio MD (more than hundreds of years)**

NNP-aided Material design — **NNP-MD (*8days)**

* Hydrated Nafion Example

x 100–10,000 faster w/ High accuracy

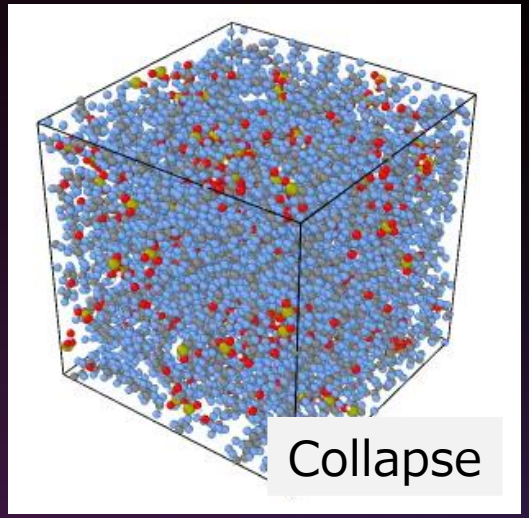
Fujitsu provides a generator of NNP for MD

MD Simulation Case Study

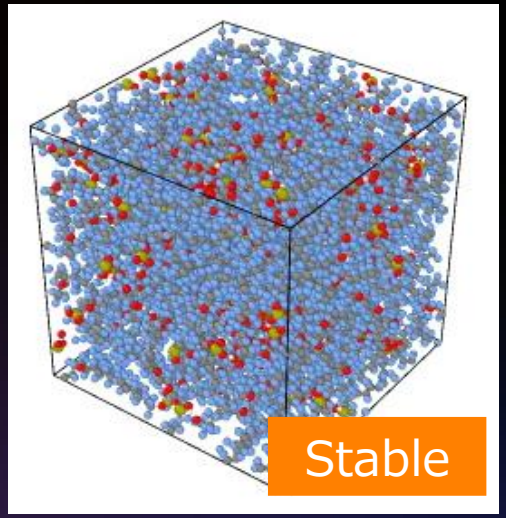
FUJITSU

MD Simulation Breakthrough: 20,000 Atoms, 30 ns (60M steps) in 8 Days, cutting simulation time from hundreds of years !

Conventional



Our technology



Note: Running 30 ns of MD takes 1.5 days

Hydrated Nafion (19,670 atoms)

Achieved sufficient performance for practical use

10k-atoms: Large enough to design entire devices

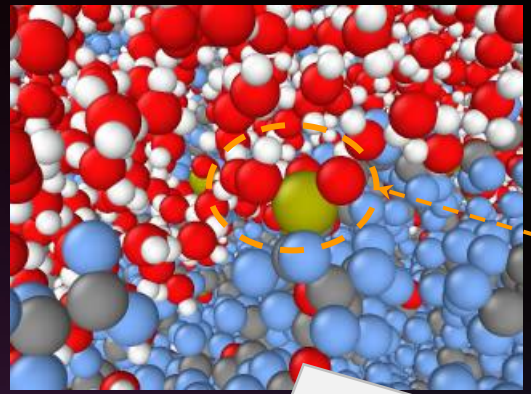
10M-steps: Reproducible chemical reactions

1-week: Acceptable computation time for developers

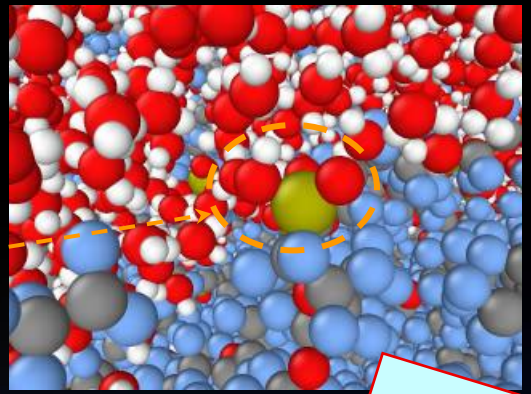
Classical MD vs NNP-MD

FUJITSU

Generated NNP can treat chemical reaction but for classical MD



Sulfone group



H: Hydrogen
C: Carbon
O: Oxygen
F: Fluorine
S: Sulfur

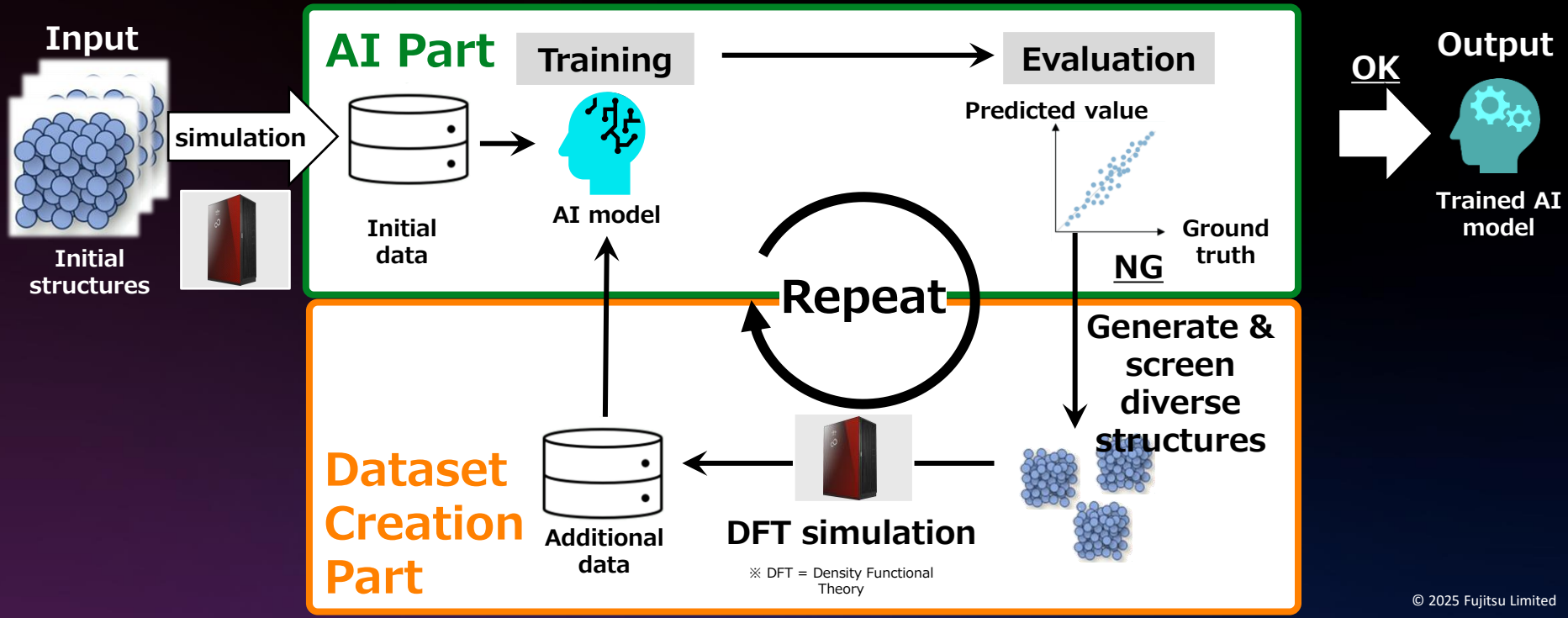
Chemical reaction cannot be simulated

Chemical reaction is simulated
H^+ dissociates from sulfone group and
H^+ combines adjacent H_2O to form H_3O^+

Feature of Our NNP Generator

FUJITSU

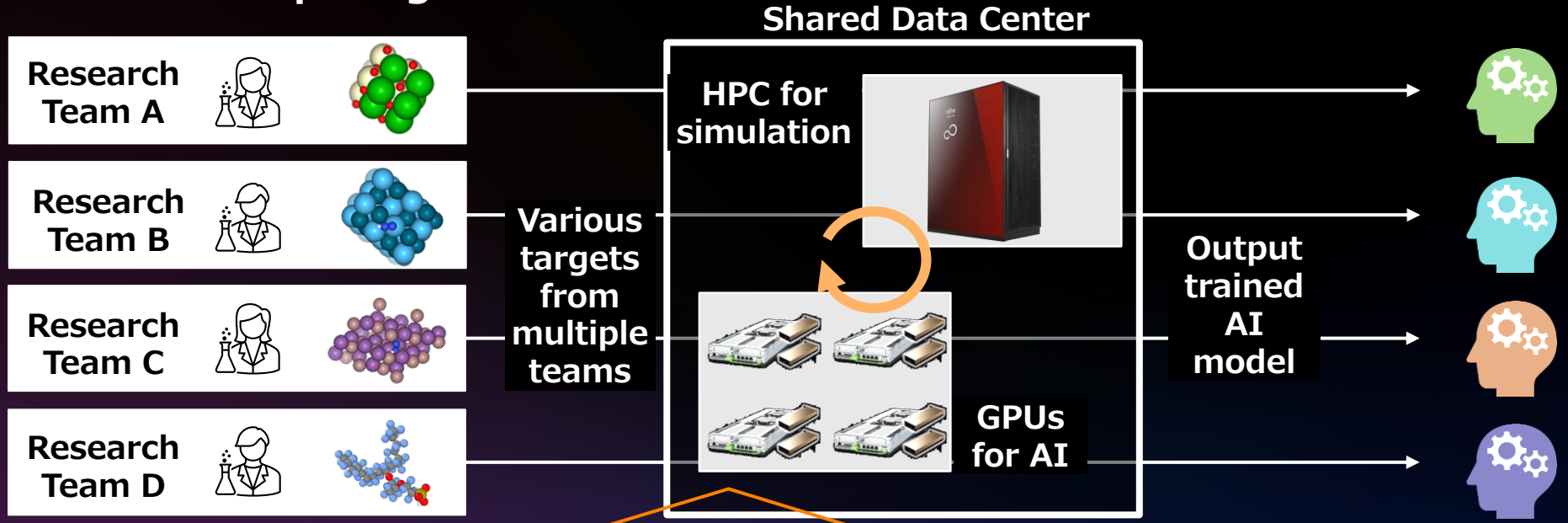
Generate NNP with high accuracy, high speed, and high MD stability with our unique training data generation technology (Active Learning)



Practical Use Case of NNP Generator

FUJITSU

Multiple AI training for each target are running simultaneously on limited computing resource



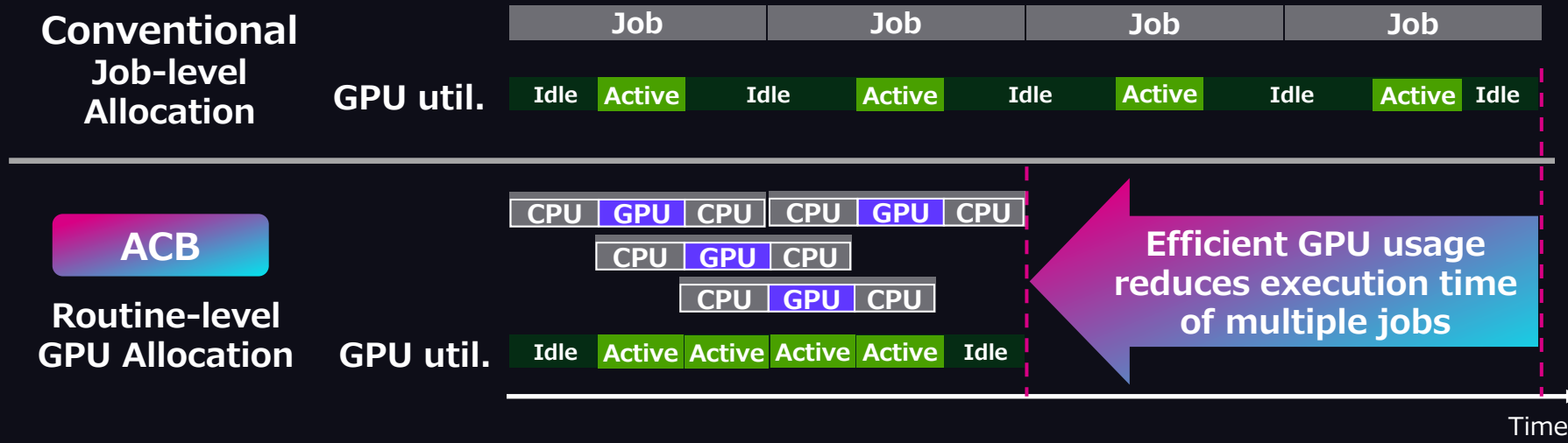
Challenge: **Heavy AI-training jobs monopolize GPU, delaying short jobs significantly**

Feature of AI Computing Broker

FUJITSU

Best-in-class GPU utilization efficiency

- “Routine-level” allocation that detects actual GPU parts of jobs and dynamically allocates GPU accordingly



Feature of AI Computing Broker

FUJITSU

Enabling full GPU memory for each job

- Allocate GPU to only one job at a time (Temporal-sharing)
- Data of other jobs on GPU is automatically swapped to CPU

Conventional: Spatial-sharing

Memory is divided among jobs
Limited to small AI models

Memory

Job A	Job B	Job C

Computing Unit

Sharing among Job A/B/C

ACB Time-sharing

Memory is occupied by each job
Large AI models can run

swap memory

swap memory

Memory	Memory	Memory
Job A	Job B	Job C
Comp. Unit	Comp. Unit	Comp. Unit
Job A	Job B	Job C

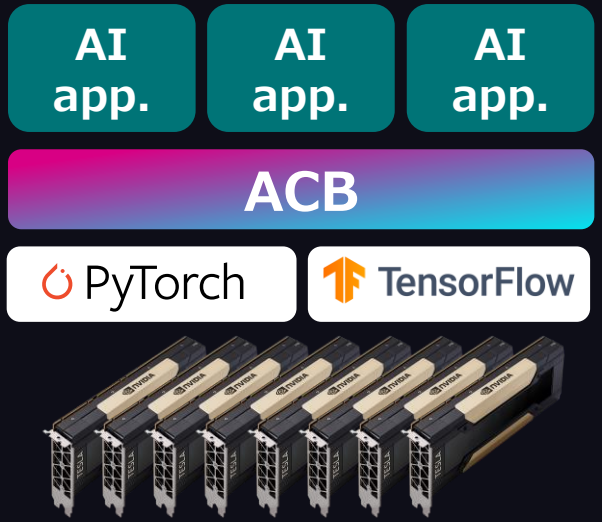
AI computing broker (ACB)

FUJITSU

A middleware to share GPUs among AI apps.

Key Features

- Best-in-class GPU utilization efficiency
- Optimized memory management across various AI applications



"Doubled model training throughput!"

"Deploying multiple jobs beyond physical GPU memory capacity!"

Success stories from ACB users
Scan here for more detail!

FUJITSU

Thank you