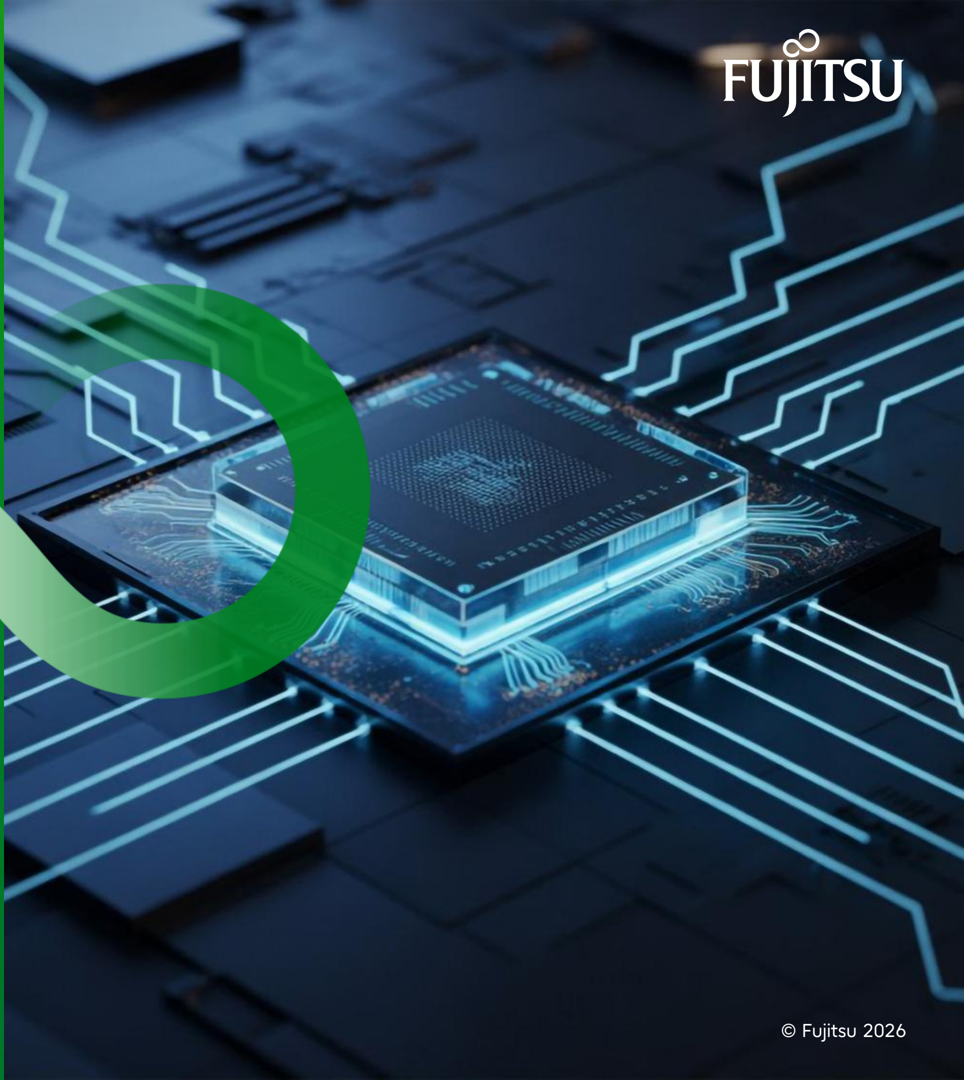
FUJITSU-MONAKA: Next-Generation Processor for Sovereign, Sustainable AI & HPC



* This presentation is based on results obtained from a project subsidized by the New Energy and Industrial Technology Development Organization (NEDO).

Market Context and FUJITSU-MONAKA Value Proposition

FUJITSU

- **Growing AI & HPC demand is reshaping requirements for computing infrastructure.**

Key Trends

Workload diversification Evolving compute demands
- Single large-scale
  → Diversified & high- throughput workloads
- AI infrastructure and HPC-AI convergence

Power constraints
- Rapid growth in compute & power consumption
- Need to maximize performance under power limits

Europe-specific requirement
- Need for control over data, infrastructure, and operations
- Driven by geopolitics, export controls, supply chain risks, and vendor lock-in concerns

FUJITSU-MONAKA Value

Powering Diverse AI/HPC Workloads
- Up to 2× higher performance in AI workloads
- High-throughput CPU architecture
- Open AI/HPC ecosystem

Over 50% TCO Reduction at the Same Performance
- High performance per watt

Sovereign & Trusted computing
- Control over data and infrastructure
- Secure execution and data protection in use

CPUs as the Foundation for Evolving AI/HPC Systems

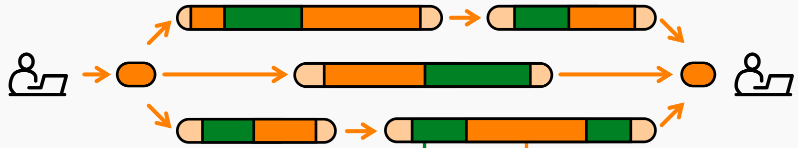
FUJITSU

CPU continues to expand its role in AI/HPC systems alongside growing GPU adoption

- AI: Agentic AI drives greater CPU usage for orchestration and peripheral processing
- HPC: CPUs remain essential for high-throughput and high-precision workloads

AI Workloads

Multi-Agent collaborative execution



AI workload CPU GPU

LLM, Speech, Video, Vision, etc.

General-purpose workloads CPU CPU

Orchestration, RAG, DB, Code execution, etc.

HPC Workloads

Precision Small-Job Processing

CPU – Small jobs x many

Startup ≈ compute → CPU

startup compute

CPU – High precision

FP64

3.14159265

Large-Scale Parallel Processing

GPU – Large jobs x few

Startup << compute → GPU

compute

startup compute

GPU – Speed over precision

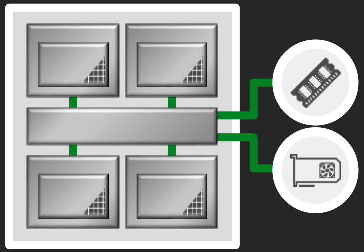
FP16/FP32

3.1415

FUJITSU-MONAKA Processor Overview

FUJITSU

FUJITSU-MONAKA



- Armv9-A Architecture
- 3D chiplet
  - Core die 2nm
  - SRAM die/IO die 5nm
- Ultra low voltage for energy-efficiency
- DDR5 12 channels
- Liquid / Air-cooling
- Arm SVE2-256bit for AI and HPC
- 144 cores x 2 sockets (288 cores per node)
- Confidential Compute Architecture for security
- PCI Express 6.0 (CXL3.0)

subject to change without notice

Selected PoC starts from 2026 Summer.
Product shipments begin in 2027.

FUJITSU-MONAKA

High-Performance and Energy-Efficient CPU for a Carbon-Neutral Digital Society

High-Performance
Achieving high-speed computing centered on AI workloads (2×competitors CPUs).

Power-Efficiency
Reducing CO_2 emissions and electricity costs (2×competitor CPUs).

Safety & Security
Leveraging mainframe RAS technologies.

Ease of Use
Leveraging Armv9 Software ecosystem.

FUJITSU-MONAKA Technology Features

Fujitsu's Made-in-Japan technology powers performance and efficiency

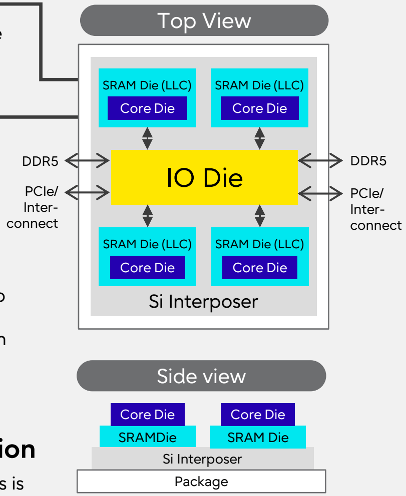
High-Performance per Cost: Optimized Multi-Die Architecture

Core Die: 2nm process
- Die integrating multiple compute cores

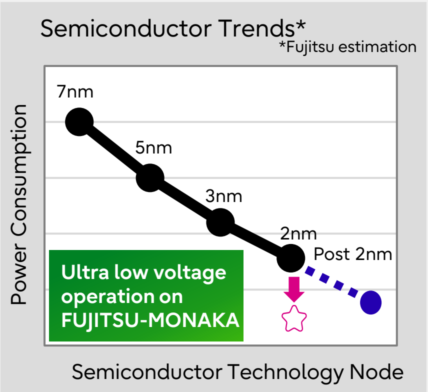
SRAM+IO Die: 5nm process
- Die integrating last-level cache (LLC)

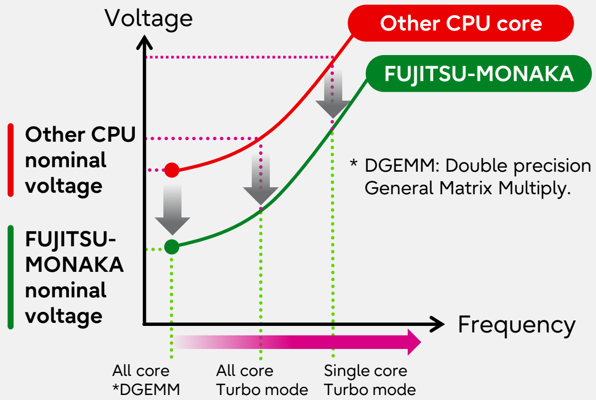
3D stacking
- Core die stacked on top of SRAM die
- Tightly coupled through TSV

Heterogeneous process integration
- The costly 2 nm process is applied only to the core die
- Limiting the cutting-edge process area to less than 30%



For High-Performance per Watt: Ultra Low Voltage Technology





Unique Ultra Low Voltage libraries for Core Die
- FUJITSU SRAM cells
- FUJITSU custom logic cells

➔ Achieves lower voltage operation **reduce core operating voltage.**

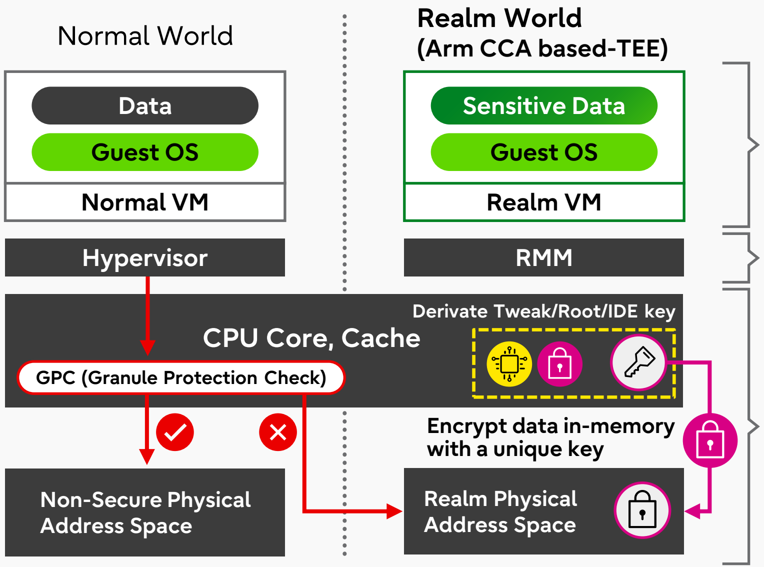
Advantage of Ultra Low Voltage design
- FUJITSU-MONAKA micro-architecture enables lower operating voltage for core at any operating frequency. (i.e. any workload)
- Since power ∝ voltage^2, FUJITSU-MONAKA core can drastically reduce power consumption, realizing **superior power efficiency.**

Arm CCA with FUJITSU-MONAKA

*CCA = Confidential Compute Architecture

Fujitsu delivers a secure and sovereign CCA platform to meet Europe's requirements

- **Security**: Protects data-in-use within VMs, using Arm Confidential Compute Architecture (CCA)
- **Integrity**: Prevents environment spoofing through hardware-based remote attestation
- **Compatibility**: VM-based CCA technology enables seamless migration of existing environment



Sovereignty: Fujitsu delivers CCA Full-Stack platform

Software Stack

Confidential VM / Container	Guest Kernel, CoCo
Remote Attestation	Veraison, Trustee
Host OS/ Hypervisor	Host Kernel, QEMU, EDK2, libvirt, OpenStack, KubeVirt

Firmware

TF-A: Establishes the trusted boot chain for RMM and Secure Monitor
TF-RMM: Manages Realm lifecycle, memory, and attestation

Hardware Root of Trust (FUJITSU-MONAKA)

- **Security ensured by: - Sub processor HW**
- **Tweak/Root/IDE key Derivation core/cache**
- **Encrypt data for out of chip**

Building the OSS Ecosystem for FUJITSU-MONAKA

FUJITSU

Advancing AI/HPC Software Readiness Through Collaboration with OSS Communities

- **Delivering High-Performance Execution Environment for Diverse AI/HPC Workloads**
- **Ensuring Performance and Continuous Validation of Major OSS Applications and ISVs**

Submitted 200+ PRs to 30+ OSS communities, doubling contribution in a year for FUJITSU-MONAKA

AI Frameworks/Libraries

ML DL

OpenBLAS | PyTorch | oneDNN

Improved performance by **1.2-3x**
through introducing SVE, Int8, thread throttling and others

LLM

vLLM | llama.cpp | OpenVINO

Improved inference performance by **2-3x**
through supporting quantization and enhancing framework interoperability

DP

Milvus | PostgreSQL

Improved similarity search by **1.3x** and encoding/decoding by **17x**
through using SVE and OpenMP

Application Software

OSS

Achieved **1.9-2.1x** performance over projected 2027 CPUs
with OSS applications: OpenFOAM, FrontISTR, LAMMPS

ISVs

Expanding validation activities with major ISVs

System Software

Collaboration with Arm, Red Hat, SUSE

Linux | PAPI/libpfm

Review&Test: 90+
Working on community development for FUJITSU-MONAKA and Armv9.

GCC | LLVM

Submitted bug report: 350+
Working on performance improvements and Fortran quality enhancements.

FUJITSU-MONAKA Accelerates Industrial AI

Achieve cost-effective and high-accuracy CAE design in the real world

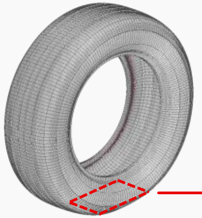
- Rapid, low-cost design evaluation with surrogate models
- Improve design accuracy in the early stages of product development

AI model for structural analysis

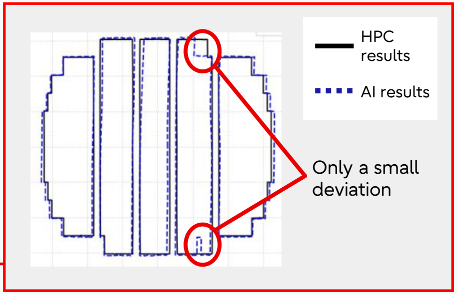
- Achieved **9x** speed-up over HPC simulation
- Evaluation Accuracy: **87.7%**
- PoC in collaboration with DUNLOP

Reference

DUNLOP × FUJITSU



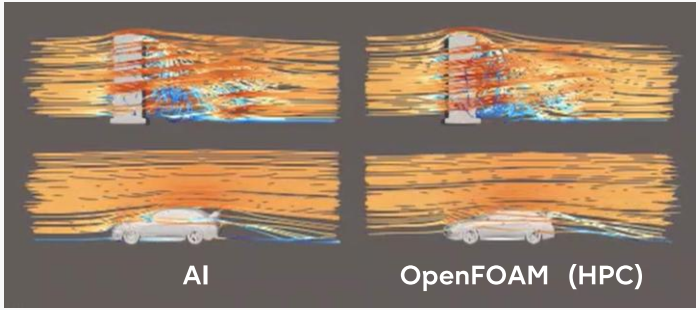
Tire Deformation



Deformation at the Tire Contact Patch

AI model for fluid dynamics

- Achieved **5x** speed-up over HPC simulation
- Evaluation Accuracy: **91.1%**



Building

Car

AI Evolution and FUJITSU-MONAKA Technology Roadmap

FUJITSU

AI workloads evolve from HPC-centric to AI-native computing, and FUJITSU-MONAKA addresses this through multi-generation innovation

Development Roadmap

	2020	2027	2029	2031
	HPC, AI	Data Center, HPC, Telecommunication	Next	
	Fujitsu A64FX / 7nm core	FUJITSU-MONAKA / 2nm core	FUJITSU-MONAKA-X / 1.4nm core	FUJITSU-MONAKA-XX / Further Advanced Node
Hardware	• Adopted in Fugaku (Top-tier supercomputer) • World's first implementation of Arm SVE	• High-speed data processing infrastructure • High performance × power efficiency	• CPU with integrated NPU • Selected for FugakuNEXT	• CPU-NPU fusion • Cutting-edge technologies
Compute Architecture	SVE	SVE2 (2027–2031)		
			SME2 (2029–2031)	
			NPU (2029–2031)	
		PCIeGen6 GPU	NVLink Fusion (2029–2031)	
Software & Ecosystem	Fugaku HPC software stack	FUJITSU-MONAKA software stack	• NPU software stack • AI/HPC optimized stack	

Thank you