

# Next Arm Processor FUJITSU-MONAKA and Its Technologies



# Fujitsu Processor Development



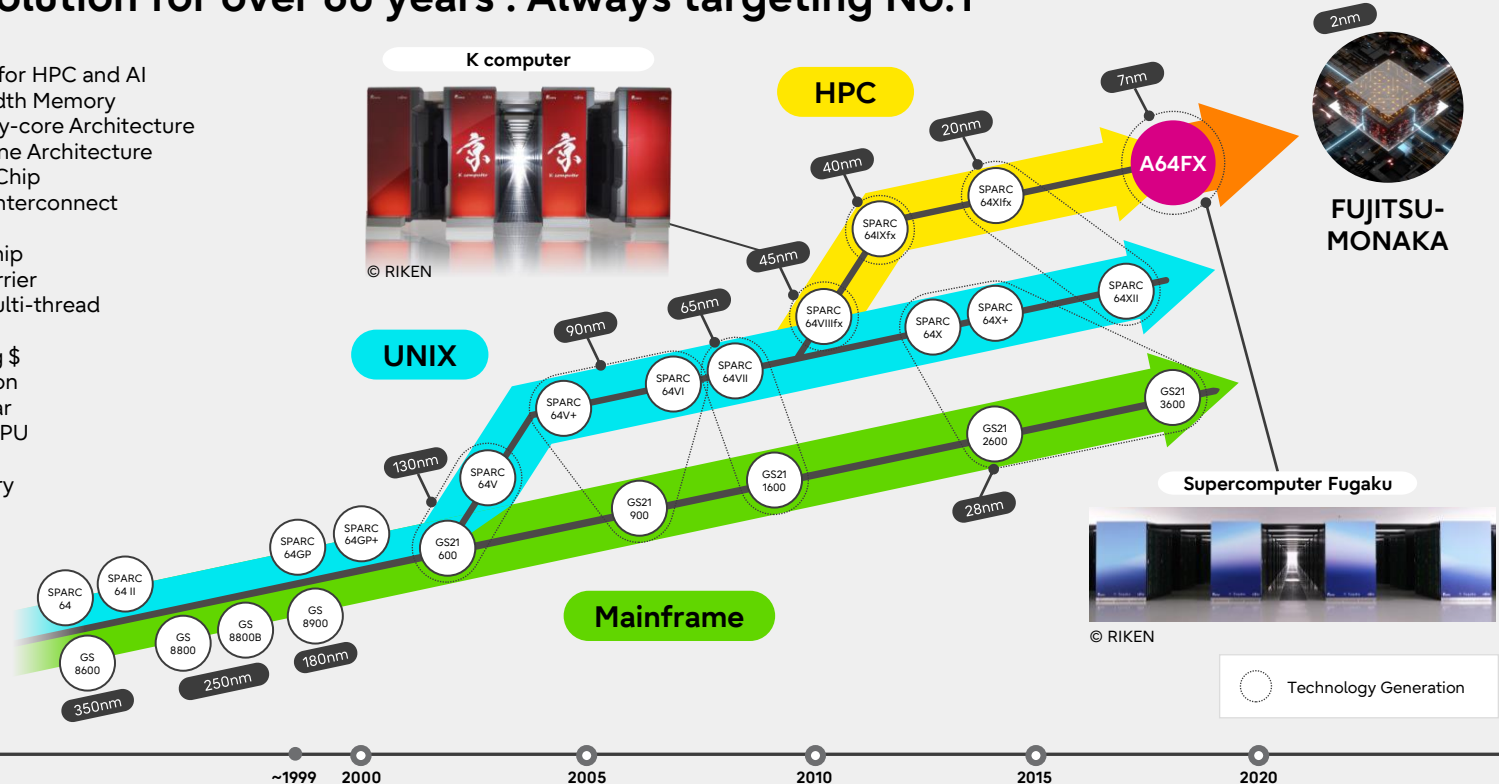
- Persistent Evolution for over 60 years : Always targeting No.1

**Performance**

- 512-bit SIMD for HPC and AI
- High Bandwidth Memory
- Scalable Many-core Architecture
- Virtual Machine Architecture
- Software on Chip
- High-speed Interconnect
- HPC-ACE
- System on Chip
- Hardware Barrier
- Multi-core Multi-thread
- L2\$ on Die
- Non-blocking \$
- O-O-Execution
- SuperO-Scalar
- Single-chip CPU
- Store Ahead
- Branch History
- Prefetch

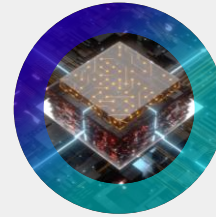
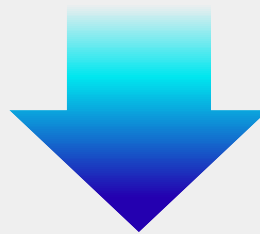
**Reliability**

- \$ ECC
- Register/ALU Parity
- Instruction Retry
- \$ Dynamic Degradation
- Error Checkers/History



# The Next Stage as World's Top Processor

- Creating a new era of computing power is mandatory for the future society with massive data generation and processing
- Ever-increasing power in datacenters is critical, and the power efficiency in CPU (consists of 60%) would be the vital factor for a sustainable future
- Fujitsu shall utilize its Supercomputer success and technology for the solution



FUJITSU-MONAKA

- **Developing the new power efficient CPU "FUJITSU-MONAKA" for datacenters, which will be shipped in 2027**
- **Targeted for wide range of usage in the datacenter including AI and HPC, and contribute to the realization of carbon-neutral society**

# Fujitsu Arm Processor "FUJITSU-MONAKA"



Fujitsu microarchitecture

3D many-core architecture

Confidential Computing



High-performance

- Cloud native 3D many-core design by Fujitsu-proven microarchitecture
- High memory bandwidths



Energy Efficient

- Leading-edge process technology
- Ultra low voltage operation



High Reliability

- Multiple VM Confidential Computing
- Mainframe class RAS for stable operation

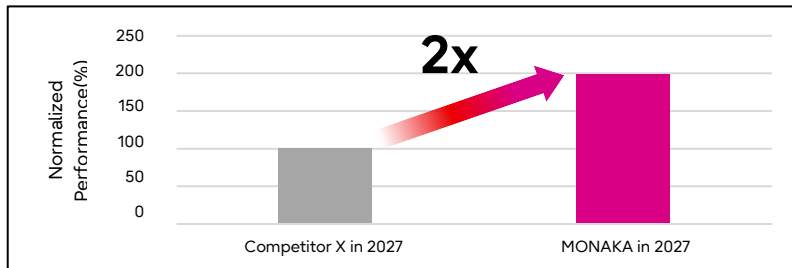


Easy to Use

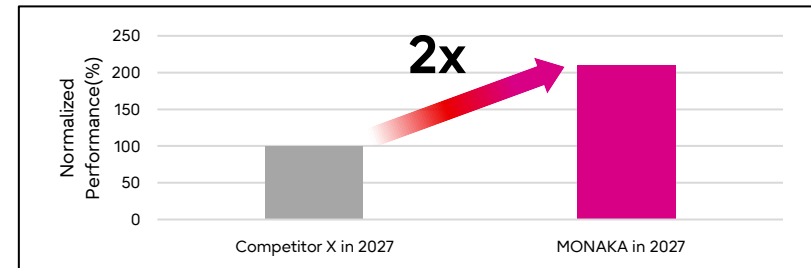
- Open & de-facto standard software stacks
- Fujitsu compiler technology
- Air-Cooling for easy deployment



## Application Performance



## Performance per Watt

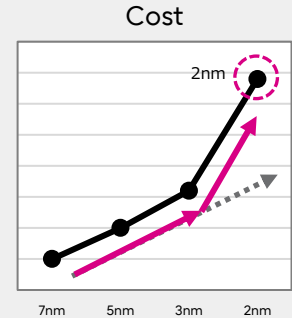
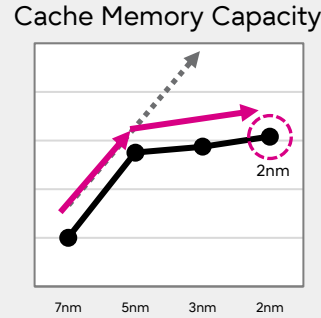
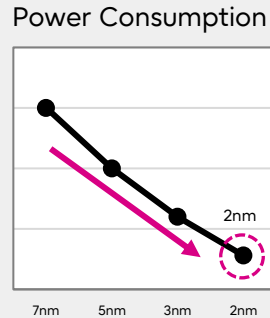
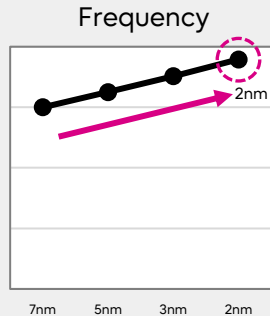


# Semiconductor Trends -High Performance-

- (Pros.)  
Continuing improvements in semiconductor performance and power consumption, due to advancements in transistor structures
- (Cons.)  
Facing challenges with the slowdown in increasing cache memory capacity and the drastic increase in cost

	Transistor Type		
	Planar	FinFET	GAA (Gate-All-Around)
Technology node	~20nm (K computer : 45nm)	16nm ~ 3nm (Fugaku : 7nm)	2nm~ (MONAKA : 2nm)
Semiconductors Structures			

## Semiconductor Trends\*



\*Fujitsu estimation

Architectural innovation is required to meet demand of performance, power and cost

# 3D Microarchitecture -High Performance-

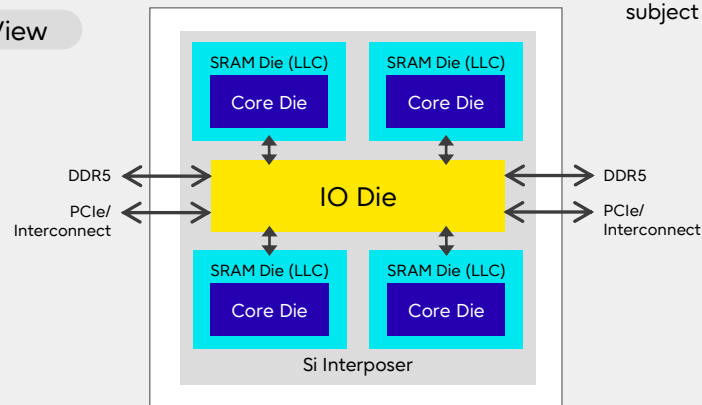
## ● FUJITSU-MONAKA adopts the innovative 3D many-core architecture

- 2nm is used only for core die (top die), achieving high performance and low power consumption
- All the last level cache are in 5nm SRAM die (bottom die), tightly coupled with core die through TSV
- 3D many-core architecture realizes more cores, low latency and high throughput
- 2nm area is less than 30% of total die size, contributing to cost-efficiency

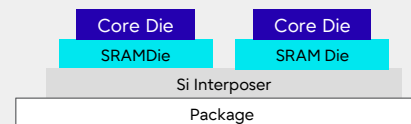
### FUJITSU-MONAKA Specification

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

Top View



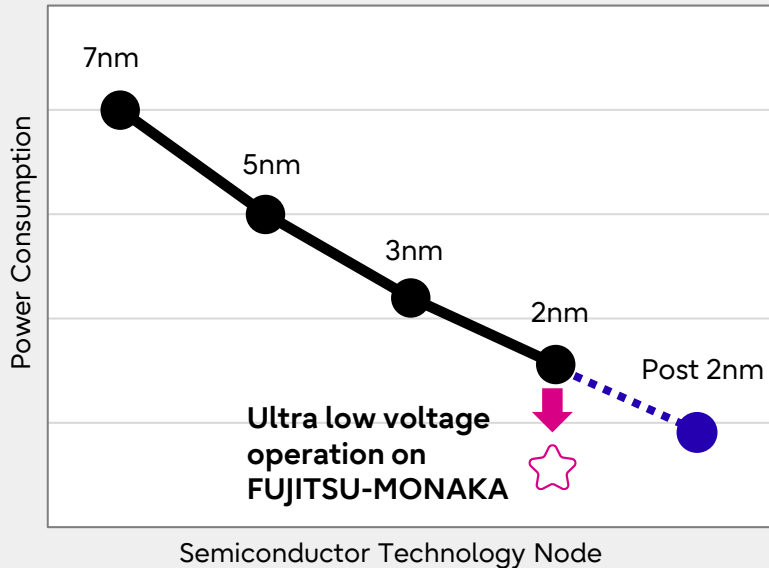
Side view



# Ultra Low Voltage Technology -Energy Efficient-

- FUJITSU-MONAKA's ultra low voltage operation technology enables energy saving comparable to one generation ahead of 2nm

## Trend of Semiconductor Power\*



- Reducing power consumption by lowering voltage of the CPU

$$P \propto C V^2 f$$

C : Capacity

V : Voltage

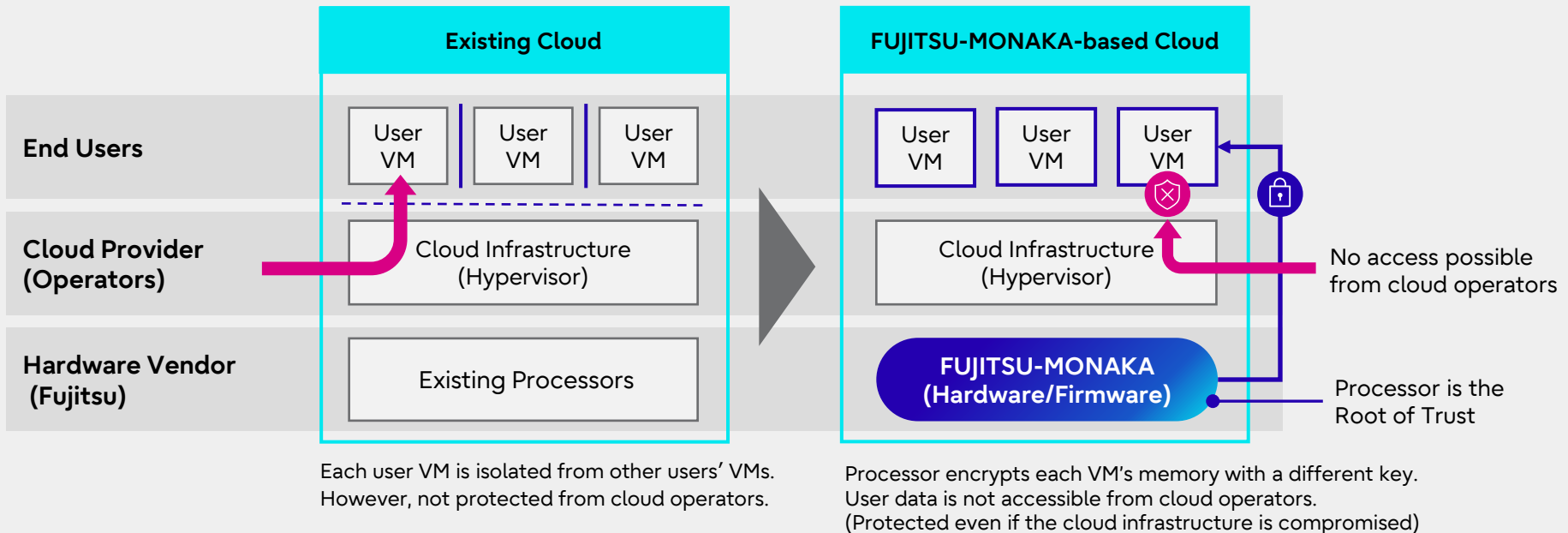
f : Clock Frequency

- Fujitsu develops custom circuits including SRAM by using our proprietary CAD, which enables stable operation at an ultra low voltage

# Security Enhancements -High Reliability-

## ● Confidential Computing

- Protect end-user data in memory by encrypting every VM with a different key generated by the processor hardware and firmware
- Expected to be an essential technology in cloud, edge and HPC environments which deals with sensitive data





# Comparison between A64FX and FUJITSU-MONAKA

A64FX	FUJITSU-MONAKA
<p>Armv8-A Architecture - SVE for HPC and AI</p>	<p>Armv9-A Architecture - SVE2 enhanced for HPC and AI - Confidential Computing</p>
<p>48 cores x 1 socket</p>	<p>144 cores x 2 sockets</p>
<p>Low voltage</p>	<p>Ultra low voltage</p>
<p>2.5D - CPU 7nm - HBM2</p>	<p>3D chiplet - Core die 2nm - SRAM die/IO die 5nm</p>
<p>HBM2 4 channels</p>	<p>DDR5 12 channels</p>
<p>PCI Express 3.0 Tofu Interconnect</p>	<p>PCI Express 6.0 (CXL3.0)</p>
<p>Air cooling and water cooling</p>	<p>Air cooling</p>

# Software Ecosystem - Easy to Use -

- **Supports industry standard software**

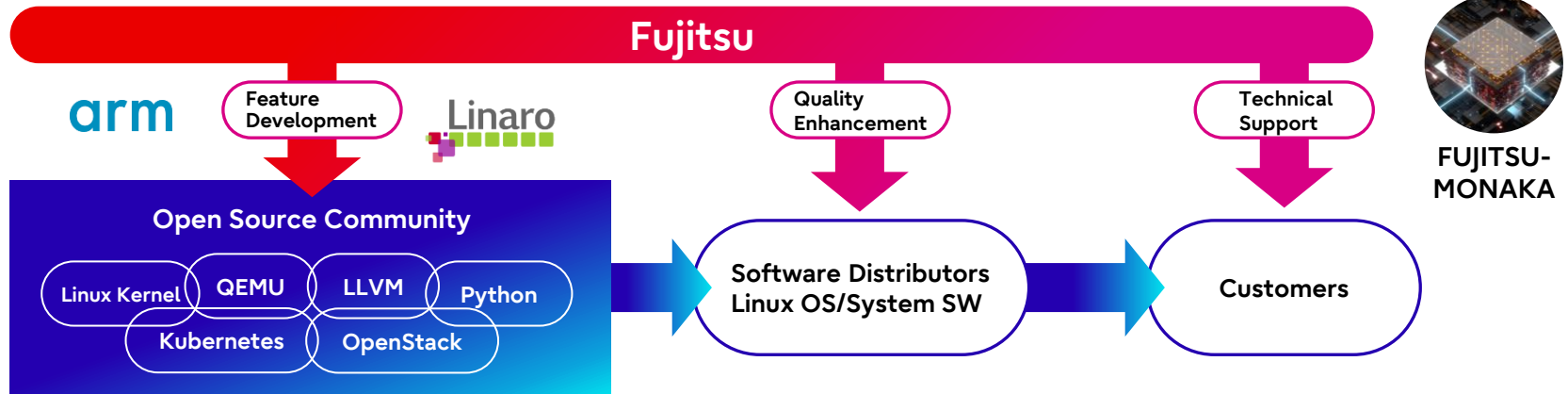
- Standard Linux OS support and system architecture

- Continue and expand OSS development activities for FUJITSU-MONAKA
  - OSS development achievements for Fugaku/A64FX: GCC, glibc, live-patch, papi, etc
- Comply with standard system architecture (Arm System Ready) and support major distributions

- Arm software ecosystem

- Working on the standard tools (Python/Java/LLVM) to provide higher performance on FUJITSU-MONAKA.

- ▶ Enabling smooth transition of customer assets and continuously enhancing performance



# Conclusion

- Fujitsu develops high performance and energy-efficient processor called FUJITSU-MONAKA using our own microarchitecture and innovative 3D many-core architecture
- We continue and expand software development with communities and partners for easy-to-use
- This processor will meet future computing demand of performance, power, reliability and usability for wide range of usage in the datacenter including AI and HPC
- We will contribute to the realization of carbon-neutral society by our computing technologies and collaboration with users and partners

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

**Thank you**

