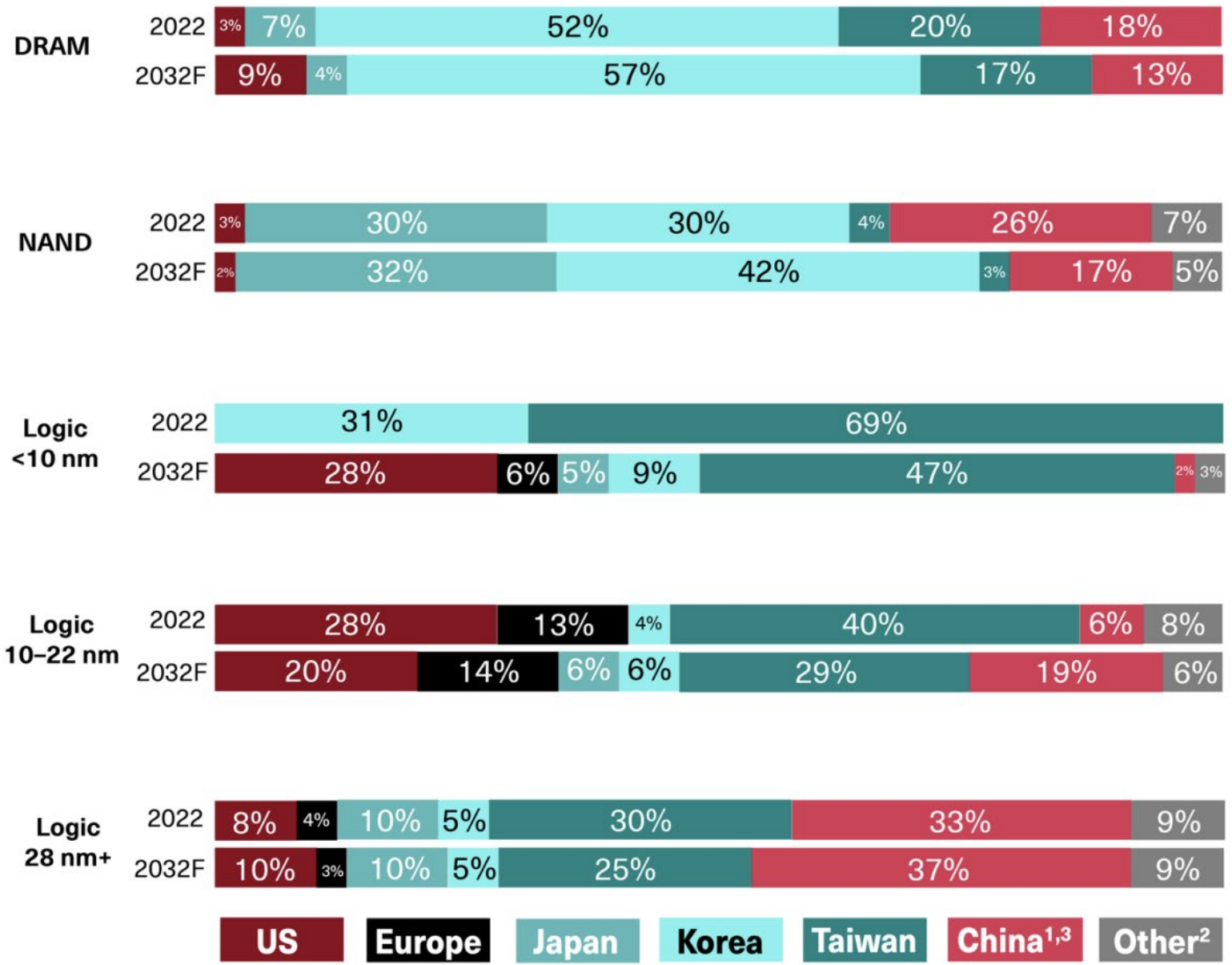
Asia Memory Markets
Trends, Threats and Opportunities:

Sovereign Data Center & Chip Mandates Reshaping Global Tech

Nilesh Shah
ZeroPoint Technologies

Global Wafer Fab capacity: 2022 - 2032

- Korea, Taiwan, China continue DRAM dominance
- China NAND share rise
- Taiwan dominance leading edge nodes
- China dominates older nodes



FMS the Future of Memory and Storage

https://cetas.turing.ac.uk/publications/chinas-quest-semiconductor-self-sufficiency

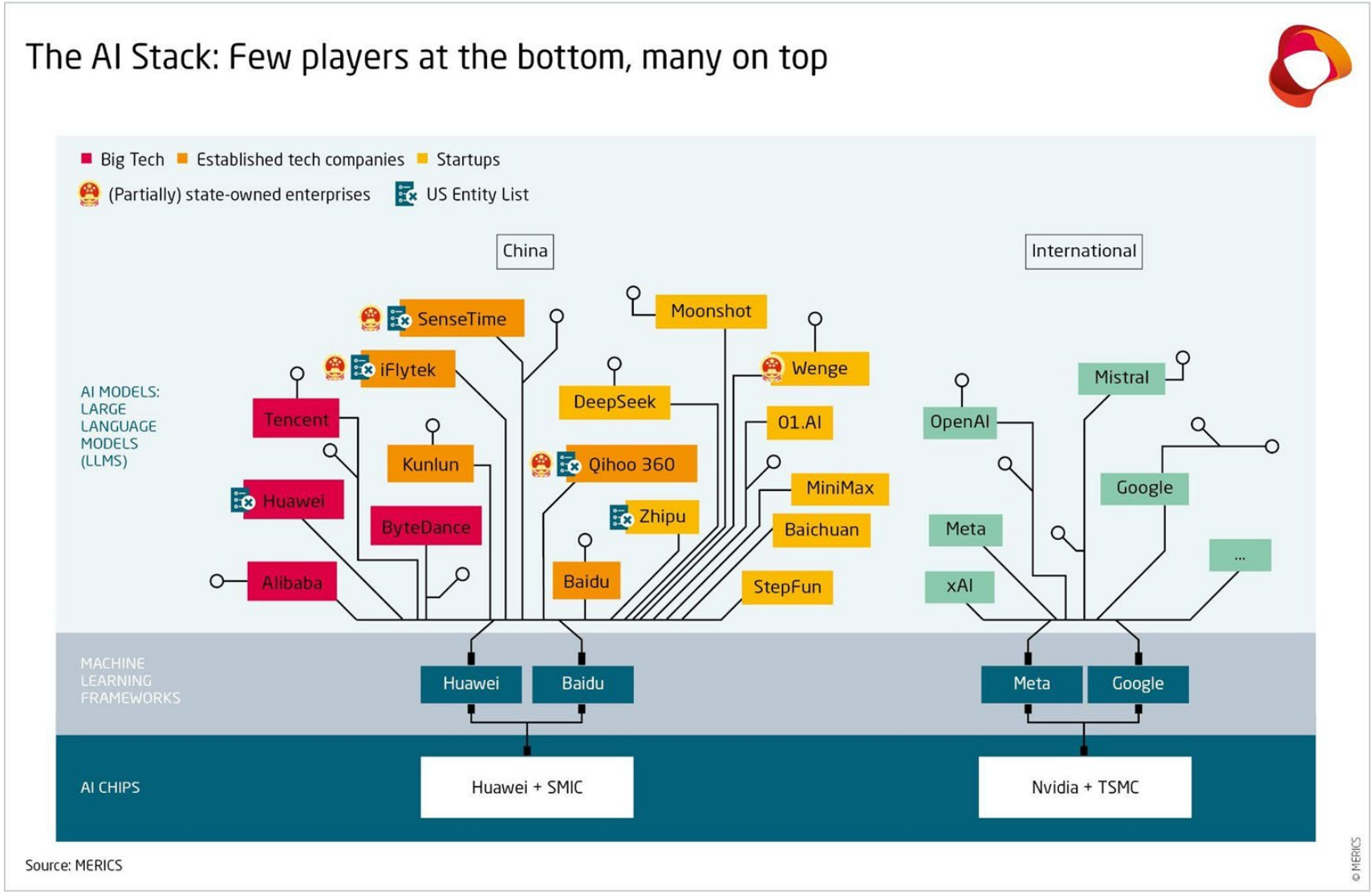
Made in China 2025 policy

Big Fund III (May 2024)

- **$47.5B Committed**: Exceeds U.S. CHIPS Act funding.
- **15-Year Horizon**: Long-term focus vs. previous 5-year cycles.
- **Centralized Strategy**: Directed capital for national semiconductor goals.

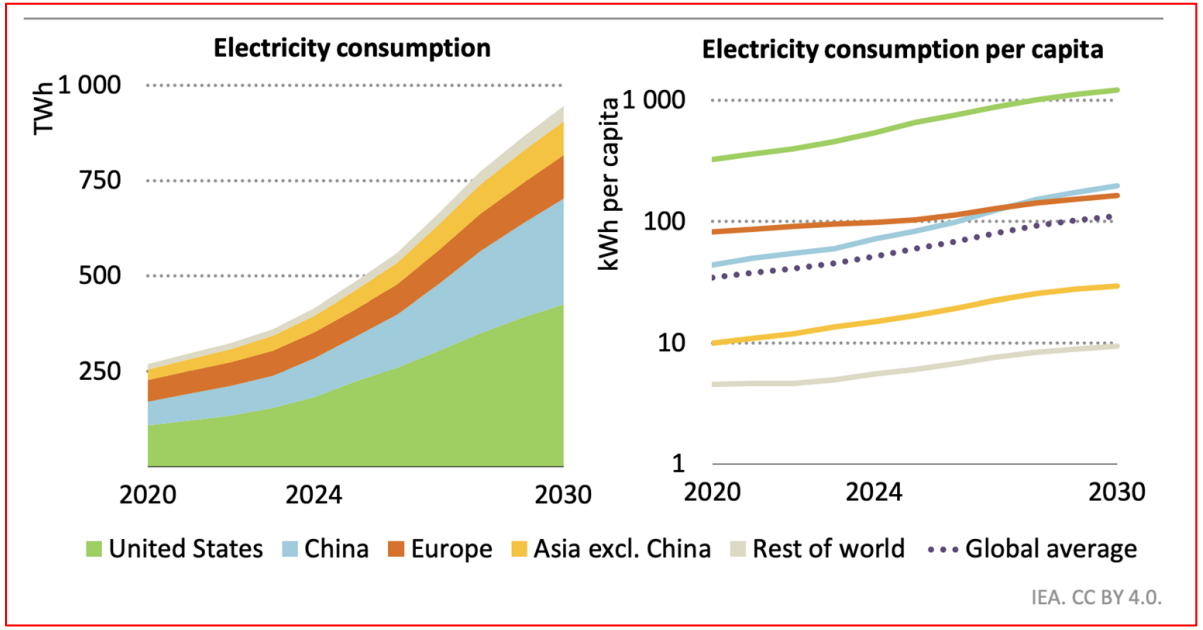
https://cetas.turing.ac.uk/publications/chinas-quest-semiconductor-self-sufficiency

AI Stack: China v/s World



https://merics.org/en/report/chinas-drive-toward-self-reliance-artificial-intelligence-chips-large-language-models

Sovereign Compute: Data Center Demand, Asian Chip Manufacturers



	China	South Korea
Inference	• **Moore Threads** (CUDA-alike GPU) • **Loongson** (CPU, edge inference)	• **FuriosaAI** (edge inference) • **Rebellions.ai** (datacenter inference)•
Training	• **Biren** (HBM-powered BR100 for LLM training)• HBM2 from CXMT (early)	NA
Memory/ Storage	• **CXMT** – DDR5, HBM2 (in development) • **YMTC** – NAND (Xtacking), 13% global share	• **SK hynix** – HBM3E, LPDDR5, enterprise SSDs • **Samsung** – HBM3, ZNS SSDs, CXL memory prototypes

US + China: 80% AI DC power consumption by 2030

Sovereign chip mandate

FMS the Future of Memory and Storage

source: https://www.iea.org/reports/energy-and-ai

China: Memory (CXMT) & Storage (YMTC)

Sovereign mandates accelerating Memory/Storage self sufficiency

Category	CXMT (ChangXin Memory)	YMTC (Yangtze Memory)
Product Focus	DRAM (DDR4 → DDR5), HBM2/3/3E	3D NAND (294-layer, XStacking)
Production Scale	**280K–300K wafers/month projected by end-2025**	**~250K WOPM (wafer-on-product); uses ~500K raw wafers/month**
Technology Node	16nm (1z); ~3–4 years behind industry leaders	**Within ~1–2 years of leaders (Samsung, Micron)**
HBM Status	HBM2 in production (late 2024); HBM3 planned for 2026	Not active in HBM; focused on advanced NAND
Market Share	0% (2020) → 5% (2023) → 10–12% projected (2025)	13% global NAND share (2024–2025)
Strategic Role	**Potential DRAM supplier for sovereign compute (e.g., Biren)**	**Silent expansion into high-density AI storage; NAND alternative to Western flash, HBF leadership**
Sanctions Exposure	Uses 16nm tech (just above U.S. export threshold)	On U.S. Entity List; vulnerable to export tightening

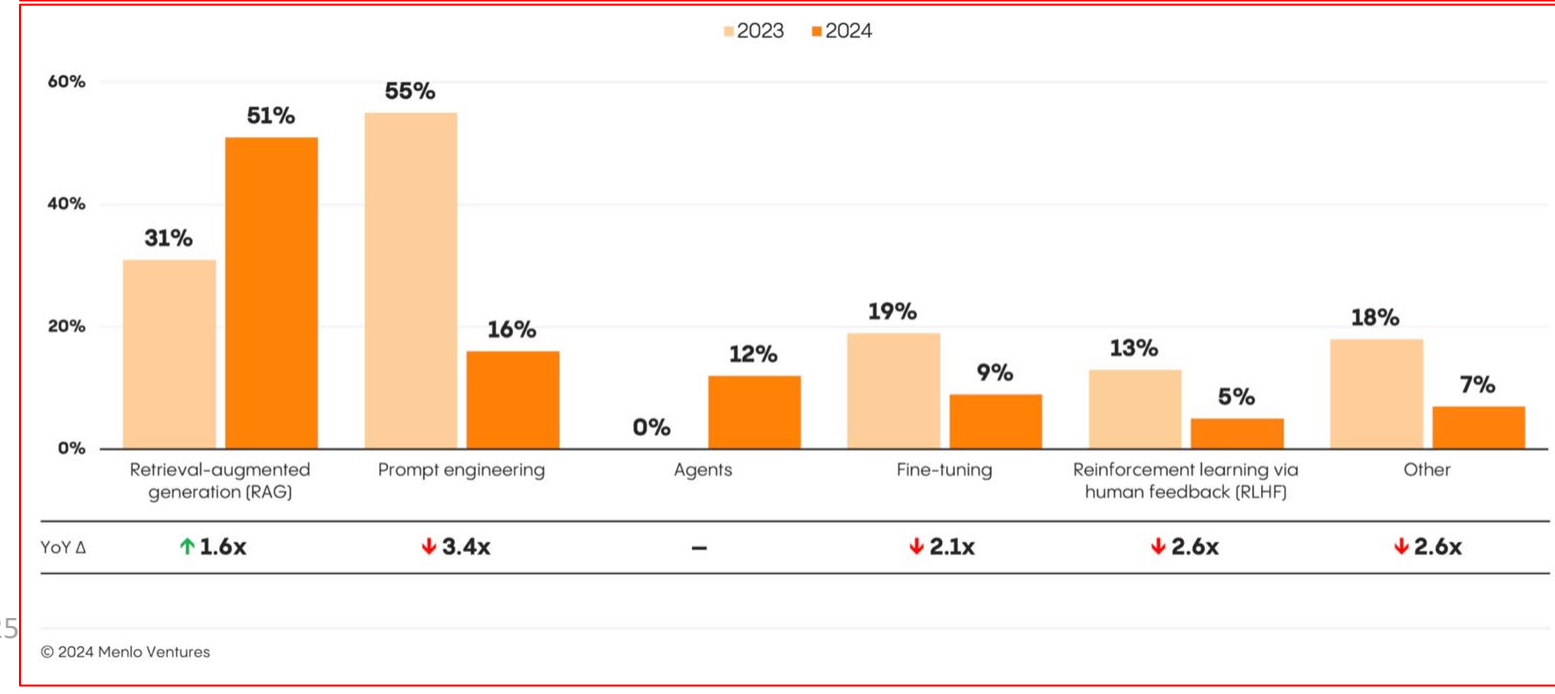
Source https://discussion.fool.com/t/notes-on-cxmt-and-ymtc-production-and-plans/118722

Enterprise Gen AI: Trends

Inference spend dominates

Inference: Memory Bound

RAG use growing



Source: https://menlovc.com/2023-the-state-of-generative-ai-in-the-enterprise-report/

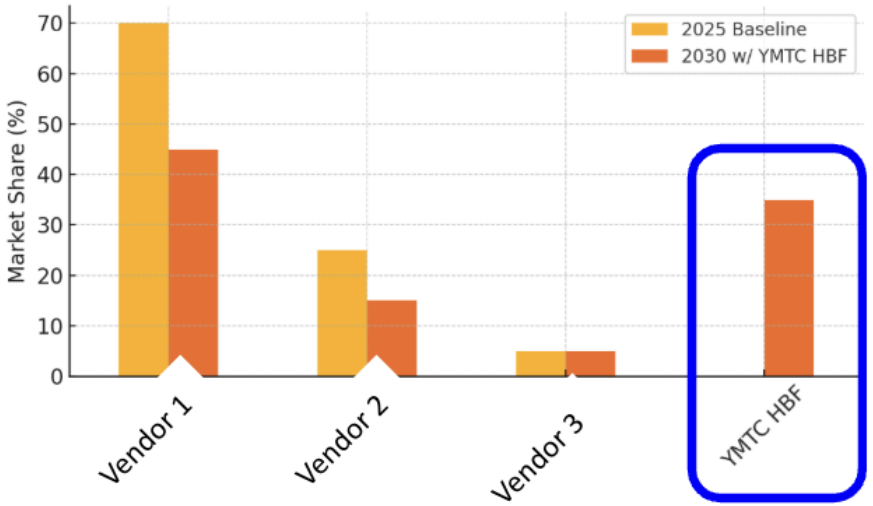
LLM Inference Unveiled: Survey and Roofline Model Insights

Threat Modeling: Hypothetical Market Share Impact: YMTC Launches HBF by 2030

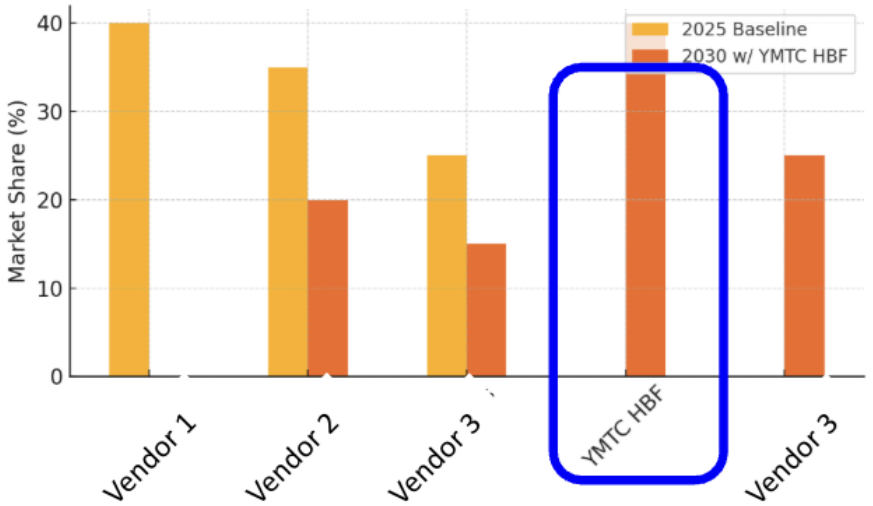
Scenario: YMTC/CXMT co-develop HBF equivalent technology

DRAM + Flash share of revenue erosion, design win loss

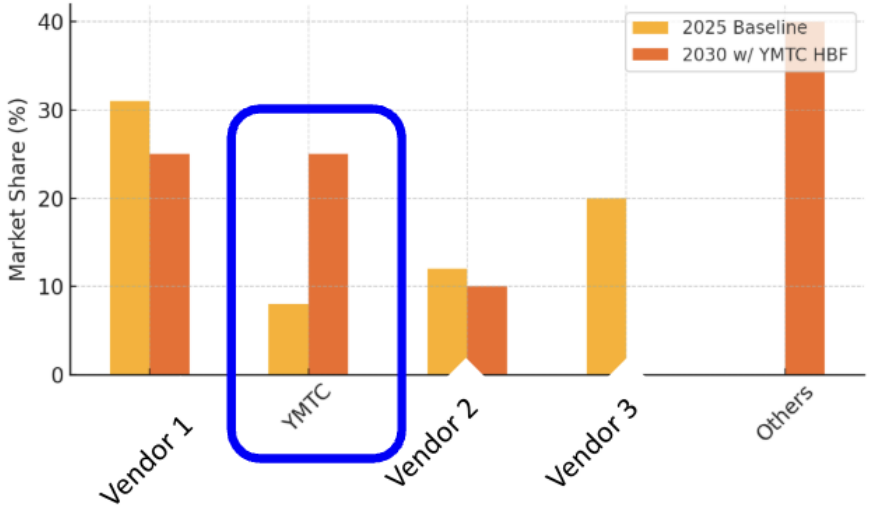
HBM DRAM



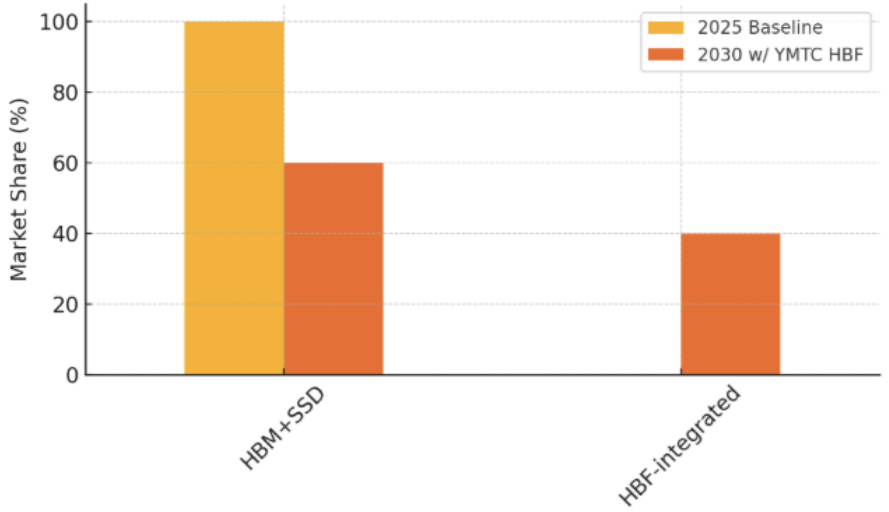
AI Inference (Flash Storage)



Flash Share of Revenue



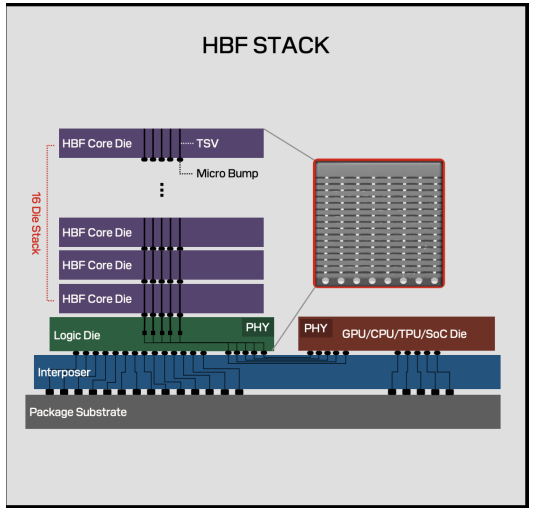
AI Accelerator Integration Base

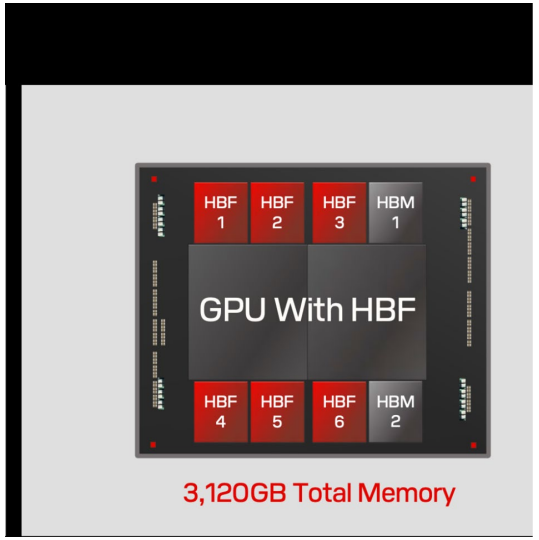


Opportunity: New Technology

Source: https://investor.sandisk.com/static-files/79481580-ada2-4e08-bdeb-4b440d08f4ab

High Bandwidth Flash (HBF), targeting 8× HBM capacity for AI inference at similar cost, Leverages **BiCS and wafer bonding 8-16X**

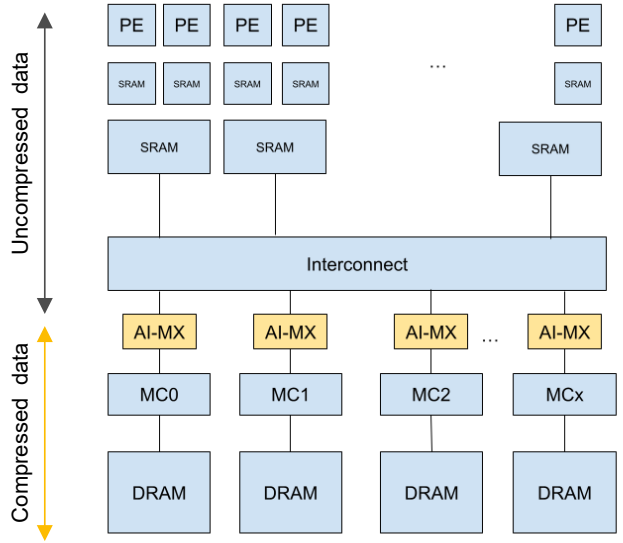


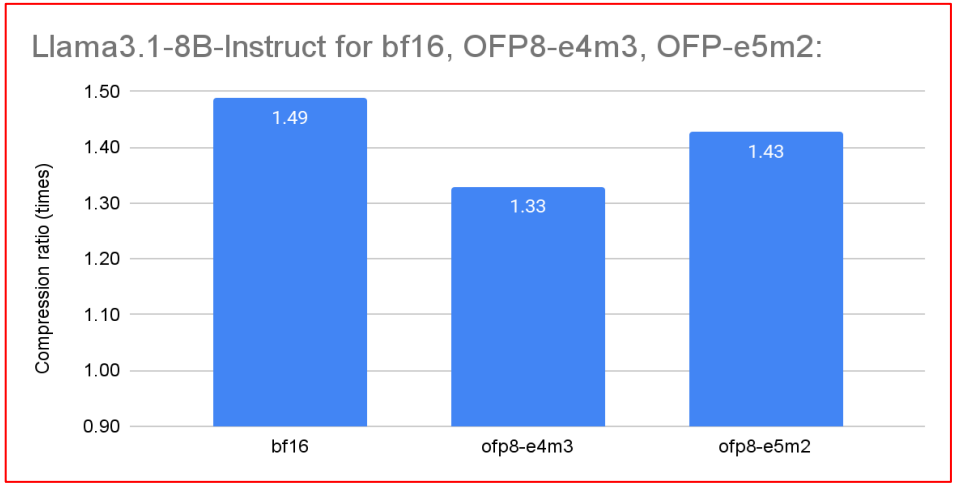


Combine HBF + inline (de)compression maintain competitiveness factor of 16-32X

Lossless HW accelerated AI Model & KV Cache (de)compression by 1.5-2X

Source: https://www.zeropoint-tech.com/products/hbm-memory-expansion





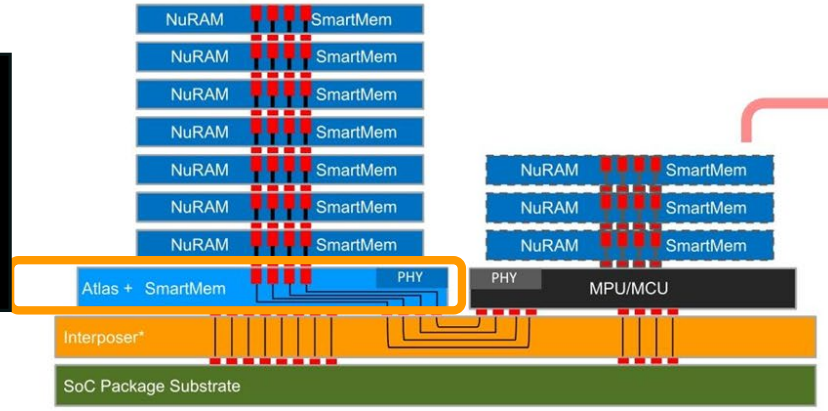
1.5X Model compression

FMS the Future of Memory and Storage

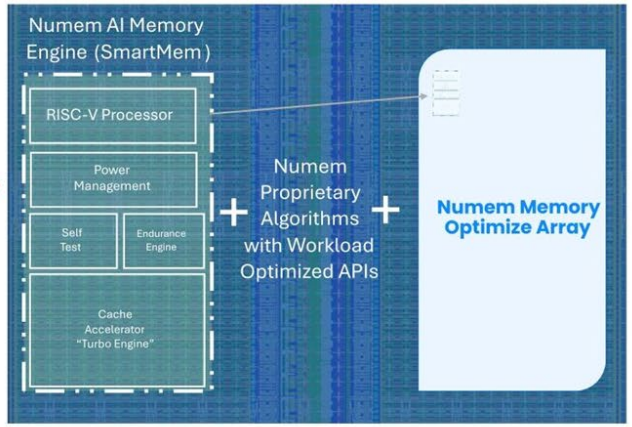
Opportunity: Managed MRAM "SmartMem" technology

Gen 1 12nm: 2-6TB/s, 3-9GB/stack

Gen 2 5nm: 4-12TB/s , 8-16GB/stack



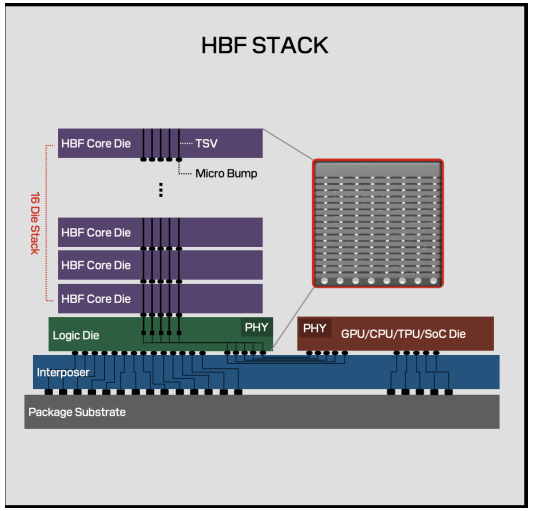
Numem SoC example (I/F : UCIe, HBM, LPDDR, etc.)



Numem Each Chiplet

Three Key Ingredients
1. AI Memory Engine
2. Numem Algorithms
3. Numem Array Design

Combine MRAM + (de)compression to gain 5X advantage

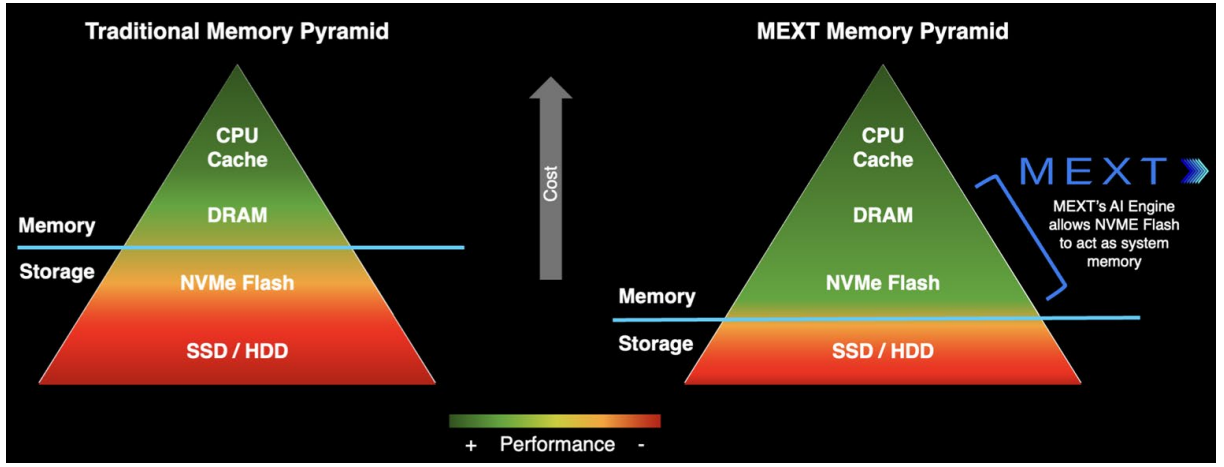


AI Memory Engine to manage HBF performance

Opportunity: New Technology

Source: https://www.mext.ai/blog/memory-finally-goes-multi-tier

Enterprise Solution: AI Powered Predictive Memory Tiering Flash + DRAM



Solve specific Hyperscale & Enterprise Memory challenges

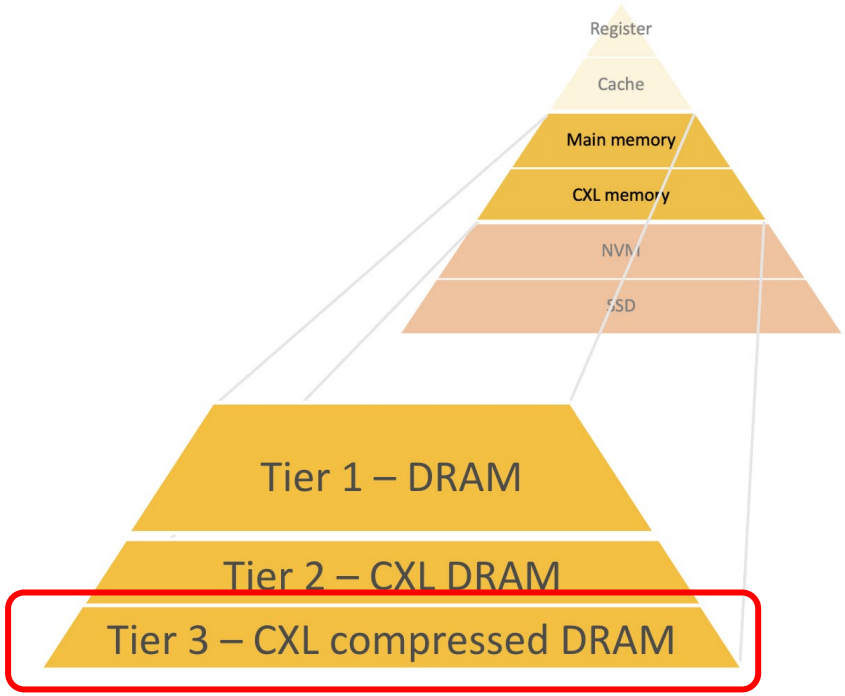
OCP Hyperscale solution: Compressed CXL Tiered Memory Tier

OPEN Compute Project

Hyperscale CXL Tiered Memory Expander Specification

Revision 1

Version 1.0
Base Specification Template v1.2
Effective October 27, 2023



FMS the Future of Memory and Storage

Summary

1. Sovereign computing is not a trend — it's a restructuring of global compute economics
2. Asia leads in **power availability, chip fab scaling, memory innovation**.

Action Required

1. US & EU must act now: **invest in NAND/HBM/CXL/HBF** ecosystems or become dependent
2. **Deep tech collaboration required**: Memory, storage, IP, software to effectively address threats

Next-gen AI infrastructure is not just compute-bound

it is **memory-sovereign**