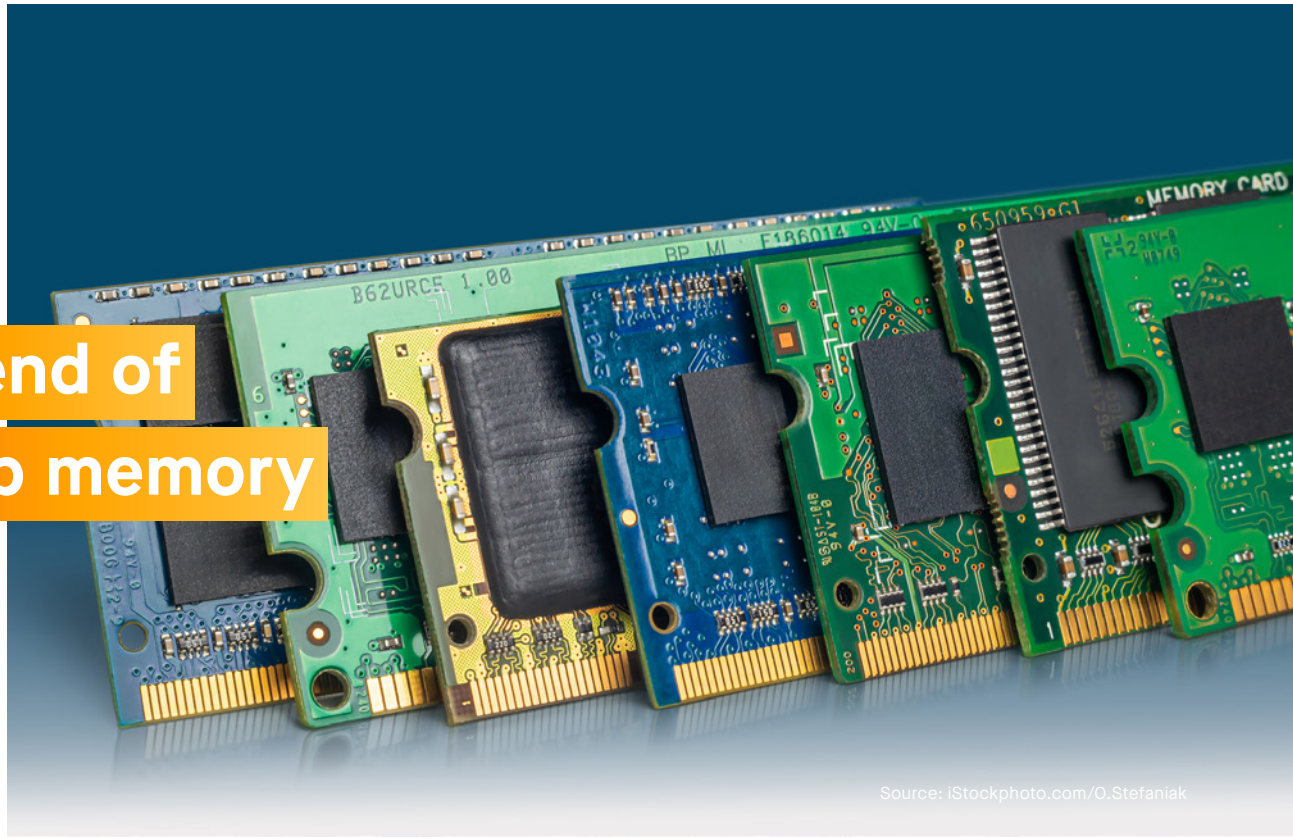
# The end of cheap memory




Source: iStockphoto.com/O.Stefaniak

Memory chips underpin every layer of the modern technology stack, yet the market that produces them has entered a period of structural constraint with no near-term resolution. The implications for companies dependent on predictable hardware access are becoming increasingly material.

**Charles-Henry Monchau**, CFA, CAIA, CMT
Chief Investment Officer
charles-henry.monchau@syzgroup.com

**Assia Driss**
Syz Research Lab Team Coordinator
assia.driss@syzgroup.com

**Hugo Morel**
Syz Research Lab Team
hugo.morel@syzgroup.com

Syz PRIVATE BANKING

## Introduction

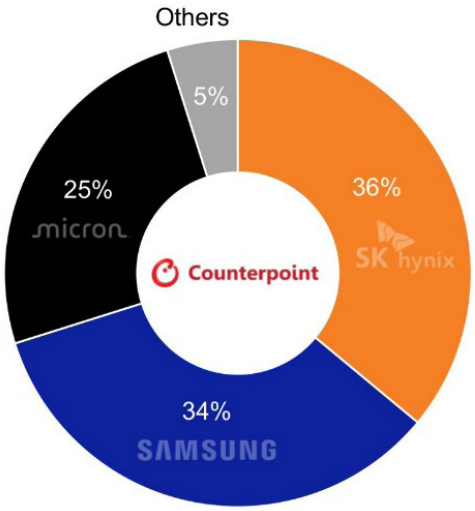
The global memory market has entered a period of sustained tension. Tight supply, rising prices, and multi-year procurement commitments are reshaping how technology companies plan, invest, and compete. For decades, memory chips were treated as a commodity: abundant and reliably cheap over time. Current conditions challenge that view. These developments affect not only the manufacturers who produce these chips, but every company whose growth strategy depends on predictable access to hardware.

## A structural shift in the memory market

For years, the memory industry followed a predictable cycle: demand rose, capacity expanded, oversupply followed, and prices corrected. It was a framework that shaped how producers invested and how markets priced the sector. In 2026, this framework is changing.

At the centre of this disruption are two chip technologies that underpin virtually every modern electronic device. The first, DRAM, or Dynamic Random Access Memory, is a device's working memory. It is what allows a laptop to multitask, a server to scale, and a modern car to function. The second, NAND Flash, is the storage technology that retains data when a device is switched off. Every photo on a smartphone, every file on a laptop, and every database in a corporate server room exists because of it. What makes this moment so far-reaching is precisely how embedded these chips have become. They are as essential to a military aircraft as to a living room television. When their supply tightens, the effects are not limited to one industry. They travel through all of them. And that supply is concentrated in the hands of just three producers: Samsung Electronics, SK hynix, and Micron, who together control roughly 90 to 95% of global DRAM production capacity.

**Q1 2025 DRAM Market Share in Revenue**




*Source: Counterpoint Research Memory Tracker.*

*Source: Counterpoint*

The origins of the current situation trace back to 2022 and early 2023, when a collapse in demand left manufacturers selling both chips below cost. Deep production cuts followed, and by the second half of 2023 prices had begun to recover: NAND first, then SSDs, then DRAM. Each previous cycle had its own trigger: the smartphone boom, the rise of solid-state drives, the expansion of cloud computing. Each time, the industry eventually found its footing.

This time is different. The rise of artificial intelligence has introduced a category of demand the industry has never previously encountered, and one that shows no sign of plateauing. Rather than managing inventory corrections or short-term pricing rebounds, executives at Samsung Electronics, SK hynix, and Micron have pivoted to forward visibility, committed AI capacity, and long-dated capital planning. Industry leaders no longer describe this as a cyclical shortage. They call it structural, and the distinction matters. Cyclical shortages resolve themselves; structural shifts reorder the market.

As Manish Bhatia, Executive Vice President at Micron, stated, the industry is facing "the most significant disconnect between supply and demand, in magnitude and time horizon, that we have seen in a quarter century." Markets have drawn their own conclusions. Micron's shares rose from around $102 at the start of 2025 to a peak of $455 in January 2026, a gain of over 340% in twelve months. SK hynix delivered returns of approximately 145% across the same period. Such performance suggests that markets view the current imbalance as more than a temporary dislocation.

## How demand is transforming the industry

Memory markets are entering a different phase of demand formation. AI infrastructure is materially increasing memory intensity per system, altering the balance between supply, pricing, and capital allocation across the industry.

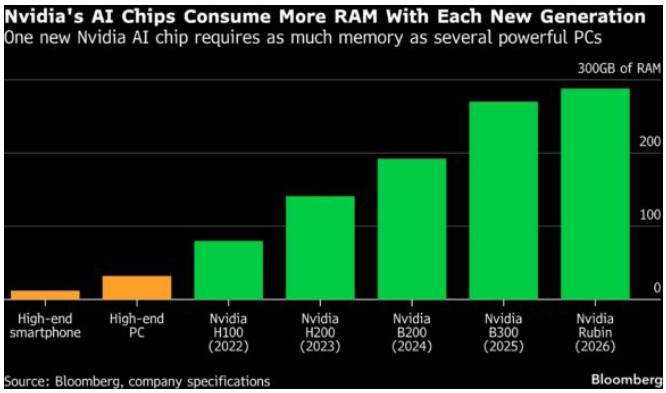
The dominant force behind this transition is the scale of memory embedded in modern AI systems. Artificial intelligence models are trained and deployed on clusters of high-performance GPUs, each of which requires large volumes of high-speed memory positioned directly alongside the processor. As model size and computational complexity increase, memory requirements rise proportionally.

Nvidia's Blackwell architecture integrates up to 192 gigabytes of memory per chip, and a single rack of AI servers can contain more than 13 terabytes. By contrast, a high-end personal computer typically operates with around 32 gigabytes of memory, and a smartphone with 8 to 12 gigabytes. AI systems therefore consume multiples of the memory used in traditional end-markets, even before accounting for the scale of deployment.

This intensity matters because AI infrastructure is being built in concentrated, industrial-scale increments. When

hyperscalers expand a data centre, they are not adding thousands of devices, but entire clusters of memory-dense systems at once. Each incremental buildout translates into a step-change in memory demand.

The scale of capital committed reinforces this dynamic. Combined data-centre investment from Microsoft, Meta, Amazon and Alphabet is projected to reach roughly $650 billion in 2026, up from approximately $217 billion in 2024.




*Source: Bloomberg*

Demand is therefore embedded in long-term infrastructure expansion rather than short-term consumer refresh cycles, giving this driver structural weight within the industry. If AI has structurally increased memory demand, the next question is how supply is responding.

At the advanced end of the DRAM market, including HBM production, capability remains confined to Samsung Electronics, SK hynix and Micron. While Chinese suppliers such as CXMT and YMTC account for an estimated 5–10%, largely in lower-end segments, they do not currently possess the advanced HBM capabilities required for leading AI systems.

In this structure, capacity allocation is decisive. A central development has been the expansion of High Bandwidth Memory, or HBM, which is required for advanced AI accelerators. HBM is built differently from conventional DRAM. Instead of arranging memory dies side by side, manufacturers stack multiple layers vertically, typically 12 to 16 layers, and connect them through through-silicon vias. This design allows data to move much faster, which is essential for AI processors.

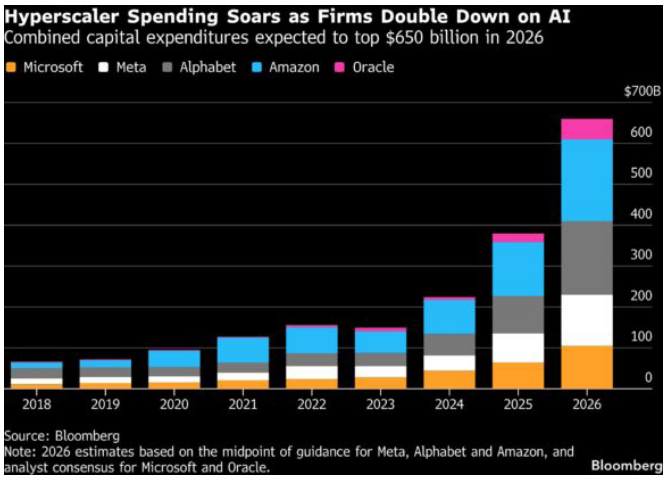
The manufacturing process is complex and capacity intensive. Producing one bit of HBM effectively displaces several bits of conventional DRAM output. By 2026, HBM is expected to account for roughly 25% of total DRAM wafer production, with demand growing around 70% year-on-year. At the same time, hyperscaler expansion has translated into forward capacity reservations. Large AI initiatives have secured substantial portions of future DRAM wafer output under multi-year agreements, tightening availability for traditional end markets.

This strategic reallocation comes at the expense of traditional DRAM and NAND used in PCs, smartphones, and automotive applications. Micron has scaled back parts of its consumer PC exposure to preserve server capacity. Samsung has announced the discontinuation of MLC NAND Flash, with final deliveries expected in 2026, and analysts anticipate a contraction of more than 40% in global MLC NAND capacity.

Given that more than 95% of market share is concentrated among three producers, and that advanced manufacturing capabilities are limited outside this group, pricing and allocation decisions remain highly centralised. The shortage is therefore amplified not only by rising demand, but by deliberate capital reorientation within a technically constrained oligopoly.

Even if producers wished to increase output rapidly, the structure of the industry constrains how fast this can happen.

Memory manufacturing is among the most capital-intensive segments of the semiconductor industry. Constructing a new fabrication facility typically requires $10–20 bn in investment and at least two to three years before meaningful production begins. After completion, additional time is required to ramp yields and reach stable output.




*Source: Bloomberg*

These timelines matter because capacity decisions are made years in advance. Supply cannot be increased in small increments, nor can advanced production lines be easily converted between product categories without further capital expenditure and engineering work.

Planned expansions from Samsung and SK hynix are expected to come online between 2027 and 2028, but these facilities will also require ramp-up periods and are largely oriented toward advanced nodes. In the near term, they do not materially ease current constraints.

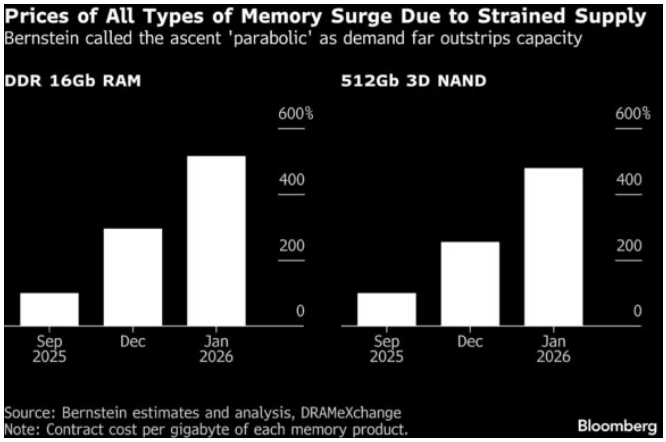
The result is limited elasticity. Demand has accelerated, capacity is concentrated, and expansion requires years. The imbalance persists not only because of allocation decisions, but because the industry cannot scale quickly enough to restore equilibrium.

Beyond AI-driven demand and capacity reallocation, the memory supply chain is also exposed to volatility in commodities and raw materials. Semiconductor production relies on metals such as copper for interconnects, gold and silver in bonding, palladium for contacts, and gallium and germanium in certain advanced components. Geopolitical tensions and export restrictions have increased

uncertainty around the cost and availability of some of these inputs. While metals are not the main driver of memory pricing, swings in commodity markets can add cost pressure and further limit supply flexibility in an already tight environment. When capacity is constrained, even secondary input risks can intensify broader supply-demand imbalances.

## The road ahead

The immediate response has been to secure available memory. Hyperscalers lock in future supply via multiyear deals with Samsung, SK Hynix, and Micron, with much production already sold HBM is sold out for 2026. DRAM prices are set to rise 50–55% this quarter versus Q4 2025.




*Source: Bloomberg*

OEMs like Acer, HP, and Dell are diversifying suppliers, including Chinese makers CXMT and YMTC. Manufacturers design multigeneration hardware (DDR4/DDR5) and optimise software memory use to extend inventories. 8 GB on an Apple M3 MacBook can match 16 GB elsewhere, highlighting efficiency over sheer volume.

The strategy focuses on securing supply, spreading risk, and using memory efficiently, rather than quickly fixing shortages.

Medium-term solutions expand industrial capacity. Micron is expanding its manufacturing footprint with new fabrication facilities in Singapore and Taiwan (2027), followed by a US site in New York (2030). SK Hynix is also increasing capacity, with additional facilities planned in South Korea (2027) and the US (2028). Samsung launches new lines in South Korea (2028). These $10bn+ projects aim to reduce East Asia dependence and improve resilience but lead times of 2-4 years mean relief would not arrive before 2027–2028.

Geopolitical tensions, export limits, and industrial policies continue to fragment supply chains, keeping costs high. Long-term stability needs geographic diversification, redundancy, nearshoring, strategic stockpiles, and memory recycling. Acts like CHIPS and European Chips Act target reduced reliance on South Korea and Taiwan.

## Conclusion

Analysts expect supply shortages to gradually ease by 2027–2028. However, pricing is likely to remain elevated, supported by sustained AI-driven demand and rising capital intensity.

Memory was once a component. It is now a strategic, capital-intensive asset in an economy built on data. What it becomes next will depend on which companies and governments move fastest to control it.

# Welcome to Syzerland®

**For further information**


**Banque Syz SA**
Quai des Bergues 1
CH-1201 Geneva
T. +41 58 799 10 00
syzgroup.com

**Charles-Henry Monchau**, CFA, CAIA, CMT
Chief Investment Officer
charles-henry.monchau@syzgroup.com

**Assia Driss**
Syz Research Lab Team Coordinator
assia.driss@syzgroup.com

**Hugo Morel**
Syz Research Lab Team
hugo.morel@syzgroup.com





This marketing document has been issued by Bank Syz Ltd. It is not intended for distribution to, publication, provision or use by individuals or legal entities that are citizens of or reside in a state, country or jurisdiction in which applicable laws and regulations prohibit its distribution, publication, provision or use. It is not directed to any person or entity to whom it would be illegal to send such marketing material.
This document is intended for informational purposes only and should not be construed as an offer, solicitation or recommendation for the subscription, purchase, sale or safekeeping of any security or financial instrument or for the engagement in any other transaction, as the provision of any investment advice or service, or as a contractual document. Nothing in this document constitutes an investment, legal, tax or accounting advice or a representation that any investment or strategy is suitable or appropriate for an investor's particular and individual circumstances, nor does it constitute a personalized investment advice for any investor.
This document reflects the information, opinions and comments of Bank Syz Ltd. as of the date of its publication, which are subject to change without notice. The opinions and comments of the authors in this document reflect their current views and may not coincide with those of other Syz Group entities or third parties, which may have reached different conclusions. The market valuations, terms and calculations contained herein are estimates only. The information provided comes from sources deemed reliable, but Bank Syz Ltd. does not guarantee its completeness, accuracy, reliability and actuality. Past performance gives no indication of nor guarantees current or future results. Bank Syz Ltd. accepts no liability for any loss arising from the use of this document.