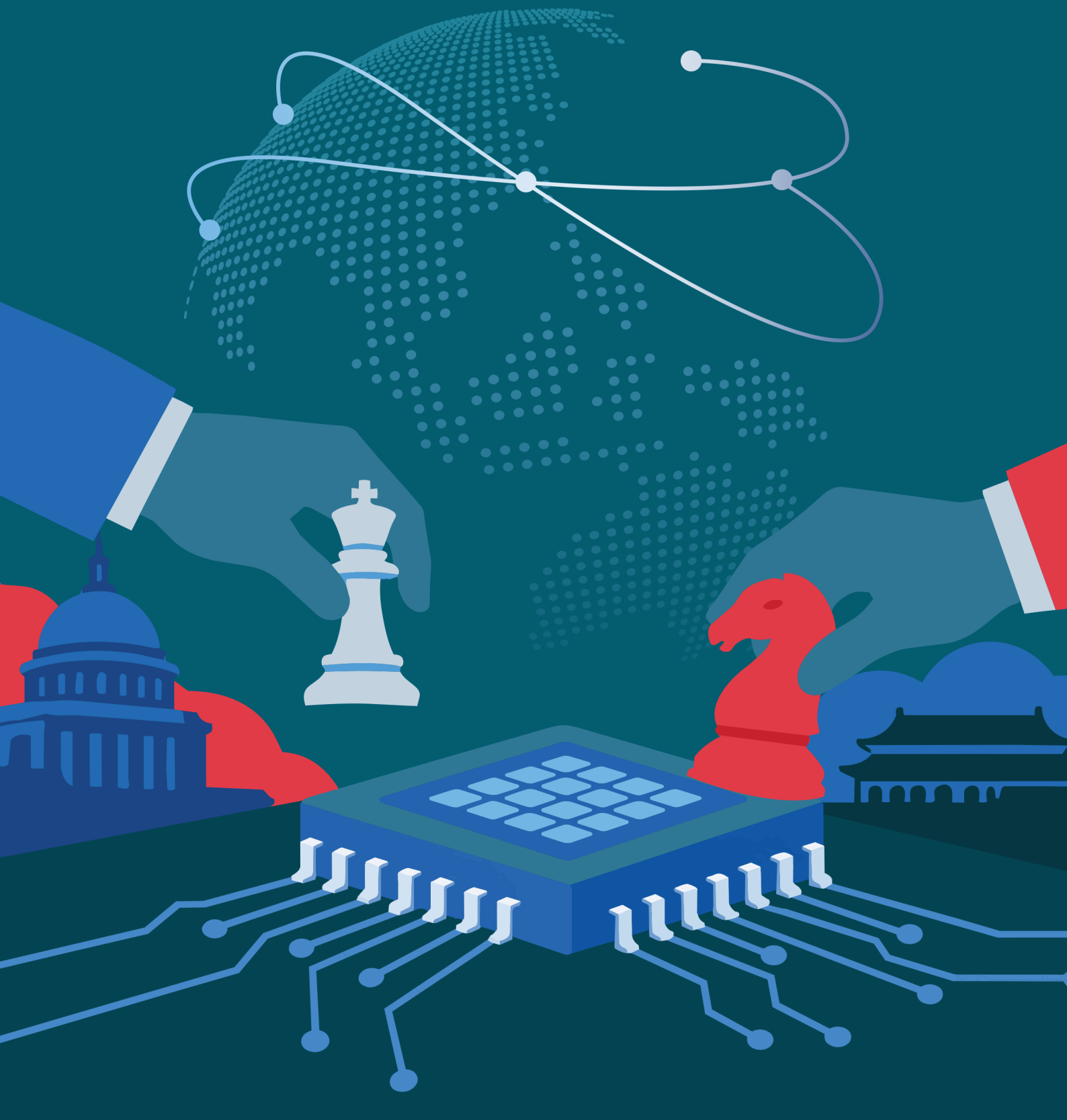
STONEBENCH

The politics of chip access



Executive summary

On 28 October 2026, SK hynix reported record quarterly revenue of ₩104.6 trillion (US$78.2 billion) and said its dynamic random-access memory (DRAM) and high-bandwidth memory (HBM) output was effectively sold out through 2029 once demand under negotiation is counted, with customer requests for HBM more than double of what it could supply. In the same quarter global smartphone shipments fell 20% year on year to their lowest third-quarter level in 14 years, and PC shipments fell 21% to 57 million units. Meanwhile, Texas Instruments, a bellwether for analogue chips, reported that third-quarter revenue was up by 28% based on broad growth in industrial, automotive and data-centre markets.

A general chip shortage of the 2021 kind would have squeezed analogue suppliers and device makers alike. The evidence points to an acute, price-rationed shortage of memory and of the advanced-packaging, substrate and tool capacity needed to build AI systems. That shortage is now spilling into a narrow set of mature components as suppliers move shared capacity to AI work, while factories keep running.

We think this squeeze will persist through at least 2027 because its binding constraint is new wafer and packaging capacity, which arrives in volume only from mid-2027 onwards. Prices may peak before physical tightness ends, as memory's cyclical history and early demand destruction in consumer devices suggest. Jensen Huang's resignation from NVIDIA on 12 November 2026 alters the politics of chip access, particularly to China. But there is no evidence that it alters physical supply.

Supply constraints and their underlying drivers

The shortage is concentrated in memory, and its primary cause is that suppliers have moved wafer capacity to AI products. Data centres are projected to consume roughly 70% of all memory devices produced worldwide in 2026, according to the Center for Strategic and International Studies. HBM stacks multiple DRAM dies beside an AI processor, so each wafer diverted to it removes commodity DRAM from phones, PCs and cars. Some have attributed supply tightness partly to suppliers concentrating production lines on HBM. Samsung said on 29 October 2026 that its 2027 HBM supply was already sold out, and Micron said on 30 September 2026 that it expects supply-demand conditions to be "much tighter in calendar 2027 and 2028 than they were in 2026".

Most of the rationing is happening through price. Memory prices more than doubled between the third quarter of 2025 and the first quarter of 2026, then rose 80%–90% in the second quarter and 15%–30% in the third, according to figures reported by Seoul Economic Daily. TrendForce expects conventional DRAM contract prices to rise 10%–15% quarter on quarter in the fourth quarter, a marked slowdown from earlier in the year. NAND flash is following the same path, with forecasts of fourth-quarter contract prices rising up 15%–20% as AI servers absorb enterprise storage. It was reported in late October 2026 that PC DRAM gains were narrowing amid weak notebook shipments and rising inventory, early evidence that consumer demand destruction is beginning to cap prices at the low end. Nonetheless, DRAM price increases are expected to regain momentum in the first quarter of 2027. Some buyers are substituting cheaper parts, and TrendForce expects reballed memory, recovered chips re-soldered onto new boards, to reach 7%–9% of smartphones in 2027 as low-end manufacturers avoid paying for new chips.

Data centres are projected to consume roughly 70% of all memory devices produced worldwide in 2026, according to the Center for Strategic and International Studies.

Leading-edge logic, and the packaging that binds it to memory, is also short of capacity. TSMC told investors on 15 October 2026 that customer demand far exceeds supply, with high-performance computing at 68% of third-quarter revenue. Substrate materials and chipmaking tools are rationed too. Unimicron's chairman warned on 20 October 2026 that lead times for some key substrate equipment had stretched to 30 months and that critical materials could be 20% short in 2027. ASML said on 14 October 2026 that the capacity it had planned for 2027 and 2028 "will not fully meet the demand" of its customers, so the equipment needed to add capacity is itself in short supply.

Mature-node logic, analogue and power chips remain broadly available but are tightening, most visibly in analogue. Texas Instruments' chief executive said on 21 October 2026 that the supply-demand mismatch is greatest in analogue, where some capacity has gone offline and some foundries are moving to other products, and that lead times have ticked up. Infineon closed its fiscal year with a record quarter and forecast fiscal 2027 revenue growth of more than 20%, while NXP's shares fell after its 27 October 2026 results, as a forecast raised concern that the automotive recovery was stalling; that is a sign of weak demand. The sharpest pressure arises where contract foundries and packaging houses have moved capacity to AI work. Nexperia's interim chief executive Stefan Tilger said the company was "totally oversubscribed", and STMicroelectronics' chief executive Jean-Marc Chery warned that some foundries and assembly firms were "disengaging" from mainstream technology. Distributor Sourceability reported select MCU and FPGA products at 40–52 weeks in its third-quarter survey, with high-end multilayer ceramic capacitor lead times stretching towards 10 months. These are pockets of allocation driven by AI's pull on shared capacity, not the across-the-board mature-node scarcity of 2021.

Power and construction delays are also holding back AI deployment. Amazon has bought NVIDIA systems that are waiting in warehouses until data-centre power is available, and Microsoft attributed its lower capital spending to delayed third-party capacity deliveries. These delays do not reduce chip output, but means that part of today's demand is early purchasing, and that stock would let buyers cut new orders if spending slowed.

The sharpest pressure arises where contract foundries and packaging houses have moved capacity to AI work.

Confidence: high that memory and advanced packaging are the binding constraints; medium on how far mature-node tightness spreads; low on when consumer-device demand destruction caps DRAM prices.

How this cycle differs from the pandemic shortage

The 2026 squeeze differs from the pandemic shortage in cause and composition, and therefore in cure. The acute phase of 2021–22 was a disruption-and-restocking event concentrated in mature-node logic. Federal Reserve researchers found that microcontroller lead times roughly doubled from about 15 weeks to about 30, with analogue chips peaking near 29 weeks. Buyers had run down buffers, with the median inventory of semiconductor products held by customers falling from 40 days in 2019 to less than five days in 2021. Carmakers paid in lost production, and AlixPartners forecast in September 2021 that 7.7 million vehicles of production would be lost that year.

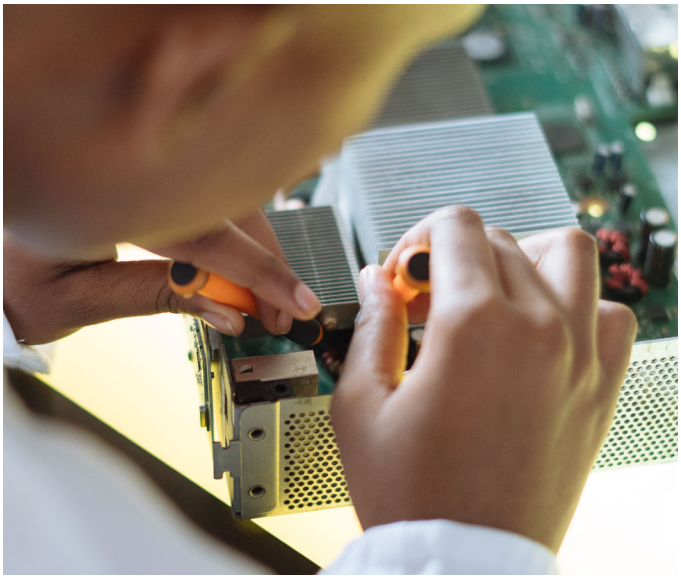


Image by RF Studio

Today's constraint is a reallocation driven by price, and it followed the glut that ended the pandemic episode. Micron's revenue roughly halved to US$15.54 billion in fiscal 2023 from US$30.76 billion a year earlier and it posted a net loss of US$5.83 billion. That downturn left suppliers cautious about new capacity just as demand shifted to buyers able to outbid everyone else. Data centres take roughly 70% of memory output, so carmakers and device makers are small customers who cannot outbid them.

Chip revenue in 2026 is far above its level during the pandemic shortage. Global semiconductor sales reached US$517.7 billion in the third quarter of 2026, and September sales were up 146.8% year on year; the whole of 2021, when the pandemic shortage was most acute, semiconductor sales generated US$555.9 billion. Much of the 2026 increase is price rather than volume, so revenue alone cannot be read as evidence of physical scarcity.

Ordering behaviour has changed in ways that both prolong the shortage and leave some room for a reversal. In 2021 the distortion came from double ordering that later unwound into excess inventory. In 2026 suppliers are locking customers into take-or-pay arrangements. By 30 September 2026, Micron had signed 26 strategic customer agreements, which it estimates at more than 35% of its revenue through 2030, backed by US$32 billion of customer financial commitments, mostly cash deposits. Contracted demand is less likely to vanish suddenly than the phantom orders of 2021, which supports a long shortage. But IDC attributes part of the third-quarter PC slump to inventory pulled forward earlier in the year, a sign that some precautionary buying is present again at the margin.



Image by Tima Miroshnichenko

Global semiconductor sales reached US$517.7 billion in the third quarter of 2026, and September sales were up 146.8% year on year; the whole of 2021, when the pandemic shortage was most acute, semiconductor sales generated US$555.9 billion.

The remedies that ended the pandemic shortage will not end this one. The earlier shortage eased as factories reopened, logistics normalised and over-ordering unwound. Today the factories and the logistics are working, and the output goes to data-centre buyers who pay more. The resilience strategies carmakers built after the pandemic centred on buffer stocks, which protect against factory and shipping disruption but cannot be built from output already sold under multi-year agreements. Relief must come from new fabrication and packaging capacity, which takes years to build.

Sector exposure and commercial implications

The consequences of the chip shortage are distributed by bargaining power rather than by geography. Buyers with scale, pricing power or long-term contracts secure supply at a higher cost, while smaller and price-sensitive buyers lose volume, specification or margin.

AI and cloud infrastructure: securing supply at a premium. The largest buyers are protected on availability but not on cost, and they are converting cash into multi-year claims on memory. NVIDIA's supply and capacity commitments rose from US$119 billion to US$279 billion in a single quarter, "primarily related to the procurement of memory", and the company described its forecast of roughly 70% revenue growth in fiscal 2028 as supply-constrained. Alphabet held its 2026 capex guidance at US$195 billion–US$205 billion and Microsoft lowered its calendar 2026 figure to about US$170 billion from US$175 billion as third-party capacity arrived late. Network suppliers pay the memory bill directly. Cisco's

adjusted gross margin fell to 65.6% from 68.1% a year earlier in its quarter reported on 12 November 2026, reflecting a higher hardware mix and elevated memory costs. AI buyers are therefore pre-empting capacity that would otherwise go to the sectors below.

Consumer electronics: higher costs and market consolidation. The shortage is most visible in consumer electronics, where memory now accounts for 54% of the bill of materials of budget smartphones, up from 20% a year earlier, and average smartphone retail prices have risen 20% this year. Apple took a record 50% share of global smartphone revenue in the third quarter, and its iPhone average selling price reached US$1,031, up 19% year-on-year. Even so, Apple guided its December-quarter gross margin to 45%–46% from 47.9% as memory costs rose; the September 2026 figure included a benefit of about one percentage point from tariff refunds. Xiaomi's share fell to 12% as it reduced exposure to lower price segments. In PCs, IDC data showed HP's third-quarter shipments down 30.9% and Dell's down 25%. The burden falls hardest on emerging markets that rely on low-priced handsets. Premium brands pass costs on, while volume brands lose units.

Automotive: cost pressures and selective disruption. Cars are less exposed to outright shortage as compared to 2021 but more exposed to cost and parts discontinuation. Aumovio, a supplier to BMW and Ford, expects at least €1 billion (US$1.1 billion) of gross cost headwinds in 2027 from higher prices for semiconductors, memory and raw materials. Visteon, which buys more than 25 different DRAMs, some of which are going out of production, expects incremental electronic-component costs of about US$500 million in 2027. Rivian's chief executive RJ Scaringe said the company now spends an "enormous amount of time" planning around memory chips, and China's Li Auto has disclosed "significant supply tightness and price volatility" in DRAM and NAND for its cockpit and driving systems. General Motors and Ford had already signed long-term supply agreements with Micron in July. Stoppages are real but localised and not always chip-driven. Stellantis said on 2 November 2026, that a shortage of an engine-control component was disrupting Fiat output in Italy and could last through December 2026, though unions attribute the extended Mirafiori halt partly to weak Fiat 500 demand. BMW cancelled shifts at Regensburg, Leipzig and Oxford because of unspecified supplier shortages. The likely business consequence is thinner margins and vehicles built to lower specifications.

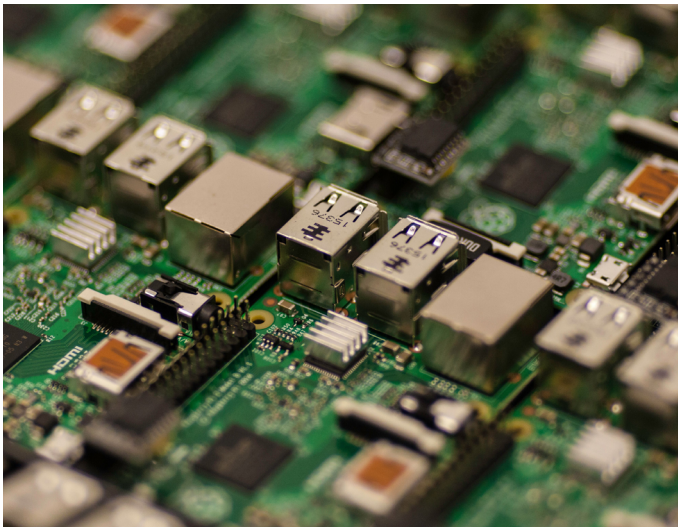


Image by Craig Dennis

Industrial, network and medical equipment: margin and qualification pressures. Rockwell Automation has said inflation, particularly in memory, is a cost pressure it is offsetting through price increases, while describing memory availability as manageable, and Cisco raised list prices on software and select hardware from 7 November 2026. AdvaMed, the US medical technology association, has warned that memory supply-demand imbalances may persist well into 2027. These sectors use long product cycles and qualified parts, so their larger exposure is to end-of-life notices that force costly redesign and requalification. Industrial demand remains strong, and Texas Instruments reported industrial revenue up more than 40% year on year.

Buyer scale and pricing power: divergent commercial outcomes. In every sector, firms that can sign multi-year contracts or pass on prices are better placed than those that cannot. Aumovio's interim finance head Claus Bauer said the data-centre segment has "no price sensitivity whatsoever", and that suppliers could sell twice their capacity to it. At the other end, weak notebook sales are pushing second-tier PC makers towards white-label models with downgraded chips and smaller orders for next year, and some smaller phone makers are reported

to be considering recycled or lower-grade memory. We judge that the shortage is accelerating consolidation in consumer hardware, as brands with pricing power gain share and second-tier assemblers lose both volume and specification. Readers assessing counterparties in Asia, Europe or North America should look first at contract coverage and pricing power.

Geopolitical constraints on supply and market access

Geopolitics chiefly determines who may buy advanced chips. Over the next six months it has less bearing on how many memory wafers the world can produce. The most consequential instrument is the licensing of NVIDIA's H200 to approved Chinese customers, announced in December 2025 with 25% of revenue going to the US government. NVIDIA's guidance for the quarter it reports on 17 November 2026 assumes no data-centre compute revenue from China, so China is a policy-constrained market rather than a driver of global scarcity.

The NVIDIA scandal has shifted US politics against easier chip access for China. Lawmakers are seeking to fold the AI Overwatch Act, the Chip Security Act and the MATCH Act into the defence authorisation bill, and David Meale of Eurasia Group described Huang as "the industry's most effective lobbyist against tighter rules". On 12 November 2026, Senator Ron Wyden and Representative Richard Neal opened an inquiry into the administration's approval and taxation of NVIDIA's China sales. Chinese AI firms are reportedly rushing to rent NVIDIA-powered capacity in data centres outside China before Washington restricts remote access, and the Commerce Department is reported to be finalising a rule to close that route, a direct exposure for operators in Singapore and Malaysia. Scrutiny in those hubs is rising. Singapore said on 12 November 2026 that it had barred Alice Huang from re-entry in February 2026, after police found she knowingly gave investigators inaccurate information, though they found insufficient evidence that she had committed offences. Malaysia, which has required permits to re-export US AI chips since July 2025, said on 13 November 2026 that it was working with customs to establish whether chips brought in by a Megaspeed subsidiary had been re-exported. Compliance costs and customer due diligence for data-centre investors in Southeast Asia are rising faster than chip availability.

The US and Chinese export-control deadlines now fall on the same day. The Commerce Department has extended its suspension of the "Affiliates Rule" on export controls through 10 January 2027, and China has extended its suspension of the rare-earth export controls it announced in October 2025 to the same date, while its earlier licensing of seven heavy rare earths stays in force. The matching dates are consistent with each side holding its leverage in reserve. Ilaria Mazzocco of CSIS has noted that China's dominance in rare earths could limit Washington's room to tighten restrictions. Donald Trump met Xi Jinping on the sidelines of the APEC summit in Shenzhen on 19 November 2026, where agreed on having minimal guardrails with respect to AI. In our assessment, tighter US chip curbs would raise the risk that China restricts inputs on which Western manufacturing depends. Paul Triolo of DGA-Albright Stonebridge Group has warned that Beijing could respond to the MATCH Act by tightening controls on rare earths and magnets.

Diversification will not change availability within the outlook horizon. Micron's first Idaho fab is scheduled for first DRAM wafer output in mid-calendar 2027. The Nexperia split shows the cost of politically driven fragmentation: its Dutch and Chinese operations remain separated and in litigation more than a year after the Dutch intervention. Until customers have qualified it, duplicated capacity raises costs without adding usable supply or reducing concentration risk.

NVIDIA's leadership transition and its implications

Huang's departure is a governance and policy shock to the dominant supplier of AI processors, estimated to control about 90% of the market for advanced AI chips, and has not disrupted supply. On 29 October 2026, the New York Times reported that Huang had been the target of a Chinese intelligence effort involving Alice Huang, a

former executive of Singapore-based Megaspeed. NVIDIA told the newspaper that Megaspeed was like any other regional customer and that it had previously investigated and debunked claims that Megaspeed was smuggling chips. On 8 November 2026, NVIDIA's board opened an independent investigation overseen by independent directors with outside counsel, and on 10 November 2026, it retained Joon H. Kim of Wachtell, Lipton, Rosen & Katz. The Commerce Department inquiry into Megaspeed was halted after NVIDIA's general counsel contacted investigators. On 12 November 2026, NVIDIA announced that Huang had resigned as president, chief executive and director with immediate effect; chief financial officer Colette Kress became interim chief executive while retaining her finance role, and Harvey Jones was named board chair.

The espionage allegations remain unproven. US Attorney General Todd Blanche said on 12 November 2026 that the Justice Department had not concluded whether Alice Huang was a Chinese spy. China's foreign ministry spokesperson Guo Jiakun called the allegation "unsubstantiated" and "nothing but smears against China", and Megaspeed, which has denied wrongdoing, has commissioned a census of its GPUs.

Operationally, the evidence so far points to continuity. The board said it expects operations, customer commitments and the product roadmap to continue without interruption, Kress has held calls to reassure customers and suppliers, according to a person cited by Reuters, and South Korea said on 9 November 2026 that contracted NVIDIA GPU deliveries would proceed. NVIDIA's output depends on TSMC wafers and packaging and on HBM, as its US$279 billion of memory-heavy commitments show, and none of those inputs changes with the chief executive.

The commercial and policy consequences are larger. The shares closed down 4.34% on 12 November 2026, and the week's reporting erased about US$500 billion of value. The larger loss is NVIDIA's most effective advocate in Washington, at a time of growing congressional scrutiny. Senate Banking Committee chair Tim Scott has asked Huang and Commerce Secretary Howard Lutnick to testify, and Lutnick has agreed to appear. House committee chairs have requested a closed-door interview with Huang on 3 December 2026 and threatened a subpoena if he does not accept. This could mean that China's access is now more likely to narrow than widen, and that customers' main exposure is execution risk on a founder-built roadmap rather than delivery risk.

At the time of writing, Jensen Huang is back at the helm of Nvidia. The board voted unanimously on 23 November to reinstate him as CEO and president, with immediate effect. The move reverses a decision taken only days earlier, when the chair said the directors did not intend to bring him back. Markets have cheered this development, at least for now, with Nvidia shares jumping about 31% on Monday. However, the board's independent investigation will continue and so will government scrutiny.

The larger loss is NVIDIA's most effective advocate in Washington, at a time of growing congressional scrutiny.

The outlook for supply normalisation

Relief will arrive unevenly, and for memory mostly beyond six months. Over the next one to three months, to around February 2027, the central case is continued tightness, with price increases slowing in the fourth quarter before picking up again in the first quarter as contract negotiations point higher. Counterpoint expects memory prices to keep rising through the first quarter of 2027, with mobile memory rising sharply again, and possibly to stabilise only around mid-year; TrendForce likewise expects DRAM increases to gain momentum in the first quarter of 2027. Over the following three months, to around May 2027, HBM4 output for NVIDIA's Vera Rubin platform, which entered production shipments in August 2026, is likely to absorb any yield gains, leaving little for commodity DRAM. Genuinely new wafer capacity begins only with Micron's Idaho fab in mid-2027, and Unimicron's NVIDIA-backed substrate expansion in Taiwan is expected to begin adding capacity only in the second quarter of 2027, reaching full capacity by mid-2028.

A faster-easing case would require consumer demand destruction to spread and AI buyers to pause. Early signs include falling PC and phone volumes and narrowing PC DRAM gains. If hyperscalers delayed orders because power is not ready, as Amazon's warehoused systems hint, commodity DRAM and mature components would loosen first, while HBM and advanced packaging would stay tight.

A prolonged-tightness case would combine tighter US export controls, Chinese retaliation through rare earths after 10 January 2027, and substrate shortfalls of the scale Unimicron has warned of. It would hit the automotive and industrial sectors through parts discontinuation and requalification as much as through price. Evidence for it would include further mature-node capacity exits of the kind Chery described and suppliers extending sold-out horizons beyond those already stated.

Different segments will ease at different speeds. Analogue, power and most microcontrollers should remain largely available, though lead times are lengthening and allocation could spread beyond the specific parts where contract foundries have withdrawn capacity. Commodity DRAM and NAND are likely to stay expensive into mid-2027 even if price increases slow. HBM, CoWoS-class packaging, substrates and lithography tools are constrained by lead times measured in years, so they are likely to be the last to ease. Falling prices for commodity memory should therefore not be read as a sign that HBM, packaging or substrates have become easier to obtain.

Strategic implications. For procurement, the evidence favours securing multi-year memory volume now and redesigning boards around fewer, standard memory parts; waiting for relief is likely to mean paying spot prices into 2027. For investors, contract price momentum and supplier capex are better cycle signals than revenue, which is inflated by price. For policymakers, schemes that pick who gets scarce supply would mostly move the shortage from one industry to another, as the Information Technology and Innovation Foundation has argued; accelerating the qualification of new fabrication, packaging and substrate capacity would add supply.

Key indicators to monitor:

Memory contract prices. Whether quarterly increases in DRAM and NAND slow, hold or pick up again. Renewed momentum in early 2027 would confirm that tightness runs deep.

Lead times and sold-out horizons. Quoted waits for chips, substrates and tools, and how far ahead suppliers say their output is already sold. Longer horizons mean the squeeze is spreading.

New capacity coming online. First wafer output from new fabs, and packaging and substrate expansions. Any delay pushes relief back, since these are the real cure.

Signs of demand cooling. Falling phone and PC shipments, rising inventories, and AI buyers trimming spending because power and sites are not ready. Commodity memory would ease first. Advanced packaging would not.

Export controls and counter-measures. Fixed deadlines on chip and rare-earth curbs, new rules on remote access to AI computing, and congressional action. These decide who can buy, even if they do little to change how much is made.

The decision facing corporate buyers is whether to pay now for certainty of supply or to wait for relief that, on present evidence, arrives no earlier than the second half of 2027.

STONEBENCH

ABOUT STONEBENCH

StoneBench is a strategic editorial advisory and thought leadership studio specialising in topics at the intersection of economics, geopolitics, trade, finance, sustainability and policy. Our rigorous approach and deep understanding of key global trends have earned the trust of leading institutions, including banks, technology companies, multinationals, consulting firms, industry associations and policy advisory firms, for whom we shape thought leadership strategy and deliver end-to-end research and editorial programmes.