Deloitte.

2024 global semiconductor industry outlook

What's inside

About Deloitte's Outlooks

Deloitte's *2024 global semiconductor industry outlook* seeks to identify the strategic issues and opportunities for semiconductor companies and other parts of the semiconductor supply chain to consider in the coming year, including their impacts, key actions to take, and critical questions to ask. The goal is to help equip companies across the semiconductor supply chain with the information and foresight to better position themselves for a robust and resilient future.

Executive summary

Chip sales look to bounce back in 2024, led by generative AI, but could be complicated by geopolitics

The infamously cyclical semiconductor industry had a challenging year in 2023, the seventh downturn since 1990, with sales expected to be down 9.4% (to US$520 billion) for the year.[1] But that's not as bad as was expected in the spring; before relatively stronger second and third quarters were in the books, the previous forecast had been for US$515 billion.[2] 2024 is now predicted to see global sales of US$588 billion.[3] Not only would that be 13% better than 2023, but it's 2.5% higher than 2022's record industry revenues of US$574 billion.

The stock market is often a leading indicator of industry performance: As of mid-December 2023, the combined market capitalization of the top 10 global chip companies was US$3.4 trillion, up 74% from US$1.9 trillion in November 2022 and 17% higher than the US$2.9 trillion we saw in November 2021.[4]

As is so often the case, the memory chip market was the biggest swing factor. In 2022, memory sales were almost US$130 billion, or just under 23% of the overall chip market, but they dropped 31% (about US$40 billion) in 2023, compared to down 1% for logic. The market is expected to get almost all of that back in 2024, with sales expected to reach 2022 levels. If we exclude memory, the rest of the industry was down in 2023, but only by about 3%.[5]

In terms of end markets, both PC and smartphone sales are expected to grow 4% in 2024, after 2023 declines of 14% and 3.5%, respectively.[6] Returning to growth for these two end markets is likely important for the semi industry: In 2022, communication and computer chip sales (which include data center chips) made up 56% of overall semiconductor sales for the year, compared to auto and industrial, which accounted for only 14% of sales each, for example.[7]

Two other important measures of the industry's health are inventories and fab utilization. As of fall 2023, inventories remained high at more than US$60 billion, about the same level as the previous year. And the process of drawing those down will be a significant headwind for sales in the first half of 2024.[8] In addition, utilization was high during the recent shortage (in the mid-90% range) and is expected to fall below 70% in Q4 2023.[9] The industry likely needs utilization to be much higher than that to be profitable, which could take some time. Meanwhile, capacity is also growing as the United States and Europe increase domestic chipmaking.

These trends and others play into our *2024 global semiconductor industry outlook*, where we drill down into five big topics for the year ahead: generative AI accelerator chips and how semiconductor companies are using gen AI; trends around smart manufacturing; the need for more assembly and test capacity worldwide; how chip industry intellectual property (IP) is a target for cyberattacks at a whole new threat level; and a final geopolitics section that looks at export controls around advanced node manufacturing equipment and technologies, as well as advanced gen AI semiconductors.

There are some interesting angles to the gen AI chip boom that affect the chip industry as a whole. With sales predicted to reach more than US$50 billion in 2024,[10] this market is a tailwind for the sector and is expected to account for about 8.5% of sales. A portion of that will come from logic processors made on advanced nodes, some of it from advanced high bandwidth memory (HBM3), some from advanced 2.5D packaging, plus some from advanced connectivity chips. In each category, these gen AI-driven chips are among the priciest of their kind. In 2022, more than a trillion chips were sold at an average selling price of US$0.57 per chip.[11] Meanwhile, some gen AI chips were selling for US$40,000 each in 2023, or 70,000x higher, and therefore US$50 billion worth of chips might only be a volume of 1.25 million chips, or less than 0.1% of total chip volumes for the year.[12]

Why does that matter? Although gen AI chips are expected to be a big part of 2024 chip revenues (and likely to be even bigger in the future), they are a relatively small part of unit volumes, and therefore manufacturing capacity. As mentioned above, with industry utilization below 70%, having gen AI chips fly off the shelves is great news for the handful of companies that sell those chips or parts of those chips, but it may not be as helpful for the overall industry.

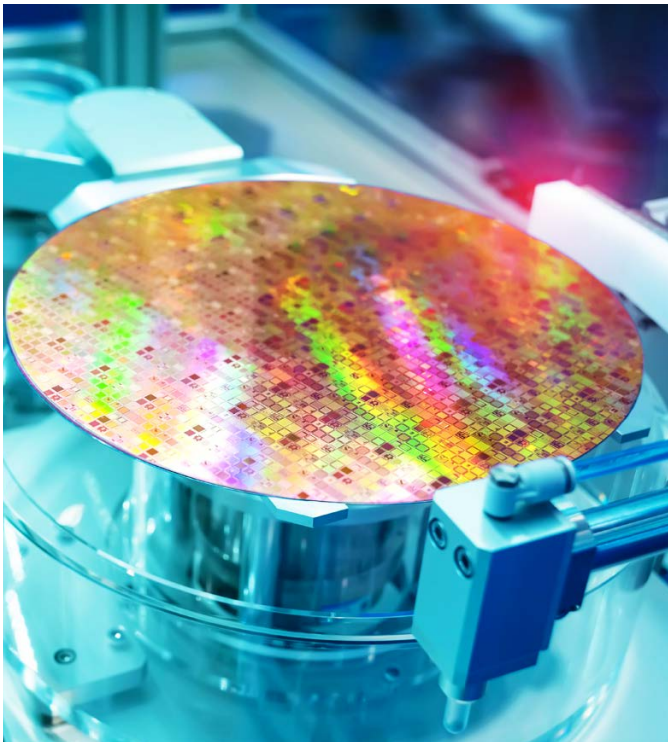
For the industry to see optimal utilization across all process nodes, other kinds of chips in addition to gen AI may need to see stronger demand.

Generative AI and the next wave of chip supply

Selling gen AI chips:
The market for chips that accelerate the training and inference of generative AI models was the semiconductor story of 2023. At a high level, gen AI chips are packages of special GPUs, special CPUs, special HBM3 in advanced 2.5D packaging, plus other special chips needed for connectivity in the data center.[13] For an industry fighting headwinds from weak memory prices, and weak demand for smartphone and computer chips, gen AI chips provided a growth area, especially at leading manufacturing nodes. As 2024 approaches, the market for these chips looks to be strong and is predicted to reach more than US$50 billion in sales for the year,[14] or 8.5% of the value of all chips expected to be sold for the year.[15]

In the longer term, there are forecasts suggesting that AI chips could reach US$400 billion in sales by 2027.[16] But what will happen in 2024? On the one hand, Deloitte predicted in November 2023 that gen AI chip sales will be more than US$50 billion in 2024.[17] On the other, there are reasons to believe quarterly gen AI chip sales growth could flatten or even decline at some point, at least for a while. Fall 2023 was a perfect storm: strong demand, buyers eager to secure supply, and relatively few choices. But this could change at some point in 2024.

1. **New entrants:** New chips are expected from existing gen AI chipmakers as well as emerging gen AI chipmakers. And new chips will likely come from companies that, up until now, have been better known as buyers of chips but have decided to start making their own. Nobody knows how any of these will fare at this point.
2. **New architectures and models**: There are many different gen AI models and approaches, and no single chip architecture is likely to be optimal for all use cases. There will likely be data center chips, edge chips, training chips, and inferencing chips, and it's expected that billions of dollars will be spent developing these "flavors" of gen AI chips.
3. **Edge chips:** It's possible that in 2024 more processing will be done at the edge, on smaller, cheaper, or different accelerators or gen AI models than the ones that dominated in 2023.
4. **The bullwhip effect:** The chip industry can be highly susceptible to over-ordering and excess inventories in times of undersupply, which then needs to unwind when supply catches up to demand, often leading to falling unit sales and falling prices per unit. This bullwhip effect[18] could be seen by the second half of 2024.
5. **Will people pay for gen AI?** Finally, much of the current demand for gen AI chips is coming from enterprise software companies either directly for those building on-premise processing, or indirectly via cloud companies that provide the software companies with gen AI processing services. They may plan on including gen AI features and services inside their existing or new software offerings and hope to charge for those features: Deloitte predicts this will provide a US$10 billion revenue uplift in enterprise software revenues by the end of 2024.[19] And they may be buying up gen AI chips now (either directly or via the cloud) to help meet anticipated processing demand. Should software buyers be unwilling to pay a premium for gen AI services, or even be slow about doing so, the software companies could abruptly reduce their orders for gen AI processing.



Using gen AI for making chips:

The chip industry has been using AI tools to help design chips for a couple of years, but that was just the beginning.[20] Gen AI can help improve operations and proliferate best practices throughout the semiconductor industry value chain: According to Deloitte's upcoming *Gen AI in Semiconductors Study* (2024), 72% of respondents believe that gen AI's impact on their industry will be "high to transformative."[21] Respondents narrowed that further, describing the following areas where gen AI might be used, in addition to chip design and code creation:

- Generating more accurate schedules and supply chain forecasts
- Enhancing research and development via research augmentation
- Improving anomaly and defect detection; existing AI solutions are already at use in the industry, but gen AI offers the potential to speed this up by creating synthetic data for model training
- Operations could benefit from manufacturing process simulation and gen AI-enabled digital twin generation that can simulate sorting, assembly, testing, and other complex manufacturing processes without the need for petabytes of data
- Enhancing sales and marketing efforts by using gen AI for better content generation, tailoring marketing content to audience and purpose

In spite of these benefits, gen AI still has its challenges. Building or buying custom models for generating circuit design, testing plans, and synthetic data can be cost prohibitive as compared to manual execution. Running costs for very large models can also outweigh the speed benefits of automation. As is well known, gen AI can create inaccurate or nonsensical outputs, so validation by humans-in-the-loop is needed to improve accuracy. Finally, when working with human-centered applications such as HR or sales and marketing, human review can help ensure that all data is sanitized, personally identifiable information (PII) removed, guardrails installed, and validation performed.

Strategic questions to consider:

- As gen AI chipmaking capacity increases, will prices and volumes continue to increase or come down? At a forecasted US$400 billion in 2027, gen AI chips would be close to half the chip industry; what would that mean? What early signals should be paid attention to from across the chip supply chain, such as movements in channel inventories or order buildups at hyperscalers?
- Many chip segments are dominated by a single player, whose preeminent role is both strong benchmark performance and a set of tools and support systems that lock buyers in. Will gen AI chips follow that trend, or will we see a more fragmented industry? To be prepared for the various scenarios, what changes will need to be made in supplier and vendor contracts, as well as order pipeline? Will there be aggressive M&A or private equity interest in smaller chip companies?
- Gen AI tools may transform the industry, but what will the effects be? Will they just lower costs, or will we actually see revenue growth?

Making smart manufacturing smarter

Over the years, semiconductor fab facilities and outsourced semiconductor assembly and test facilities (OSATs) have leveraged IoT devices, robotics tech, and artificial intelligence/machine learning (AI/ML) and analytics with the goal of achieving smart, lights-out chip factories that are fully automated.[22] Wafer fab equipment makers, integrated device manufacturers (IDMs), foundries, and back-end AT facilities all continue to invest more in smart manufacturing practices, digital tools, and technologies, but the basic smart manufacturing objectives more or less remain the same. They connect things on the factory floor, automate material movement and data collection, and apply analytics to prompt decisions and actions. However, from a chip manufacturing perspective, two things are expected to be different for smart manufacturing in the semi industry in 2024.

One is the availability of sophisticated and highly advanced AI tools (including gen AI) to analyze large datasets and offer sharp insights. Deloitte's upcoming *Gen AI in Semiconductors Study* noted roughly seven in 10 semi execs acknowledge gen AI will have a "high to transformative" impact on their business, indicating they view it as more than just another new tool.[23] Across manufacturing, operations, and maintenance functions, semi execs surveyed believe gen AI has the potential to add the greatest value through analyses and insights pertaining to processes and equipment (noted by 28% of respondents), followed by predictive maintenance and smart diagnostics and troubleshooting (18%).

The second relates to enhancing performance and sustainability of fabs and buildings. Moving from a mature technology node (e.g., 28 nm manufacturing) to an advanced node (2 nm) needs 3.5 times as much energy, consumes 2.3 times as much water, and emits 2.5 times as much greenhouse gases.[24] The semi industry should consider implementing manufacturing transformation on older plants (brownfields), in addition to building brand-new greenfield plants, to help realize even greater sustainability benefits.[25] Moreover, fabs can look at investing in smart manufacturing tools such as 6D BIM (building information modeling) to help improve cost management, simulate and analyze energy consumption, enhance efficiency, and streamline facility management—making a positive impact on the bottom line and the planet.[26]

As fabs explore using smart manufacturing tools and virtual models such as 6D BIM throughout 2024, they should identify specific environmental and social factors and metrics to be measured for a more holistic sustainability assessment and environmental, social, and governance (ESG) reporting. Additionally, companies should consider exploring where and how technologies and tools such as gen AI, private 5G networks, and digital twins could be integrated in the manufacturing plants.[27]

Semiconductor executives polled as part of the Deloitte *2023 Semiconductor Transformation Study* identified that lack of enterprise-wide alignment on digital capabilities is a key challenge as part of their transformation initiatives.[28] Though the level of business and IT alignment showed improvement in the 2023 survey (56%) compared with the prior 2021 survey (37%), it may need to be revisited in 2024.[29] Companies should continue investing in smart databases and systems to acquire disparate sets of data across business units and an increasingly complex and extensive supply chain. They should bolster capabilities in data modernization, unified data platforms, advanced analytics, and next-generation SaaS-based applications (e.g., procurement, planning, customer, and enterprise resource planning)—all of which may require better integration with their manufacturing execution system (MES), with emphasis on clean, accurate, and good quality data. And these aspects are vital to help enable companies to benefit fully from their prior investments made in AI/ML and data management solutions.

Even as fabs and AT facilities increase the degree of automation and use smart technologies, they should look to determine what specialized skills their workers and operators should possess—especially those who oversee the machinery, equipment, processes, and control systems. 2024 could be the year when the semi industry takes stock of where and how the level of human involvement can be limited—at least for some period of time in select aspects of the fab operations—to truly achieve "smart" lights-out factories that operate efficiently in the dark.

Strategic questions to consider:

- As part of a digital vision, what aspects should be considered by an enterprise—including design, production, sales, support, and product end-of-life/recycling—to make it a holistic digital approach?
- To make the variety of manufacturing equipment, machinery, and tools in the facilities talk to each other and benefit from the broader data modernization efforts across the other business units (e.g., in supply chain, finance, procurement, sales), what other line-of-business applications should the production systems and processes be integrated with? What communication protocols should be standardized? And what interface programs should be developed?
- In view of the goal to be lights-out, how can the next-generation smart factory implementation be more human-centric, and what specific aspects of manufacturing processes should have human workers doing the job alongside machines? Additionally, how can companies integrate worker well-being to make the semi industry not only sustainable but more human-centric?

Semiconductor assembly and test venture into new geographies

More than 75% of the global semiconductor fab capacity is in Asia (the front-end),[30] but the region's market share is even higher (90%) in chip assembly and testing (the back-end).[31] Except for large IDMs, most chip players have been outsourcing AT processes to third-party vendors, or OSATs.[32] The majority of the big OSATs are based in China and Taiwan, commanding roughly 80% of the OSAT market share in 2022. Although the United States is aiming to bolster domestic AT capacity,[33] almost all actual AT work is done in Asia.[34] The lines between traditional front-end and back-end are increasingly blurring, with each attempting to capture more of the value chain. Advanced packaging is also increasingly becoming a strategic enabler to build the most sophisticated leading-edge chips.[35]

Going forward, as the United States and Europe look to expand domestic chip fabrication capacity, they should look to build up their back-end capacity to avoid lengthening and making their supply chains more complex. To help stay on the leading edge of product performance and flexibility, IDMs in the United States and South Korea are increasing efforts to bolster their packaging capabilities, which are usually provided and enabled by their respective assembly operations and facilities. Concurrently, leading fabless companies are pushing for nearshore AT. Further, complex gen AI chips are fueling demand for advanced packaging, exposing an acute capacity shortage for this technology.[36]

Deloitte's *2023 Semiconductor Transformation Study* noted that strong demand for AI and compute- and memory-intensive applications will spur innovation in advanced packaging techniques such as 2D, 2.5D, and 3D.[37] To keep pace with the growing packaging innovation needs, both the EU Chips Act and the US CHIPS and Science Act have funds allocated for advanced packaging technology development.[38] However, as noted in the 2024 Prediction *Gen AI chip demand fans a semi tailwind, for now*, governments may need to add to existing incentive programs to expand domestic or nearshore AT manufacturing capacity.[39] Packaging is becoming more complex, often requiring new techniques such as data modeling and simulations to proactively help discover and predict errors and other packaging issues in advance.

In 2024, the back-end AT market could experience significant transformation, as prominent IDMs and foundries move even further into advanced packaging, while traditional OSATs also continue to enhance their packaging capabilities. Simultaneously, US- and EU-based semiconductor companies are expanding their front-end wafer fab facilities on their home turfs.[40] Alongside this expansion, steps are being taken to shift their back-end AT services to new countries. For instance, new AT capacity is being built in Vietnam, Malaysia, India, and Poland,[41] reflecting how IDMs and OSATs are diversifying and de-risking their supply chain; this trend is in line with Deloitte's perspective in the *2023 global semiconductor industry outlook*.[42]

But the emerging AT facilities face distinct challenges. New advanced packaging technologies and test solutions should be delivered with high-quality performance within stringent time-to-market constraints. Also, such technologies often require distinct skills and experiences. For example, packaging and testing engineers need to have specializations in electrical and electronics engineering, material sciences, capacity planning, and yield processes.[43] Additionally, back-end players are challenged to provide a range of novel but intricate advanced packaging options; for example: 2.5D/3D, fan-outs, chiplets, SiP, and hybrid bonding.[44]

In 2024, IDM AT units and pure-play OSATs could look to shortlist from those several options, and gain mastery of specific packaging technologies. They should be agile and constantly innovate to help allow branded semiconductor companies to launch superior products more rapidly and at competitive performances and prices. One other aspect that AT facilities should consider is the energy, materials, and other resources used in assembly, test, shipping, and distribution operations—which are often equally important parts of the semiconductor sustainability equation.[45]

To stay competitive in the dynamic AT landscape throughout 2024 and beyond, OSATs and captive AT facilities should strengthen their core enterprise IT systems. Additionally, integrating AI and ML into their operations can help develop advanced packaging technologies and features, improve demand planning, manage inventory effectively, and streamline information flow across the extended supply chain. Testing is also expected to gain prominence, as complex chip and module designs could require captive AT and OSATs to advance capabilities like system-level test, adaptive or dynamic test, and AI/ML-based bin prediction.

Strategic questions to consider:

- Which peer companies and academic institutions should companies partner with to take advantage of government incentives and shared research infrastructure to collaboratively build, test, and pilot next-generation innovative packaging approaches?
- What parts of AT processes should be reshored domestically, which ones should be outsourced, and what countries (within Southeast Asia, South Asia, Eastern Europe, Latin America) should be considered to build new back-end facilities?
- In back-end assembly, how should metals, plastics, and other materials be handled as part of broader sustainability efforts and to comply with ESG regulatory needs?
- Should more foundries and IDMs integrate packaging into their value chain? Will the increased strategic relevance of packaging manifest itself through new alliances and partnerships that fabless and IDMs could consider?

Bolstering cybersecurity to combat intensifying cyberthreats

The semiconductor industry faces a different level of cyberthreats compared to other industries. In addition to the usual profit-seeking ransomware attacks that every industry deals with, semi companies possess unique, valuable, and restricted IP. Due to the increasing importance of semiconductors for multiple industries, it's often targeted by state-backed actors. As a result of geopolitical issues and restrictions on advanced chipmaking tech, the IP of semi companies is one of the world's most important targets for cyberattacks.[46]

If geopolitical tensions continue to escalate in 2024—resulting in further restrictions around IP, chips, and raw materials—cyberattacks may intensify, disrupting production in the industry.

Cyberthreat actors are not only targeting the core semi companies, but attacks are increasingly directed at extended channel partners, too (e.g., suppliers, distributors, and contract manufacturers).[47] Sophisticated threat actors could use advanced methods, like masking as ransomware groups,[48] to fabricate attacks and cause business disruptions across the supply chain.

Here's where the semiconductor industry could experience an asymmetric battle in 2024: The industry may need to face off with sophisticated threat actors that have much more advanced resources than other industrial cyberthreats. This asymmetric battle between state-backed actors and broader semi industry participants could present new challenges to industry executives in 2024 and beyond.

Therefore, in 2024, semi companies should consider accelerating their efforts to bolster cyber defense capabilities of their own as well as the digital and cyber infrastructure of their extended supply chains.

Companies should be more agile in realigning their cyber programs with these emerging threats and invest in customer data protection solutions. For instance, Deloitte's *2023 Global Future of Cyber Survey* noted that more than 75% of respondents believed that investing in cybersecurity strategy, cyber cloud, and data protection solutions generated the highest value for their organizations.[49] In addition, semi companies should look at using artificial intelligence to help combat cyberthreats: AI can not only enable cybersecurity teams to respond faster than the pace of cyberattackers, but using an AI-based and data-driven approach could also help sense and anticipate threats and counter cyberattacks proactively.[50]

Apart from bolstering their own IT infrastructure, semi companies should help suppliers across regions bolster their cyber defenses as well. This could allow the broader semiconductor ecosystem to operate with a greater degree of assurance and resilience, even in high-risk regions. This may be especially critical as semi companies' manufacturing and supply facilities will likely be far more widespread across the globe, as we wrote in Deloitte's 2023 global semiconductor industry outlook.[51] Besides, advanced cyber solutions can provide a secured digital environment for fabs, foundries, equipment vendors, back-end facilities, and suppliers to interact and exchange real-time information. As a start, they should work together to set up security standards and processes to prevent data leaks and mitigate breaches.

Strategic questions to consider:

- Given the regional differences in cybersecurity threats and incidents, what strategies will help improve resilience and enable digital trust for organizations and suppliers globally? How can companies tailor strategies that best address cybersecurity needs and threat vectors that are specific to each of the regions in which they operate? Can a shift to smart manufacturing help chipmakers become more cyber resilient, in addition to its other benefits?
- What types of distinct methods should be considered when compartmentalizing IP and crown jewels using security systems and access management?
- What holistic skills and capabilities—such as cyber forensics, cryptography, AI/deep learning, DevSecOps—do companies need, in order to develop a robust organization-wide cybersecurity program?

Geopolitics, export controls, semiconductors, advanced nodes, and AI

In recent years, US policymakers have enacted export controls on various types of semiconductors and semiconductor manufacturing technologies to China, many of which are restricted because policymakers argue that they have a military application.[52] A handful of other countries have enacted similar restrictions with respect to China.[53] What was relatively new in 2022–23, and may be increasingly important in 2024, is the perception around two key semiconductor technologies: advanced node manufacturing itself and chips that accelerate AI. Both of these technologies were restricted to some extent in 2022, with restrictions tightening further in October 2023.[54]

We expect that 2024 will see the ramifications of those restrictions, as well as possible further restrictions. Companies are expected to try to balance making sales while still complying with ever-evolving restrictions, possible counter-restrictions, and new developments in domestic advanced node capability in the countries impacted by these restrictions. They should not only comply with the current situation, but increasingly plan for or attempt to anticipate future rules and restrictions.

Advanced node manufacturing restrictions:
China has had a series of government initiatives and invested billions to become self-sufficient (or at least more self-sufficient) in chips for a decade now.[55] China is a major player in trailing node manufacturing (approximately 50–180 nm) and intermediate node (14–45 nm), with about 25% of global capacity—and growing.[56] However, for advanced node manufacturing (10 nm and below, with current leading-edge chips being made at 3 nm in 2023), it's generally believed that extreme ultraviolet lithography (EUV) is essential for making these chips.[57]

Many of the US restrictions have been around deep ultraviolet (DUV) and EUV machines and supporting technologies, which were believed to be an effective barrier to making advanced node chips.[58] In the summer of 2023, one of China's leading chipmakers was reportedly producing 7 nm chips for a high-end smartphone brand—possibly using pre-restriction advanced DUV lithography tools with multiple patterning and precision alignment.[59] It may be possible to make a 5 nm chip using these approaches, but it is predicted to be expensive, slow, and likely to affect yield negatively.[60]

Gen AI chip restrictions:
The fastest-growing market in semiconductor sales in 2024 is expected to be chips designed to accelerate gen AI workloads: Deloitte predicts that this market will reach more than US$50 billion for the year, representing close to 10% of all global chip sales for the full year, up from zero in 2022.[61] The United States is restricting the export of these chips to select locations, but those restrictions have been rapidly evolving.[62] To be clear, restrictions are in general about high performance computing chips of all kinds, not just gen AI.

So far, the United States' approach is to subject chips with certain features to export restrictions. In 2022, export controls were mainly around the maximum interconnect speeds, which meant that some gen AI chips were restricted, but some older gen AI chips, as well as newer chips designed with lower interconnect speeds, were able to be exported.[63] But the United States further tightened the rules in October 2023, and new ceilings on performance density—specifically FLOPS/mm²—were imposed, resulting in several gen AI chips that had previously been unrestricted becoming subject to export controls.[64] Further evolving tensions between government restrictions and industry's desire to sell chips are likely.

A consideration for these ongoing and evolving restrictions is that some of the restricted destinations are large chip and equipment markets. Like export restrictions, which could substantially affect chip and equipment company revenues. This leaves the industry with a key question to consider: Even if restricting chips or chipmaking technologies achieves short-term geopolitical goals, does it have the long-term effect of leading restricted countries to develop their own solutions and become self-sufficient in these areas? Equally, increasing restrictions on chips or chipmaking technologies could lead to China or other countries introducing their own export restrictions for older generation chips and important raw materials.[65]

In recent news, the US Commerce Department announced in late December 2023 that it will conduct a survey on US supply chains of the current generation but *also mature-node or legacy chips* in 2024.[66] These new findings could potentially result in a significant shift, especially for industries that are heavy users of these chips, such as automotive and defense.

Strategic questions to consider:

- Can countries under export restrictions on DUV and EUV technology develop their own DUV and EUV solutions? If not, if they make advanced chips using DUV-based techniques, can they do so economically?
- Will restrictions on the export of gen AI chips themselves continue to tighten? How will this affect chip company revenues and growth?
- What will be the findings and ramifications of the new proposed survey around US procurement of legacy chips?
- Will these restrictions lead to responses from restricted countries that could adversely affect companies, industries, and the global economy?

Signposts for the future

For 2024, semiconductor industry executives should be mindful of the following signposts.

Inventory management amid fluid demand situation:
Demand from traditional end markets may remain muted, but could specialized chips demand become even stronger in 2024? Semi companies should manage inventory levels in tandem with these dynamic and varying end-market demand drivers—including creating buffer stocks for some product categories while making room to re-negotiate contract terms on orders for others. Semi companies should look to become ecosystem drivers and orchestrators; for instance, by establishing more direct lines of communication and foresight with end customers to gain greater visibility into demand and strengthening their engineering relationships.

Opportunities for strategic M&A:
With a range of niche players and startups across semiconductor design, gen AI chips, advanced materials and components, and advanced packaging and test, the semi companies should plan for and pursue targeted M&A deals and look for opportunities to buy attractive assets, perhaps alongside or in competition with private equity investors.

Potential adverse impact on channel partner relationships:
As geopolitical issues escalate and Western countries' relationships with China continue to evolve in 2024, semi companies—as part of their strategies to build resilient supply chains—should keep close tabs on alternate sourcing locations and periodically review their contracts with global suppliers and distribution channel partners across key regions and countries. Additionally, they may need to assess taking some customers direct, where channel relationships are potentially disrupted by geopolitical tensions.

Tax and regulatory aspects in cross-border deals:
While the United States and European Union are aiming to further strengthen their domestic and regional chip manufacturing capabilities through onshoring, friendshoring, and nearshoring, chip companies should bear in mind that cross-border M&A deals and strategic business investments in these regions could present more stringent tax and regulatory considerations.

Interest rate movements and valuation multiples:
When funding growth and expansion in 2024, semi companies should be mindful of preserving cash for future investments as well. They should especially consider the impact from interest rate changes when raising debt and assess valuation multiples when targeting niche players spun off from private equity and venture capital exits or corporate divestitures.

Talent challenges that would demand agility:
As the talent competition continues and permeates across the semi supply chain in 2024, what alternative talent choices should industry execs look for? For instance, sourcing talent from friendshore regions, recruiting skilled immigrant workforce, upskilling and cross-skilling in-house staff, hiring gig workers, or even joining hands with startups and accelerators. Gen AI may help address some talent shortages but, in turn, may also create one, as chip companies look to hire those with gen AI skill sets in a highly competitive market.

Contacts

Christie Simons
Global Semiconductor
CoE Leader
csimons@deloitte.com

Brandon Kulik
US Semiconductor Leader
bkulik@deloitte.com

Jan Nicholas
APAC Semiconductor Leader
jnicholas@deloitte.com

Jordan Bish
EMEA Semiconductor Leader
jbish@deloitte.nl

Duncan Stewart
TMT Director of Research
dunstewart@deloitte.ca

Karthik Ramachandran
Senior Research Manager
karramachandran@deloitte.com

Christie Simons, Brandon Kulik, Jan Nicholas, and Jordan Bish would like to thank Duncan Stewart and Karthik Ramachandran from Deloitte's Center for Technology, Media & Telecommunications (TMT Center) for their contributions to the research and writing of this outlook. They would also like to thank Dr. Bobby Mitra, Uwe Lambrette, Dan Hamling, Jeroen Kusters, Upen Sachdev, Meer Hussain, Karthik Amirtharajan, and Mark LaViolette for their contributions and insights.

About the TMT Center
Deloitte's Center for Technology, Media & Telecommunications (TMT Center) conducts research and develops insights to help business leaders see their options more clearly. Beneath the surface of new technologies and trends, the TMT Center's research can help executives simplify complex business issues and frame smart questions. The TMT Center can help executives better discern risk and reward, capture opportunities, and solve tough challenges amid the rapidly evolving TMT landscape.

Endnotes

1. Worldwide Semiconductor Trade Statistics (WSTS), "WSTS semiconductor market forecast – Fall 2023," press release, dated November 28, 2023.
2. Ibid.
3. Ibid.
4. Deloitte analysis of public market capitalizations in 2023, 2022, and 2021.
5. Deloitte analysis of WSTS, "WSTS semiconductor market forecast – Fall 2023."
6. International Data Corporation (IDC), "Global PC shipments expected to return to growth in 2024 albeit below 2019 pre-pandemic levels, according to IDC," press release, August 28, 2023; IDC, "Worldwide smartphone shipments finally turn the corner with 7.3% growth forecast in Q4 2023 and 3.8% in 2024, according to IDC Mobile Phone Tracker," press release, November 29, 2023.
7. Robert Casanova, "Chip sales rise in 2022, especially to auto, industrial, consumer markets," Semiconductor Industry Association (SIA), March 27, 2023.
8. SEMI, "Global semiconductor manufacturing industry set for Q4 2023 recovery, SEMI reports," press release, November 13, 2023.
9. Ibid.
10. Duncan Stewart et al., "Gen AI chip demand fans a semi tailwind … for now," Deloitte Insights, November 29, 2023.
11. Semiconductor Industry Association, "Global semiconductor sales, units shipped reach all-time highs in 2021 as industry ramps up production amid shortage," press release, February 14, 2022.
12. Kif Leswing, "Nvidia's top A.I. chips are selling for more than $40,000 on eBay," CNBC, April 14, 2023; Deloitte analysis.
13. Stewart et al., "Gen AI chip demand fans a semi tailwind … for now."
14. Ibid.
15. WSTS, "WSTS semiconductor market forecast fall 2023."
16. Analysis based on data sourced from multiple publicly available sources: Martin Baccardax, "Nvidia jumps higher as Mizuho analysts see $300 billion AI chip potential," *TheStreet*, July 24, 2023; Patrick Seitz, "Intel on track with data center chip lineup, touts play in artificial intelligence," *Investor's Business Daily*, March 30, 2023; WSTS, "WSTS semiconductor market forecast spring 2023"; Deborah Yao, "Analysts' take: Nvidia widens its total addressable market," *AI Business*, May 31, 2023.
17. "Gen AI chip demand fans a semi tailwind … for now."
18. Dan Hamling, Chris Richard, Duncan Stewart, Karthik Ramachandran, "Five fixes for the semiconductor chip shortage," Deloitte Insights, December 6, 2021.
19. Duncan Stewart et al., "Generative AI and enterprise software: What's the revenue uplift potential?," Deloitte Insights, November 29, 2023.
20. "Gen AI chip demand fans a semi tailwind … for now."
21. Deloitte, *2023 Smart Manufacturing: Gen AI for Semiconductors Survey* (to be published in 2024).
22. Anne Meixner, "Smart manufacturing makes gains in chip industry," *Semiconductor Engineering*, June 15, 2023.
23. Deloitte, *2023 Smart Manufacturing: Gen AI for Semiconductors Survey* (to be published in 2024).
24. Marie Garcia Bardon and Bertrand Parvais, "The environmental footprint of logic CMOS technologies," *EE Times*, December 14, 2020.
25. Duncan Stewart et al., "Semiconductor sustainability: Chips take a smaller byte out of resources," Deloitte Insights, November 29, 2023.
26. *The BIM Engineers Blog*, "From 3D BIM to 7D BIM," June 8, 2023.
27. Deloitte, Smart factory for smart manufacturing, accessed January 8, 2024.
28. Brandon Kulik et al., "2023 Semiconductor Transformation Study: Business transformation in the semiconductor industry," Deloitte, November 2023.
29. In particular, the 2023 semiconductor transformation study finds that only 12% of executive respondents believe full process automation currently exists, reflecting there's a general lack of organization-wide integration of digital proficiency—digital capabilities spanning the entire product life cycle from design to manufacturing—during transformation initiatives. To read further, see: "2023 Semiconductor Transformation Study," 2023.
30. Brian Lawson et al., "The shifting global semiconductor landscape in Asia-Pacific," *S&P Global Market Intelligence Blog*, February 10, 2023.
31. SIA, *2022 state of the U.S. semiconductor industry*, November 2022, figure, p. 21.
32. IDC, "IDC: Worldwide semiconductor OSAT market grew 5.1% YoY in 2022, growth expected in 2024 due to accumulated advanced OSAT demand," press release, July 25, 2023.
33. Mackenzie Hawkins, "US launches $3 billion effort to boost advanced chip packaging," Bloomberg, November 20, 2023.
34. IDC, "IDC: Worldwide semiconductor OSAT market grew 5.1% YoY in 2022, growth expected in 2024 due to accumulated advanced OSAT demand."

35. Matt Borden, "Why advanced packaging is vital to the future of semiconductors," *Microcontroller Tips*, October 31, 2023.
36. Dylan Patel, Myron Xie, and Gerald Wong, "AI capacity constraints – CoWoS and HBM supply chain," *SemiAnalysis*, July 5, 2023.
37. Kulik et al., "2023 Semiconductor Transformation Study."
38. H.R. 4346 – Chips and Science Act, 117th Congress (2021–2022); European Commission, "Statement by President von der Leyen on the European Chips Act," press release, February 8, 2022.
39. "Gen AI chip demand fans a semi tailwind ... for now."
40. Michelle Adams, "Where are all the new semiconductor fabs in North America & Europe?," Z2Data, September 12, 2023.
41. Based on Deloitte's research and analysis of publicly available information about select major semiconductor companies including IDMs and OSATs.
42. Christie Simons et al., *2023 global semiconductor industry outlook*, Deloitte, January 2023, p. 6.
43. Nidhi Singal, "Here's how India can fill the semiconductor talent gap for the world," *Business Today*, July 28, 2023.
44. Mark Lapedus, "Advanced packaging's next wave," *Semiconductor Engineering*, May 20, 2021. Notes on select advanced packaging technologies discussed in this article: 2.5D/3D stacks logic and memory dies in one package; fan-out integrates logic and more memory cubes; chiplet enables modular die-to-die interconnects; SiP (system in package) integrates multiple components like antennas, dies, power ICs, and RF modules into a single package.
45. Stewart et al, "Semiconductor sustainability: Chips take a smaller byte out of resources."
46. Based on our analysis of publicly available information related to cyberattacks on select major semiconductor companies during 2023. To read further, see: Insikt Group, "Semiconductor companies targeted by ransomware," Recorded Future, September 29, 2022; *Hacker News*, "Chinese hackers target semiconductor firms in East Asia with cobalt strike," October 6, 2023.
47. At least 50% critical security threats and issues are believed to originate from suppliers. To read further, see: Anne Meixner, "Securing chip manufacturing against growing cyber threats," *Semiconductor Engineering*, August 17, 2023.
48. Insikt Group, "Semiconductor companies targeted by ransomware"; *Hacker News*, "Chinese hackers target semiconductor firms in East Asia with cobalt strike."
49. Kevvie Fowler et al., "Cybersecurity threats and incidents differ by region," Deloitte Insights, 2023.
50. Ed Bowen et al., "Cyber AI: Real defense," Deloitte Insights, December 7, 2021.
51. Simons et al., *2023 global semiconductor industry outlook*. Moreover, an estimated 70+ front-end wafer fabs spanning multiple process tech nodes are coming up between 2024–26, with more than half of them in Asia and roughly one-third in the United States. To read further, see: Michelle Adams, "9 key statistics on new semiconductor fabs being built around the world," Z2Data, September 22, 2023.
52. Bureau of Industry and Security (BIS), US Department of Commerce, "Commerce implements new export controls on advanced computing and semiconductor manufacturing items to the People's Republic of China (PRC)," press release, October 7, 2022.
53. Japanese Ministry of Economy, Trade and Industry, "Press conference by Minister Nishimura (Excerpt)," news release, March 31, 2023; Government of the Netherlands, "Government publishes additional export measures for advanced semiconductor manufacturing equipment," news release, June 30, 2023.
54. BIS, "Commerce strengthens restrictions on advanced computing semiconductors, semiconductor manufacturing equipment, and supercomputing items to countries of concern," press release, October 17, 2023.
55. Chris Arkenberg, "China inside: Chinese semiconductors will power artificial intelligence," *Technology, Media, and Telecommunications Predictions 2019*, Deloitte, 2018, pp. 86–96.
56. TrendForce, "Despite export ban on equipment, China's semiconductor expansion in mature processes remains strong, says TrendForce," press release, July 6, 2023.
57. Nilay Patel, "Inside the global battle over chip manufacturing," *The Verge*, January 31, 2023.
58. Ibid.
59. Jeff Pao, "SMIC bypasses US curbs to make 7nm chips," *Asia Times*, September 5, 2023.
60. Gu Erde, "Has China's chip sector reached the end of the line?," *ThinkChina*, October 10, 2023.
61. "Gen AI chip demand fans a semi tailwind ... for now."
62. BIS, "Commerce strengthens restrictions on advanced computing semiconductors, semiconductor manufacturing equipment, and supercomputing items to countries of concern."
63. Dylan Patel et al., "Wafer wars: Deciphering latest restrictions on AI and semiconductor manufacturing," *SemiAnalysis*, October 24, 2023.
64. Karen Freifeld, "Exclusive: US tackles loopholes in curbs on AI chip exports to China," Reuters, October 16, 2023.
65. Duncan Stewart et al., "A raw deal: Will materials shortages, supply chain challenges threaten tech's future?," Deloitte Insights, November 29, 2023.
66. Office of Public Affairs, US Department of Commerce, "Commerce Department announces industrial base survey of American semiconductor supply chain," press release, December 21, 2023.

Deloitte.

About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited (DTTL), its global network of member firms, and their related entities (collectively, the "Deloitte organization"). DTTL (also referred to as "Deloitte Global") and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see www.deloitte.com/about to learn more.

Deloitte provides industry-leading audit and assurance, tax and legal, consulting, financial advisory, and risk advisory services to nearly 90% of the Fortune Global 500® and thousands of private companies. Our people deliver measurable and lasting results that help reinforce public trust in capital markets, enable clients to transform and thrive, and lead the way toward a stronger economy, a more equitable society, and a sustainable world. Building on its 175-plus year history, Deloitte spans more than 150 countries and territories. Learn how Deloitte's approximately 457,000 people worldwide make an impact that matters at www.deloitte.com

8352803