

Don't Forget about Memory

Brian Shirley

**STEPS: SCIENCE, TECHNOLOGY,
ENGINEERING, AND POLICY STUDIES**

ISSUE 7, 2022

STEPS (Print) ISSN 2158-3854
STEPS (Online) ISSN 2153-3679

Brian Shirley. "Don't Forget about Memory,"
STEPS 7 (2022): 30-42.



POTOMAC INSTITUTE PRESS

Copyright © 2022 by Potomac Institute for Policy Studies

STEPS: Science, Technology, Engineering, and Policy Studies
is published by Potomac Institute Press of the
Potomac Institute for Policy Studies.

Disclaimers: The Publisher, Institute and Editors cannot be held responsible for errors or any consequences arising from the use of information contained in this publication; the view and opinions expressed do not necessarily reflect those of the Publisher, Institute and Editors. The Potomac Institute is non-partisan and does not take part in partisan political agendas.

Copyright Notice: It is a condition of publication that articles submitted to this magazine have not been published and will not be simultaneously submitted or published elsewhere. By submitting an article, the authors agree that the copyright for their article is transferred to the Potomac Institute Press if and when the article is accepted for publication. The copyright covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microfilm, or any other reproductions of similar nature and translations.

Access to *STEPS* is available free online at:
www.potomacinstitute.org/steps.




DON'T FORGET ABOUT MEMORY

BRIAN SHIRLEY

**MEMBER, BOARD OF REGENTS
POTOMAC INSTITUTE FOR POLICY STUDIES**

Article first published online as the white paper: *Don't Forget about Memory*



Introduction

Semiconductors come in various flavors—logic, analog, sensors, memory, and others. While all play important roles in modern electronic systems, the insatiable demands of today's aptly named "data economy" make memory one of the most critical. Everything from data centers to communication networks to cellphones are dependent on large amounts of memory to safely store the information on which we rely. And while ongoing shortages of critical components remind us of the importance semiconductors play in our consumer products, we cannot take for granted the critical memory devices that make our world of instant information possible.

The US semiconductor Industry is at an urgent inflection point. Geopolitical tensions are shining a light on the growing risk from decades of declines in US semiconductor manufacturing, eroding from 37% of worldwide capacity in 1990 to just 12% today.¹ Efforts are underway to explore how the United States can revitalize growth in onshore, sustainable semiconductor production, using the same kinds of incentives that continue to tilt the playing field towards Asia.²

Semiconductor memory has been at the center of this maelstrom for decades. While less visible than the semiconductor logic sector, reliable access to secure memory technology is critical to the US economy and national security.

A glance at nearly any printed circuit board will reveal a sea of memory packages, both DRAM (traditional "fast" main memory) and NAND ("storage" memory). Each package often contains multiple stacked chips, most often making memory the largest ingredient of system silicon. For example, in an average cell phone, over half of the total silicon area is memory.³ Data center servers now resemble large memory buffers—leading-edge machines can hold the equivalent of nearly three full 300mm wafers of DRAM and more than five of NAND.⁴ And memory requirements continue to increase rapidly.

For example, artificial intelligence (AI) applications rely on massive stores of captured data residing in memory to hunt for the most critical patterns. Larger databases, faster networks, more powerful sensors—innovations at the core and edge—all are part of the trend demanding a need for more memory.

However, as the memory industry continues to grow in importance, it is also escalating in risk. As one of the first semiconductor segments to globalize, it has experienced a history of vicious economic cycles and spectacular bankruptcies, where suppliers often find themselves competing against entire countries. Decades of asymmetric nation-state subsidization combined with skyrocketing capital and research and development (R&D) requirements have left a dwindling number of suppliers, all with challenging margins and a manufacturing base located almost exclusively in Asia. China is now the latest in a long list of Asian countries to subsidize memory production: The “Made-in-China 2025” initiative specifically named domestic semiconductor production as a top priority⁵ with a goal of protecting national security.⁶ Initial efforts have focused on memory, and the emergence of two well-funded Chinese memory suppliers in 2019 demonstrates they are serious.⁷

Accordingly, the challenges of the memory industry present an urgent risk to US national security and well-being.

Today’s Memory Landscape

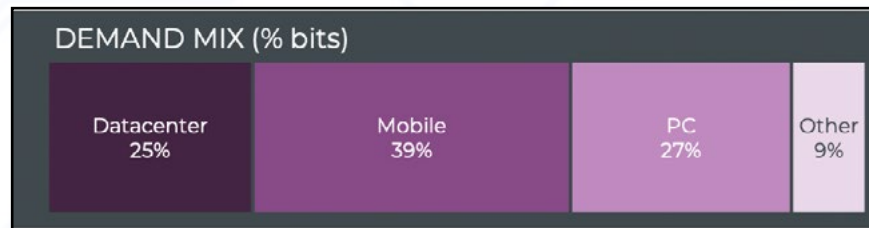
The memory industry accounts for \$154B in sales in 2021, comprising 28% of the global \$556B semiconductor market, and is equivalent in size to the entire category of logic (comprising CPUs, accelerators, FPGAs, etc.).⁸ However, memory drives a disproportionate 34% of the industry capital expenditures,⁹ yet with lower average margins than the logic industry.^{10,11} The vast bulk of memory revenue is driven by product sales of DRAM and NAND into data centers, cellphones, and PCs, as illustrated in Figures 1 and 2.^{12,13}

Five decades of intense, globalized competition and asymmetric state-sponsored subsidies have led to a hyper-competitive oligopoly (see **History of the Semiconductor Memory Industry** sidebar). Today, Samsung of South Korea commands roughly double the market share of the two next largest suppliers: SK Hynix (also of South Korea) and Micron Technology (of the US). These three companies together comprise 94% of DRAM revenue and 60% of NAND revenue.¹⁴ The notables in the remainder are all single-technology NAND providers: Kioxia Memory of Japan (spun out from Toshiba Corporation in 2018); Western Digital (arising from their 2015 purchase of SanDisk and sourcing NAND from the same fabs as Kioxia in Japan); and Solidigm, the previous NAND operation of Intel (recently acquired by SK Hynix, however not yet consolidated, with fab operations in Dalian, China¹⁵). The most recent entrants of note are NAND-focused Yangtze Memory Technologies (YMTC), in production with competitive 128-layer 3D NAND,¹⁶ as well as the DRAM-focused ChangXin Memory Technologies (CXMT),¹⁷ both funded by a mixture of Chinese private equity and government sources including “The Big Fund.”^{18,19}

The concentration of semiconductor fabs in Asia is even more extreme for memory than other semiconductor segments (see Figure 3). Not a single advanced DRAM or NAND fab remains outside Asia due to ongoing national subsidization, expansions to existing facilities, a stable utility infrastructure, and other factors. Samsung and SK Hynix operate fabs in South Korea and China, producing memory used by their own system products such as cellphones, as well as for merchant sales. Even the advanced fabs of pure-play memory supplier Micron are exclusively located in Asia, principally by virtue of acquisition, joint ventures, and expansion over the years.²⁰ The lone 300mm memory fab on US soil is operated by Micron in Manassas, Virginia, and is focused on production of older-generation devices. These more mature devices are of growing interest to markets such as defense, automotive, industrial, and medical, that no longer need the most advanced, highest density devices.²¹

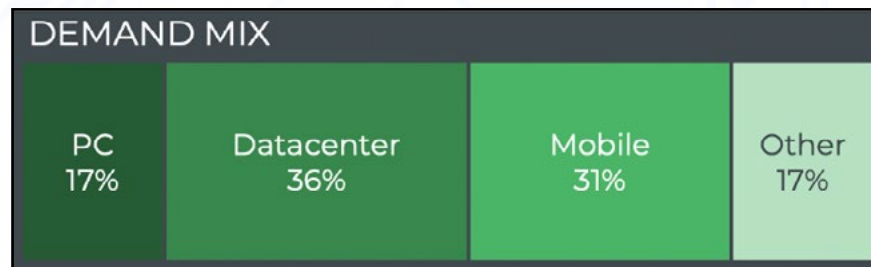
Continued page 37.

Figure 1. NAND Market Segments



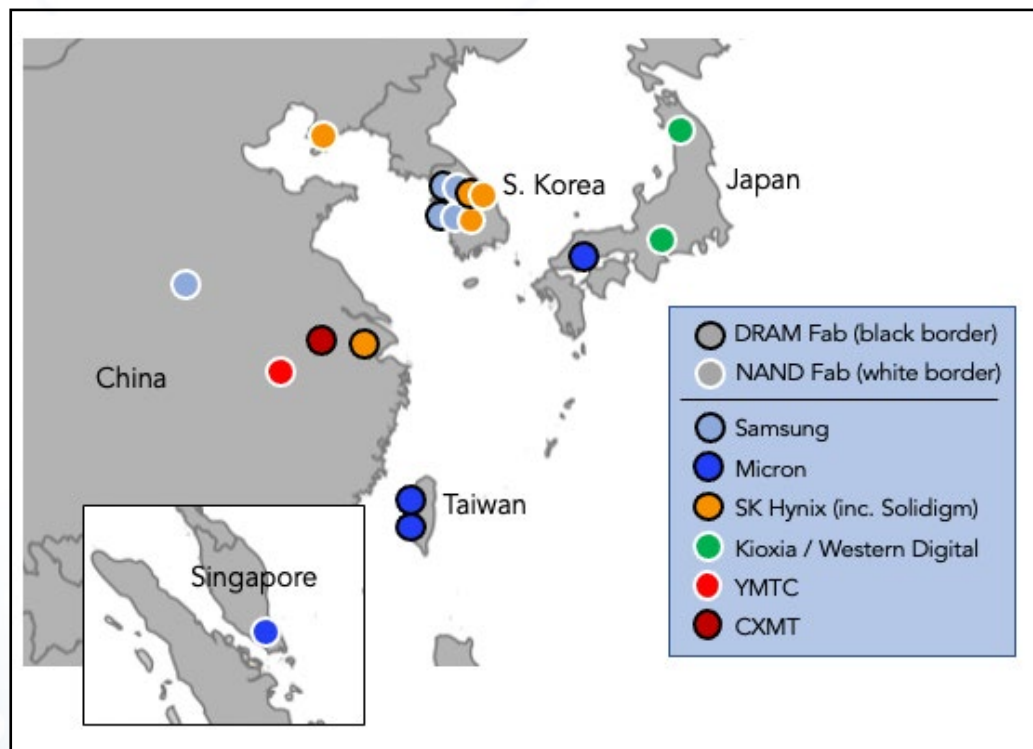
NAND Market Monitor Q2 2022. www.yolegroup.com © Yole Développement 2022
Reprinted here with permission from Yole Group.

Figure 2. DRAM Market Segments



DRAM Market Monitor Q2 2022. www.yolegroup.com © Yole Développement 2022
Reprinted here with permission from Yole Group.

Figure 3. Advanced DRAM and NAND Memory Fabs



Based on *Yole Q1 2022 DRAM/NAND Market Monitors*. Excludes fabs focused on mature/legacy technologies Solidigm included with SK Hynix based on acquisition.

History of the Semiconductor Memory Industry

The commercial memory industry dates back to the early 1970s, when Robert Noyce and Gordon Moore, in the first seminal years of Intel, helped spur the invention of the first commercially viable “Dynamic Random Access Memory” (DRAM).²² This 1-kilobit DRAM design was Intel’s first mass-produced product, conceived as a small but fast working memory between a computer’s processor and the magnetic core main memory of the time, taking advantage of the recent advent of silicon metal oxide semiconductor (MOS) transistor technology.²³ The DRAM concept quickly evolved thanks to improving yields, key technical advancements from Bob Dennard of IBM²⁴ and others to reduce the size of the cell, and the dawning realization that silicon transistor densities had a path to continue growing exponentially at a stable cadence (the birth of Moore’s Law).^{25,26}

The memory industry thus became one of the key enablers of modern computing systems, helping to drive the overall semiconductor industry. Through decades of brutal, boom/bust cycles as successive Asian countries sought to dominate the market, it has never been an industry for the faint of heart. The importance of memory has only continued to grow, supplying the linchpin components critical to powering today’s data economy.

The 1970s saw a dramatic expansion in DRAM densities, growing to 64 Kbit designs by the end of the decade. Demand increased with the continued growth in corporate computing and the phase-out of commercial magnetic core memories. The decade also saw an explosion of DRAM manufacturers, with system companies such as IBM, AT&T, and Motorola developing DRAMs for their own end systems, as well as several merchant suppliers joining the party, such as Texas Instruments, Mostek, and National Semiconductor. Standardization of the DRAM interface helped lower the market barriers to entry, while the easily testable, repeating patterns of DRAM arrays provided an easy gateway product to prove out production lines for more sophisticated products.²⁷

The 1980s saw the memory market become both international and commoditized, with the rise of the personal computer (PC) and DRAM manufacturing by several well-funded Japanese companies driving cost competition. By the end of the decade, companies such as Toshiba, NEC, and Hitachi dominated the DRAM market, forcing out most US suppliers.²⁸ Intel exited in 1984, as well-documented by Intel CEO Andrew Grove.²⁹ IBM continued large-scale internal production for their own systems, while other merchant suppliers such as Texas Instruments pursued overseas operations with the help of partner cost-sharing and subsidies.

Urgent congressional inquiry into the declining competitiveness of US semiconductor manufacturing drove the start of Sematech, originally conceived as a national memory production center, but quickly re-chartered as an R&D consortium, initially led by Robert Noyce. Further calls for protections around US manufacturing culminated in the US Japan Semiconductor Trade Agreement of 1986, spurring Japan to significantly reduce DRAM production. However, with most US suppliers already having left the business, and with DRAM demand soaring, the moves paradoxically led to a DRAM shortage that left many system companies unable to purchase sufficient supply.^{30,31}

Throughout the 1990s the demand for memory continued to skyrocket, driven increasingly by the exploding PC market and the “Win-Tel” PC cycles of Microsoft and Intel, demanding an ever-higher level of internal memory with each new product release and resultant improvements in computing power. The decade also saw a continuation of hypercompetitive boom-bust cycles. Wild price swings were caused by relatively modest supply and demand imbalances coupled with 1) the time lag to bring on new capacity when needed (due to fab construction lead-times), and 2) the unwillingness to idle capacity when not needed (due to high fixed-cost investments).

The decade also saw the emergence of South Korea as the new DRAM manufacturing powerhouse, with Samsung, Hyundai, and Lucky Goldstar supplanting many of the Japanese leaders by the end of the decade. Hyundai and LG operations merged to form Hynix after the 1997 South Korean financial crisis.^{32,33} Multiple Japanese entities exited DRAM production, with NEC and Hitachi combining operations into a pure-play DRAM spin-out in 1999 named Elpida Memory. Texas Instruments likewise retreated from the memory market in 1998 by selling its DRAM operations to the United States’ Micron Technology.^{34,35}

The 2000s saw multiple Taiwan vendors enter the market,³⁶ and a push from 200mm to 300mm wafers, requiring massive capital investments for brand new fabs and equipment. The 2001 tech crash and the 2008 financial crisis drove even wilder market gyrations, with significant industry losses. Investigations of price fixing during the 2001 downturn led to guilty pleas by all five of the largest international memory vendors.³⁷ In 2009, Qimonda of Germany declared bankruptcy after multiple bail-out packages had been extended, drawing an end to Europe’s DRAM ambitions.³⁸

A new form of memory called NAND Flash was commercialized in the 2000s, loosely descended from the original Erasable Programmable Read-Only Memory (EPROM), another early Intel

innovation.^{39,40} Although slower speed than DRAM and with each cell able to be rewritten only a handful of times, NAND devices offer much larger densities than DRAM and can retain data without power (i.e., non-volatility). NAND flash eventually overtook large portions of the hard-disk drive market in the form of solid-state drives (SSDs). NAND memory plays a critical role in PCs, cellphones, and server systems, occupying a layer of the “memory hierarchy” between large but slow hard disk drives, and the fast but low-density DRAM main memory.^{41,42}

The 2010s saw the slowing of the PC market. Demand for DRAM and NAND continued, however, driven by the consumer cellphone market and the enterprise datacenter market. By this point, the increasingly difficult technology transitions required massive annual R&D investments similar to those in the semiconductor advanced logic industry. Combined with the large capital requirements, only the largest operations were able to survive, effectively leading to an oligopoly.^{43,44} Samsung retained dominant market share in both DRAM and NAND, expanding beyond South Korea into mainland China for NAND production, and benefitting from in-house consumption of memory thanks to its large cellphone and systems divisions. Similarly, Hynix of South Korea expanded production in China with their DRAM facility in Wuxi. After defaulting on its loans in 2009, Hynix was put up for sale and subsequently rescued by the SK Group, which, at that time, was the third largest of the South Korean family-owned businesses. After a Japanese government

bail-out in 2009 and declaring bankruptcy in 2012,⁴⁵ Elpida Memory was purchased in 2013 by Micron, now the only combined DRAM and NAND pure-play operation left in the world.⁴⁶

China’s announcement in 2015 of the “Made in China 2025” policy initiative made clear however that both DRAM and NAND were of significant interest to Chinese national security, and that the 2020s would see the rise of indigenous Chinese memory production. Indeed, by the end of the decade, two Chinese memory companies had come forward—DRAM-focused CXMT and NAND-focused YMTC.⁴⁷

DRAM and NAND densities have continued to increase exponentially, achieving single component densities of 16 Gigabit and 512 Gigabit respectively by 2022. Both technologies pioneered vertical fabrication well before microelectronic logic devices, in DRAM’s case with a foray into deep substrate trench capacitors, and eventually with all vendors settling on narrow skyscraper-like “stacked” capacitors.⁴⁸ In the case of NAND, the push for more density first drove multiple bits per cell, leading in some cases to mere single digit numbers of stored electrons distinguishing between stored data states. Then NAND memory devices also “went vertical,” with the NAND cells assembled into long vertical chains, increasing density even further, and helping surmount noise issues that had stalled any continued scaling in planar structures.⁴⁹

Figure 4. DRAM Market Share 1970-2020

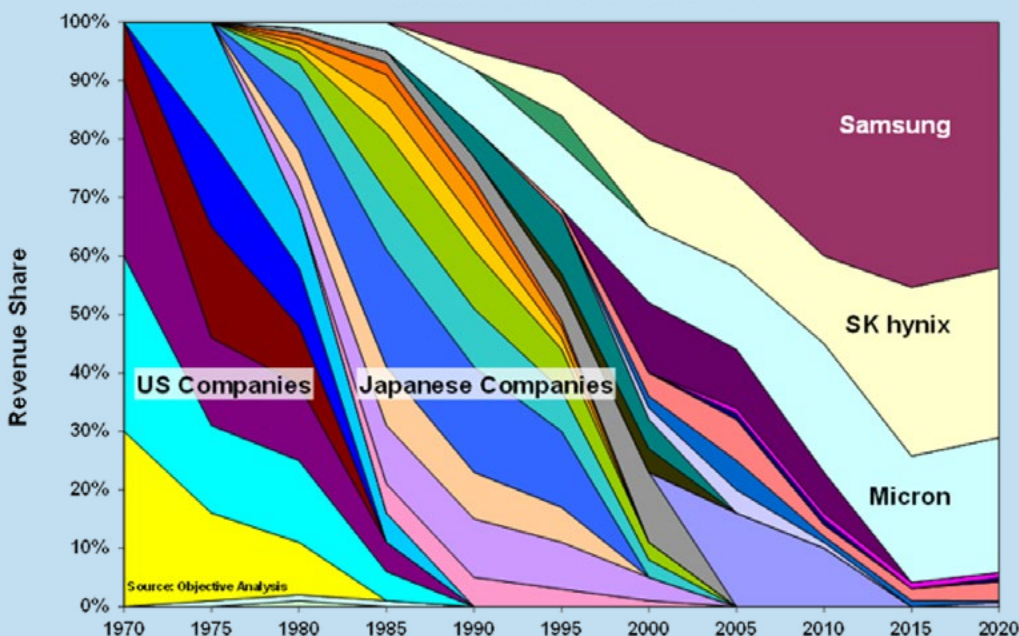
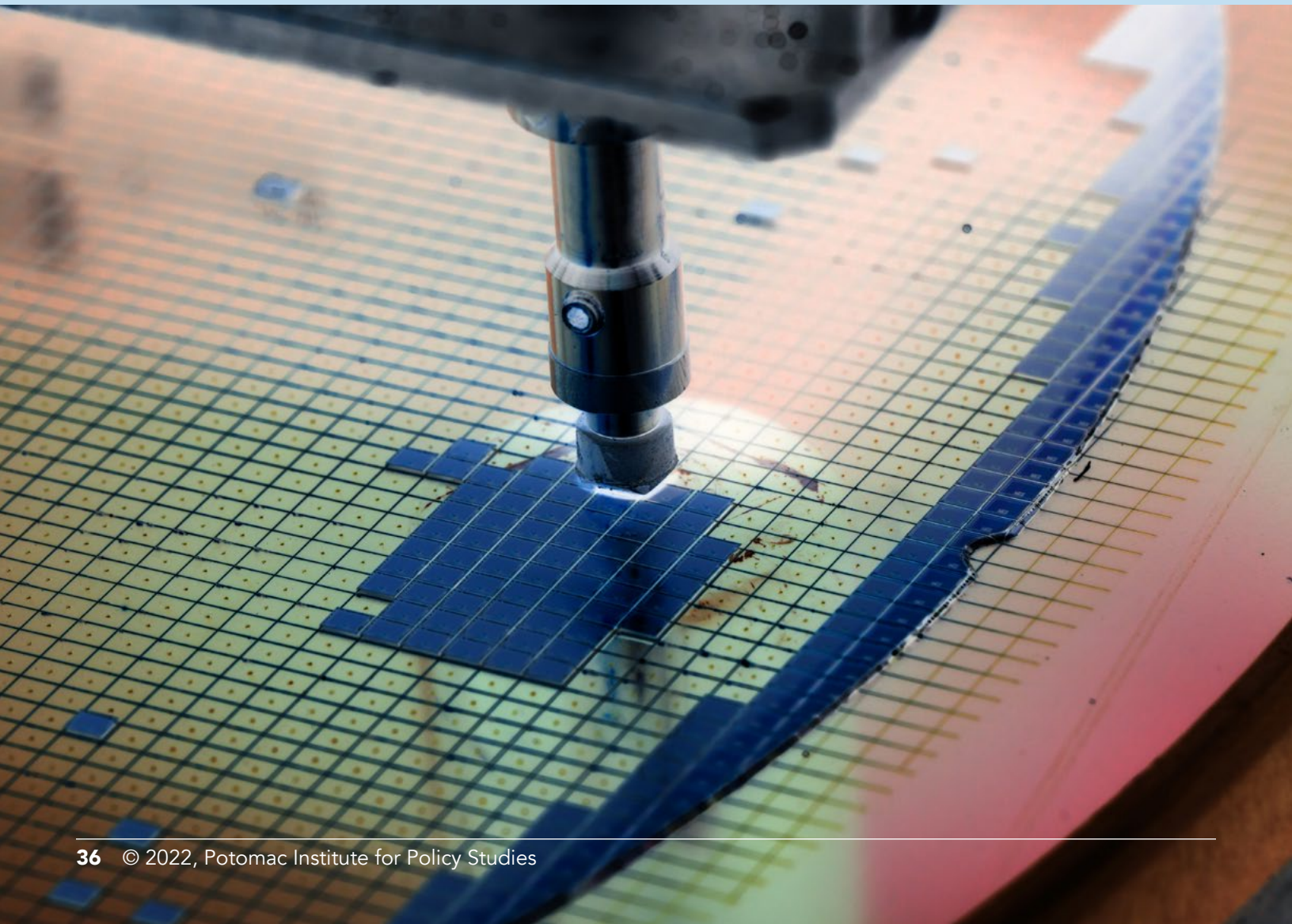


Chart Courtesy of Jim Handy, Objective Analysis www.objective-analysis.com.

Starting in the 1980s, it has been a recurrent drumbeat that memory “likely can only scale one, possibly two more generations.”⁵⁰ These concerns also helped spur a continual search to find “the holy grail” of memory technology—a new memory mechanism that could accomplish a better set of tradeoffs in cost, scalability, speed, endurance, and non-volatility, without relying on electric charge as the storage mechanism. Massive R&D efforts and dozens of dedicated startups were committed to finding winning alternatives. Many technology concepts appeared in research papers within industry journals. A smaller number made their way to actual prototypes and an even smaller number made it to low-volume production, but very few made it to levels of production that could sustain commercial interest for stand-alone components. While the future scaling of both DRAM and NAND is as challenging as ever, it is a testament to the strength of the original concepts as well as continuing innovation over the decades that DRAM and NAND have kept barreling through every prediction of imminent demise.

Despite the technical success of the memory industry, and the increasingly critical role memory plays in electronics, the economics of the industry have always been severely challenging. The industry earned essentially zero economic value from 1996 to 2012,⁵¹ due to significant asymmetric national investments by multiple Asian countries. Consolidation, as well as the growing technical and financial barriers to entry, helped somewhat stabilize the industry; However, decades of turmoil have left a production base for advanced memory located exclusively in Asia, and China’s investments are coming on strong, driving countries such as South Korea to double-down on semiconductor incentives benefitting indigenous memory suppliers.^{52,53,54} The current price tag for a new fabrication plant is around \$15B and growing, with an ongoing requirement for any serious vendor to invest multiple billions annually both in R&D and ongoing fab upgrades.⁵⁵ Storm clouds are indeed growing again for memory.



Labor cost is no longer a major factor in locating fabs, given the high levels of automation common to all 300mm semiconductor operations. However, competitive cost pressures have driven a need for massive scale. New memory fabs today aim for minimum capacities of 100,000 wafers per month, which is significantly larger than most logic fabs. This drives capital and construction costs well past \$15B for a new facility.⁵⁶ At that scale, and with challenging and volatile industry margins, organic funding is unrealistic. Thus, a major consideration in the industry is access to investment incentives.

The investment for a new fab is so large that any delay to the project can bankrupt the supplier. A new fab must be pulled into production as fast as possible to start generating cashflow. Suppliers drive fast ramps and minimize risk by focusing the fab on high-volume customer-qualified products. Speed of permitting and fab construction are also crucial. Countries such as South Korea, China, Taiwan, and Singapore help speed time-to-market, offering pro-business regulatory climates, as well as agencies that help with permitting and construction fast-tracking.^{57,58}

Once a memory fab is operational, it must continually be recapitalized with new equipment to produce the most advanced technology, upgrading much more frequently than fabs in other semiconductor sectors. Ongoing annual upgrades can easily cost over a third of annual revenue.⁵⁹ While a steep bill to pay, especially in market downturns, these investments are necessary to stay competitive as multiple bankruptcies over the decades can attest.

The size of both the initial and ongoing fab investments, as well as China's leap-ahead investments in memory have driven other countries to continue subsidies and favorable tax treatment—notably South Korea, home to both Samsung and SK Hynix. In 2021, South Korea announced a suite of capital, R&D, and other incentives for domestic semiconductor manufacturers meant to spur total investment of over \$450B through the next decade.⁶⁰

Research and Development in the Memory Sector

The need for advanced higher density memory pushes suppliers to the edge of Moore's Law, all to continually pack more memory cells on each chip. This is notoriously costly R&D that is growing more prohibitive every year, requiring

investment in new prototype equipment and dedicated cleanrooms as established technologies run out of steam. (This is true for all semiconductor sectors, with only a small handful of suppliers able to afford advanced R&D and tooling.)⁶¹ Experiments, short-loop trials, and eventually full-flow prototype wafers must be refined through multiple iterations, pushing the limits of advanced photolithography, etching processes, and new materials, all driven by brutally competitive time-to-market pressures. While both advanced memory and logic share the same mandate to continuously shrink feature sizes, the unique specialization of each field limits opportunity for synergy or cost-sharing.⁶²

Memory companies also invest R&D in product and system design enhancements to optimize for applications such as graphics, AI, and communications. Reducing power consumption is a constant priority, attempting to offset the added power from exponentially increasing chip densities and faster interface speeds. New data center interface standards such as Compute Express Link (CXL) allow larger banks of various types of memory to elegantly operate outside of traditional main memory slots.⁶³ Active research is attempting to move portions of logic into the memory fabric, to help reduce data movement in systems optimized for artificial intelligence, thus saving energy and reducing system bottlenecks.⁶⁴ Innovations in solid-state drives (SSDs) help deliver large performance gains to data centers and PCs, integrating dozens of NAND components together with logic controllers and firmware. These memory sub-systems also highlight the unique security risk of information storage, and the need for secure, high-integrity memory modules.⁶⁵

Even packaging processes, once considered relatively mature, now require significant investment in both R&D and capital. Applications such as cellphones and "wearable" devices must pack multiple memory devices close together to achieve small form-factors. Techniques placing advanced DRAM and NAND packages directly on top of the processor, as well as stacking thinned layers of NAND into a single terabyte package are already in high volume production.^{66,67} DRAM was one of the first technologies to achieve "3D" packaging, specifically for data center AI applications. Such "high-bandwidth memory" packages route signal connections through the DRAM stack itself, using wafers that have been thinned to less than the width of a human hair, and thousands of carefully aligned "Through-Silicon-Vias" that pass the signals through the

stack, buying speed and reducing power.⁶⁸ Many of these advanced packaging processes must be performed inside the fab itself, belying the perspective that packaging is “old technology.”

Taken in aggregate, the investment required to enable next-generation technology in the memory industry (adding together annual R&D and capital expenditures [capex] as a percentage of revenue), is higher or equal to other sectors, which also feature significantly larger margins. For instance, Micron (as the last remaining pure-play DRAM/NAND supplier), invested on average 50% of annual revenue into R&D and capex over 2019-2021, with an average operating margin of 23% over that time, and with all capital reinvested in existing fabs. By comparison, Texas Instruments, which is a leading US provider of analog components, invested 18% of revenue over the same time, with a 3-year operating margin of 44%. TSMC, a leading logic foundry based in Taiwan, invested a comparable 53% of revenue (of which a portion went to construction of new fabs) but with 40% operating margins.^{69,70} While point

comparisons across sectors always have caveats, the results highlight the steep investments required in the memory industry, despite thinner profit margins.

The Need and the Way Forward

Memory is not only more important than ever, but more difficult than ever. To be competitive, suppliers must continually invest larger sums in capital and R&D, despite challenging margins. The industry’s advanced fabs are all located in Asia, with several in China already and more coming, and with the biggest suppliers using large portions of their memory production in their own systems. Organic affordability of new memory fabs is no longer feasible. The global playing field is not level, due to ongoing asymmetric subsidization in Asian countries that have recognized memory’s importance. Absent appropriate attention, the US risks loss of access to the ever-growing amounts of secure memory that is foundational to the nation’s critical infrastructure, and overall economic and national security health.

Recommendations

1. **ELEVATE THE PRIORITY OF ADVANCED MEMORY PRODUCTION**
on US soil as a critical national security priority
2. **EXTEND INVESTMENT TAX CREDITS**
competitive with Asia to drive sufficient domestic memory fab reinvestment
3. **EXTEND R&D TAX CREDITS COMPETITIVE WITH ASIA**
to help defray the exploding costs for development of next-generation memory technology
4. **STREAMLINE THE SEMICONDUCTOR FAB REGULATORY PROCESS**
to ensure competitive memory fab construction “time-to-market”
5. **PROVIDE INCENTIVES FOR ONSHORE MEMORY PACKAGING CAPABILITIES**
necessary to complete a secure supply-chain

We thus propose a set of recommendations meant to ensure sustainable US access to advanced memory sufficient for national security.

These recommendations would help ensure US national and economic security are not at risk from multiple decades of asymmetric offshore incentives. While many argue against any form of active US industrial policy, these proposals merely level the playing field with countries that have methodically grown thriving semiconductor manufacturing environments over decades, using these same incentives for their domestic suppliers.⁷¹ And a thriving semiconductor fabrication and R&D base drives significant, well-documented benefits in high-value job creation, driving domestic economic activity.⁷²

These same incentives would also benefit other semiconductor sectors. In fact, several are now starting to confront the risks from similar asymmetric offshore incentives.⁷³ However, memory is the most urgent sector, given its recognized importance as the key enabler of the data economy, the exploding investments required to remain competitive, and the results from a long history of distortions. The following are comments on each recommendation in the context of the memory sector.

1. Elevate advanced memory production on US soil as a critical national security priority.

The US should prioritize advanced memory production on US soil as a key priority of incentives such as those contemplated in the CHIPS bill. Secure production of memory (DRAM and NAND) in sufficient quantities for US critical infrastructure needs is crucial for enabling our digital systems. Advanced memory production remains under threat from ongoing nation-state subsidies that continue to tilt the playing field outside the US. The US cannot tolerate risk to supply-chain disruptions of such an important ingredient of critical infrastructure. Supplies of other semiconductors from on-shore facilities will not matter if supplies of memory chips are limited.

2. Extend investment tax credits competitive with Asia to drive sufficient domestic memory fab reinvestment.

Such credits are critical to the memory sector, enabling crucial ongoing reinvestment in existing fabs and helping to ensure that new fabs can be built when market conditions warrant. While proposed legislation such as CHIPS could provide a “jump-start” to competitive US manufacturing, the total budgets are insufficient to offer grants to more than a few firms for new fabrication facilities, and will not help with the required ongoing upgrades. Investment tax credits would have the added advantage of incentivizing “skin-in-the-game,” benefiting companies only in proportion to the investments they make in US manufacturing operations. Such incentives are common in Asia; in fact, South Korea’s recent announcement highlighted 20% tax credits for all capital invested in new fab construction.⁷⁴

3. Extend R&D tax credits competitive with Asia to help defray the exploding costs for development of next-generation memory technology.

The memory industry requires an urgent level of innovation due to the slowdown in Moore’s Law, while R&D necessary for next-generation memory fabrication processes becomes ever more expensive. Accordingly, tax credits for R&D activity in the US are needed to remain competitive against Asian countries that offer R&D incentives to their own domestic semiconductor manufacturers.⁷⁵ These credits help spur the purchase of next-generation advanced semiconductor tools—an industry the US still leads.⁷⁶ The institution of a competitive US semiconductor R&D tax credit would also send the message that semiconductor manufacturing is welcome in the US again, helping revitalize critical engineering training programs at US colleges that have withered as these jobs moved overseas. As a reference, South Korea’s recently announced program highlights R&D tax credits of 40% (up from 30%), and provides benefit primarily to the region’s indigenous memory chip suppliers.⁷⁷

4. Streamline the semiconductor fab regulatory process to ensure competitive memory fab construction “time-to-market.”

While new fab “time-to-market” is important to all semiconductor manufacturers, it is crucial to memory suppliers. Costs are at least \$15B for a competitively sized memory fab, which carries the risk of insolvency if any delays forestall production. The US must streamline the regulatory process for new US semiconductor fabs to be competitive with fabs achievable in Asia. The US Department of Commerce has published an informative brief acknowledging the issue and proposing a set of recommendations.⁷⁸ Without waiving necessary regulation, the United States needs a fast-track permitting process for domestic semiconductor fabs to inject urgency into the maze of bureaucracy currently navigated when bringing a semiconductor plant to fruition.⁷⁹

5. Provide incentives for onshore memory packaging capabilities necessary to complete a secure supply-chain.

Packaging, and the adjacent steps of printed-circuit-board fabrication and system assembly, are important manufacturing steps for all semiconductors. However, they are mission-critical for memory. Memory components and modules drive significant component and board volumes, with a need for cost-effective packaging across a variety of technologies. The US domestic outsourced assembly and test (OSAT) industry was an early casualty of offshoring (based primarily on labor cost considerations), so that today there are few high-volume factories outside of Asia. However, this critical last link in the component supply chain drives significant security and supply-chain access risk.⁸⁰ US wafer fabs would still be under significant threat of disruption if the wafer and components continue to cross the ocean multiple times for finalization. As noted above, however, recent trends are injecting more innovation into packaging, but also driving up capital requirements. Opportunities exist to leverage more advanced automation technologies in this process, which makes it more feasible to reshore. The US should consider incentives to onshore sufficient OSAT capabilities as a national security imperative.

Summary

The semiconductor memory industry needs urgent national attention. While owing its start to US innovation, the industry painfully left US shores as the first case of semiconductor globalization. Asymmetric national subsidization then led to decades of brutal economic cycles that eviscerated the supplier landscape, leading to a risky concentration of fabs based in Asia, including multiple sites in China. Vicious cost pressures require massive ongoing investment. A slowing of Moore’s Law drives skyrocketing R&D costs, challenging already thin margins. These factors have caused every non-US country that still has a domestic memory supplier to double-down on subsidies. Meanwhile, China’s indigenous producers are coming on strong. Memory suppliers are truly competing against the might of entire countries, while ever-present market volatility continues.

While it is not an industry for the faint of heart, these components are at the heart of every electronic system on which our data economy relies, driving performance benefits and the lion’s share of system silicon content. Ensuring the integrity of these components, as well as the security of a viable, complete supply chain—from R&D to final system—is a national imperative. The US should urgently work to ensure a level playing field for onshore production and development of memory, with immediate priority on securing US supply for our most critical systems. Although there are no quick fixes, implementation of these incentives will make the US a viable home for advanced memory manufacturing again, ensuring that we have not forgotten about memory.

Acknowledgements

The author would like to thank the following individuals and acknowledge their contributions:

Bob Hummel, Dan Marujo, Mike Fritze, Al Shaffer, Ted Glum, and all of Potomac Institute for Policy Studies for subject matter contribution, discussion, and review.

Jim Handy of Objective Analysis, a leading memory industry analyst firm, for historical market share data and other insights: www.objective-analysis.com. Mr. Handy’s blogs can be found at TheMemoryGuy.com and TheSSDGuy.com.

Ivan Donaldson and team of the Yole Group, a leading semiconductor market research firm, for current memory industry market data: www.yolegroup.com.

Endnotes

- 1 "2022 Factbook," Semiconductor Industry Association.
- 2 "CHIPS BILL," US Congress, and "COMPLETE BILL," US Congress.
- 3 "Design Win for Smartphone...," SystemPlus Consulting, subsidiary of Yole Group.
- 4 Based on Yole Group Q1'22 DRAM and NAND Market Monitors – Industry average die/wafer in FQ1'22 applied to AMD EPYC 7002 2-socket server max memory of 4TB and PCIe direct attach storage max capacity of 184TB.
- 5 Keith Bradsher and Paul Mozur. "China's Plan to Build Its Own High-Tech Industries Worries Western Businesses," *New York Times* March 7, 2017.
- 6 Dan Kim and John VerWey. "The Potential Impacts of the Made in China 2025 Roadmap on the Integrated Circuit Industries in the U.S., EU and Japan," *U.S. International Trade Commission* August 2019.
- 7 Roslyn Layton. "White House Report On China: Short-Term Profits Undermine Long-Term Resilience," *Forbes* June 10, 2021.
- 8 "World Semiconductor Trade Statistics," WSTS.
- 9 "Semi Capex on Pace for 34% Growth in 2021 to Record \$152.0 Billion," IC Insights.
- 10 Harald Bauer, et al. "Memory: Are Challenges Ahead," McKinsey & Company, March 8, 2016.
- 11 Kotaro Hosokawa. "Micron Challenges Samsung's Dominance in Memory Market," *Nikkei Asia*, May 23, 2021.
- 12 "Q1'22 DRAM and NAND market monitors," Yole Group, www.yolegroup.com.
- 13 NAND is a high-density, non-volatile "storage" memory, able to retain its data without power, however significantly slower than DRAM, and of limited endurance, with each cell only able to be written a limited number of times. See: https://www.semiconductors.org/wp-content/uploads/2021/02/High-end-Vol-Stream-Memory_Omdia.pdf.
- 14 Results for Q1'2022, based on "Q1'22 DRAM and NAND market monitors," Yole Group, www.yolegroup.com.
- 15 Ryan Smith. "Intel Sells SSD Business to SK Hynix as New Subsidiary Solidigm," Dec 28, 2021.
- 16 Cheng Ting-Fang and Lauly Li. "US-China Tech War: Beijing's Secret Chipmaking Champions," *Financial Times* May 12, 2021.
- 17 Anton Shilov. "ChangXin Memory Technologies (CXMT) is Ramping up Chinese DRAM Using Qimonda IP," *AnandTech* Dec 2, 2019.
- 18 "China Big Fund Stepping Up Investment in Homegrown Memory Sector," China Pulse.
- 19 "Taking Stock of China's Semiconductor Industry," Semiconductor Industry Association, July 13, 2021.
- 20 Fabs in Taiwan, Singapore, and Japan, acquired from Qimonda, Texas Instruments, and Elpida (Micron 10K annual reports).
- 21 Q1'22 DRAM and NAND market monitors," Yole Group, www.yolegroup.com.
- 22 "1-Kbit DRAM (Intel, U.S.A)," Semiconductor History Museum of Japan.
- 23 George Gilder. *Microcosm: The Quantum Revolution in Economics and Technology* (New York: Simon and Schuster) 1989: 90-100.
- 24 Bob Dennard's DRAM innovation occurred in 1966, prior to the commercial start of the DRAM industry. Dennard then went on to publish a pioneering set of observations and rules about semiconductor scaling, later termed "Dennard Scaling," providing the industry a multi-decade roadmap on geometry, voltage, and power trends. For more information, see: <https://history.computer.org/pioneers/dennard.html>.
- 25 Michael Malone. *The Intel Trinity: How Robert Noyce, Gordon Moore, and Andy Grove Built the World's Most Important Company* (New York: Harper Business) 2014: 131-140.
- 26 Note: In the late 1960s, several firms had designed semiconductor "Static RAM" memories, much lower density, significantly lower volume, and without the commercial success that the 1K DRAM enjoyed.
- 27 George Gilder. *Microcosm: The Quantum Revolution in Economics and Technology* (New York: Simon and Schuster) 1989: 7, 11, 13.
- 28 Thomas Lueck. "Mostek, Big Chip Maker, Shut," *New York Times* Oct 18, 1985.
- 29 Andrew Grove. *Only the Paranoid Survive* (Penguin Random House). 1998.
- 30 Gilder. *Microcosm*, Ch. 12, pages 150-160.
- 31 Andrew Pollack. "The Selling of the 256K RAM," *The New York Times* Jun 3, 1983.
- 32 Leslie Helm. "A New Force in Chip Wars: Korean Chip Exports are Growing 35% a Year, and the US and Japan are Worried," *Los Angeles Times* Aug 17, 1992.
- 33 B.H. Seo. "Hyundai, LG Complete Merger," *EETimes* Apr 22, 1999.
- 34 Staff, et al. "TI to Sell DRAM Business to Micron," *EETimes* June 18, 1998.
- 35 Laurie Flynn. "Samsung to Pay \$300 Million Fine for Price Fixing," *New York Times* Oct 13, 2005.
- 36 Claire Jim. "Analysis-Taiwan's DRAM Industry: Desperately Seeking a Future," *Reuters* Dec 22, 2011.
- 37 Laurie Flynn. "Samsung to Pay \$300 Million Fine for Price Fixing."
- 38 Qimonda was the DRAM spin-out of Infineon, which itself was the semiconductor products spin-out from Siemens; See: <https://www.marketwatch.com/story/qimonda-files-for-bankruptcy>.
- 39 "1971: Reusable Programmable ROM Introduces Iterative Design Flexibility," Computer History Museum.
- 40 Note: Both NOR and NAND memory also benefitted from significant innovations at Toshiba in the early 1980s.
- 41 Jim Handy, "A History of Flash Memory and its Rise in the Enterprise."
- 42 Datacenter SSDs today are routinely rewritten so frequently as to approach cell wear-out limits, making them in some cases a "consumable" product.
- 43 Jim Handy/Objective Analysis, "Why DRAM Vendors Must Consolidate; How Spiraling Fab Costs Force Vendors to Merge," 2009.
- 44 Harald Bauer, et al. "Memory: Are Challenges Ahead?," McKinsey&Company, March 8, 2016.
- 45 Hiroko Tabuchi. "In Japan, Bankruptcy for a Builder of PC Chips," *New York Times* Feb 27, 2012.
- 46 Dylan McGrath. "Micron Closes Elpida Acquisition," *EETimes* Jul 31, 2013.
- 47 "Taking Stock of China's Semiconductor Industry," Semiconductor Industry Association (SIA), July 13, 2021.
- 48 Solid-State Circuits Society. "The DRAM Story," *IEEE Solid-State Circuits Society News* Vol. 13. No. 1, (Winter) 2008.
- 49 Maarten Rosmeulen, et al. "How 3D NAND Flash Works, What Lies Ahead in its Density Roadmap," *EDN*, March 8, 2022.
- 50 IEEE "The DRAM Story."
- 51 Harald Bauer, et al. "Memory: Are Challenges Ahead?," McKinsey&Company, March 8, 2016.
- 52 "Taking Stock of China's Semiconductor Industry," Semiconductor Industry Association (SIA), July 13, 2021.
- 53 "Q1'22 DRAM and NAND market monitors," Yole Group, www.yolegroup.com.
- 54 Samuel K Moore. "South Korea's \$450-Billion Investment Latest in Chip Making Push," *IEEE Spectrum* May 26, 2021.
- 55 Harald Bauer, et al. "Memory: Are Challenges Ahead" March 2016.
- 56 Harald Bauer, et al. "Memory: Are Challenges Ahead" 2016, with cost extrapolations to 2022.
- 57 VerWey. "No Permits, No Fabs: The Importance of Regulatory Reform for Semiconductor Manufacturing," *Center for Security and Emerging Technology*, October 2021.
- 58 Vincent Chin. "Sustaining Singapore's Success: Lessons of Growth and Delivery," *Centre for Public Impact* June 11, 2015.
- 59 Micron 2021 10K, capital expenditures over 2019/20/21 equaled 39% of revenue.
- 60 Kim Jaewon. "South Korea Plans to Invest \$450bn to Become Chip 'Powerhouse'," *Nikkei Asia* May 13, 2021.
- 61 Rick Merritt. "GlobalFoundries Halts 7nm Work," *EETimes* August 27, 2018.
- 62 Mark Lapedus. "DRAM Scaling Challenges Grow," *SemiconductorEngineering*, November 21, 2019.
- 63 Gary Ruggles. "Introduction to the Compute Express Link Standard," *Synopsys* 2022.
- 64 Bob Beachler. "The Advantages of At-Memory Compute for AI Inference," *EETimes* January, 24, 2022.
- 65 Chris Stokel-Walker. "Hackers are Now Dropping Malware into SSD Drives through Firmware Updates," *cybernews* January 12, 2022.
- 66 Gary Hilson. "Samsung Makes New Package for Smartphones," *EETimes* February 17, 2015.

- 68 Gary Hilson. **"AI Expands HBM Footprint,"** *EETimes*, January 20, 2022.
- 69 Calculated from respective SEC 10K's and 20F forms for Micron, TI, and TSMC, using three-year weighted averages.
- 70 Had TSMC 2021 capex spending remained constant to 2019/20, three-year intensity would have been 45%; TSMC 20F form; See: Nick Flaherty. **"TSMC's \$100bn Fab Plans to Expand Capacity,"** *EE News* April 2, 2021.
- 71 **"Global Semiconductor Incentives,"** Semiconductor Industry Association (SIA), February 2022: 2-4.
- 72 **"Robust Federal Incentives for Domestic Chip Manufacturing Would Create an Average of Nearly 200,000 American Jobs Annually as Fabs are Built, Add Nearly \$25 Billion Annually to U.S. Economy,"** Semiconductor Industry Association (SIA), May 19, 2021.
- 73 **"Taking Stock of China's Semiconductor Industry,"** Semiconductor Industry Association (SIA), July 13, 2021.
- 74 **"Global Semiconductor Incentives,"** Semiconductor Industry Association.
- 75 **"Global Semiconductor Incentives,"** Semiconductor Industry Association.
- 76 However, with today's uneven playing field, these tools (ironically) often help enable non-US suppliers first.
- 77 Kim Jaewon. **"South Korea Plans to Invest \$450bn to Become Chip 'Powerhouse,'"**
- 78 **"Streamlining Permitting and Reducing Regulatory Burdens for Domestic Manufacturing,"** US Department of Commerce, October 6, 2017.
- 79 John VerWey. **"No Permits, No Fabs: The Importance of Regulatory Reform for Semiconductor Manufacturing,"** *Center for Security and Emerging Technology*, October 2021.
- 80 Jordan Robertson, et al. **"The Long Hack: How China Exploited a U.S. Tech Supplier,"** *Bloomberg* February 12, 2021.

