AI Master Research

The 8-Layer Stack · Three Transitions · Physical AI · 98 Names Ranked

Damac Group · Research

May 22, 2026

Contents

AI MASTER RESEARCH 4

- Table of contents . . . 4

Part 1 — The 8-Layer AI Stack 4

- Why the stack matters more than the theme . . . 4
- The eight-layer stack . . . 5
- The two cascades . . . 11
- How to read the rest of this document . . . 12

Part 2 — Three Transitions Reshaping the Stack 13

Part 2b — NVIDIA Q1 FY27 · Physical AI as the Next Wave 21

- Headline . . . 21
- Macro summary . . . 22
- The five-bullet Physical-AI thesis . . . 22
- What this means for picks-and-shovels . . . 22
- What's NOT in this call but should be . . . 23

Part 2c — Physical AI · Cross-Layer Theme Linking AI Semis + Robotics 23

- Why this matters now . . . 23
- How Physical AI maps onto the 8-layer AI stack . . . 24
- Linked picks-and-shovels — names already in AI Dashboard with Physical AI tailwind . . . 27
- New names — Physical AI pure-plays to add to the watchlist . . . 30
- Validation table — our prior Robotics-MASTER report vs Jensen's May 20 call . . . 34
- Updated thesis statement . . . 36

Part 3 — Tier A · Highest Conviction 37

- Layer 1 · NVDA — NVIDIA Corporation . . . 37
- Layer 1 · AVGO — Broadcom Inc. . . . 41
- Layer 1 · MRVL — Marvell Technology . . . 44
- Layer 2 · TSM — TSMC ADR . . . 48
- Layer 2 · KLAC — KLA Corporation . . . 52
- Layer 2 · LRCX — Lam Research . . . 56
- Layer 2 · CAMT — Camtek Ltd. . . . 60

Layer 2 · ONTO — Onto Innovation . . . 63
Layer 3 · 000660 — SK Hynix . . . 67
Layer 3 · MU — Micron Technology . . . 71
Layer 3 · SNDK — SanDisk Corporation . . . 75
Layer 3 · WDC — Western Digital . . . 79
Layer 3 · 285A — Kioxia Holdings . . . 83
Layer 4 · ALAB — Astera Labs . . . 87
Layer 4 · LITE — Lumentum Holdings . . . 91
Layer 4 · 522 — ASMPT Limited . . . 95
Layer 4 · BESI — BE Semiconductor Industries . . . 99
Layer 7 · GEV — GE Vernova . . . 103
Part 1 portfolio synthesis . . . 107

Part 4 — Tier B · Add on Pullback / Hold **109**

Layer 0 · ANET — Arista Networks . . . 109
Layer 0 · CSCO — Cisco Systems . . . 110
Layer 0 · DELL — Dell Technologies . . . 111
Layer 0 · IREN — IREN Limited . . . 112
Layer 0 · NBIS — Nebius Group . . . 113
Layer 0 · WULF — TeraWulf . . . 114
Layer 1 · AMD — Advanced Micro Devices . . . 115
Layer 1 · ARM — Arm Holdings . . . 116
Layer 1 · INTC — Intel Corporation . . . 117
Layer 1 · QCOM — QUALCOMM Inc. . . . 118
Layer 2 · AMAT — Applied Materials . . . 119
Layer 2 · ASML — ASML Holding . . . 120
Layer 2 · TSEM — Tower Semiconductor . . . 121
Layer 3 · HY9H — SK Hynix Sponsored GDR . . . 122
Layer 3 · 005930 — Samsung Electronics . . . 123
Layer 3 · SMSN — Samsung Electronics GDR . . . 124
Layer 4 · 007660 — ISU Petasys . . . 125
Layer 4 · 009150 — Samsung Electro-Mechanics . . . 126
Layer 4 · 067310 — Hana Micron . . . 127
Layer 4 · 3037 — Unimicron Technology . . . 128
Layer 4 · 4062 — Ibiden . . . 129
Layer 4 · 6315 — TOWA Corporation . . . 130
Layer 4 · AAOI — Applied Optoelectronics . . . 132
Layer 4 · AXTI — AXT Inc . . . 133
Layer 4 · CIEN — Ciena Corporation . . . 134
Layer 4 · COHR — Coherent Corp . . . 135
Layer 4 · CRDO — Credo Technology . . . 137
Layer 5 · 2327 — Yageo Corporation . . . 138
Layer 5 · 3443 — Global Unichip Corp (GUC) . . . 139
Layer 5 · 3661 — Alchip Technologies . . . 140
Layer 5 · 6762 — TDK Corporation . . . 141
Layer 5 · 6981 — Murata Manufacturing . . . 142
Layer 5 · APH — Amphenol Corporation . . . 143

Layer 5 · CDNS — Cadence Design Systems . . . 144
Layer 5 · GLW — Corning Incorporated . . . 145
Layer 5 · SNPS — Synopsys . . . 146
Layer 6 · 4063 — Shin-Etsu Chemical . . . 147
Layer 6 · 6857 — Advantest Corporation . . . 148
Layer 6 · FORM — FormFactor . . . 149
Layer 6 · TER — Teradyne . . . 150
Layer 7 · CEG — Constellation Energy . . . 151
Layer 7 · VST — Vistra Corp. . . . 152
Layer 7 · BE — Bloom Energy . . . 153
Layer 7 · MOD — Modine Manufacturing . . . 153
Layer 7 · POWL — Powell Industries . . . 154
Layer 7 · PWR — Quanta Services . . . 155
Layer 7 · STRL — Sterling Infrastructure . . . 156
Layer 7 · VICR — Vicor Corporation . . . 157
Layer 7 · VRT — Vertiv Holdings . . . 158

Part 5 — Tier C · Watch / Pass **159**

Layer 0 · CRWV — CoreWeave . . . 159
Layer 4 · 033640 — Nepes . . . 160
Layer 4 · 042700 — Hanmi Semiconductor . . . 160
Layer 4 · 110990 — DIT Corp . . . 161
Layer 4 · 489790 — Hanwha Vision (Semitech) . . . 162
Layer 4 · IQE — IQE plc . . . 162
Layer 4 · LWLG — Lightwave Logic . . . 163
Layer 4 · POET — POET Technologies . . . 164
Layer 4 · SIVE — Sivers Semiconductors . . . 164
Layer 4 · SOI — Soitec . . . 165
Layer 5 · 4004 — Resonac Holdings . . . 165
Layer 5 · 5201 — AGC Inc. . . . 166
Layer 5 · 5333 — NGK Insulators . . . 167
Layer 5 · 5802 — Sumitomo Electric Industries . . . 167
Layer 5 · 6920 — Lasertec Corporation . . . 168
Layer 5 · PLAB — Photronics Inc. . . . 168
Layer 6 · 005290 — Dongjin Semichem . . . 169
Layer 6 · 2408 — Nanya Technology . . . 170
Layer 6 · 3436 — SUMCO Corporation . . . 170
Layer 6 · 357780 — Soulbrain . . . 171
Layer 6 · 3711 — ASE Technology . . . 172
Layer 6 · 402340 — SK Square . . . 172
Layer 6 · 4186 — Tokyo Ohka Kogyo (TOK) . . . 173
Layer 6 · 6488 — GlobalWafers . . . 174
Layer 6 · 8150 — ChipMOS Technologies . . . 174
Layer 6 · AEHR — Aehr Test Systems . . . 175
Layer 6 · AMKR — Amkor Technology . . . 176
Layer 6 · VECO — Veeco Instruments . . . 176
Layer 7 · NVTS — Navitas Semiconductor . . . 177

Layer 7 · POWI — Power Integrations . . . 178
Layer 7 · WOLF — Wolfspeed Inc. . . . 178

Appendix A — Coverage Audit **179**
Dedupes resolved . . . 179
Names that fell through and were stubbed in assembly . . . 180

Appendix B — Data Refresh Provenance **180**
Tickers flagged needs_manual_check . . . 181

AI MASTER RESEARCH

The 8-layer AI semiconductor stack — bottlenecks, transitions, and 98 names ranked Tier A/B/C.

Updated: May 22, 2026 · v1.1 · Physical AI section added post Q1 FY27 print **Universe:** 98 tickers **Ranking method:** Fresh data refresh (live price + last 30-day news + most recent earnings) → tier discipline applied (Tier A = highest conviction, target 15-20 / Tier B = add on pullback, hold / Tier C = watch / pass)

Table of contents

- Part 1 — The 8-Layer AI Stack
- Part 2 — Three Transitions Reshaping the Stack
- Part 2b — NVIDIA Q1 FY27 · Physical AI as the Next Wave
- Part 2c — Physical AI · Cross-Layer Theme Linking AI Semis + Robotics
- Part 3 — Tier A · Highest Conviction
- Part 4 — Tier B · Add on Pullback / Hold
- Part 5 — Tier C · Watch / Pass
- Appendix A — Coverage Audit
- Appendix B — Data Refresh Provenance

Part 1 — The 8-Layer AI Stack

An opening lecture before we get to the stock picks.

Why the stack matters more than the theme

Let's start with what's actually happening when ChatGPT answers your question. The electrons that make that work flow through eight distinct layers of hardware, each with its own choke point. Some of these layers have one or two dominant suppliers. Some have a dozen names fighting over scraps. A few have a single Japanese factory in a small town that the entire industry quietly prays does not catch fire. That asymmetry — the difference between a layer with three credible suppliers and a layer with one — is where the money gets made. Themes get you to the right neighbourhood. Bottlenecks get you to the right house.

I have spent twenty-five years watching capital chase themes — biotech, internet 1.0, China growth, shale, cloud, EVs, now AI. The pattern repeats. The thematic narrative is correct. The first-order beneficiaries are usually obvious, expensive, and crowded by the time you read about them in the Financial Times. The real returns sit two or three layers down the value chain, in companies that look boring on a Bloomberg screen and whose CEOs cannot explain what they do at a cocktail party without putting people to sleep. **Picks and shovels** is not a clever framing; it is the historical default winner of every industrial buildout from the railways onward. The novelty in AI is not that picks-and-shovels works — it is that the shovels themselves have shovels, and those sub-shovels are sometimes the best businesses in the entire stack.

The structure of this document follows that logic. We are going to walk down the stack from the buyers — the hyperscalers spending half a trillion dollars a year in 2026 — through the chips, the fabs, the memory, the bonding, the obscure sub-components, the test gear, and finally to the electrons themselves, where a 765kV grid line ends in a 0.65V transistor and somebody has to make the impedance work. At each layer we will identify what physically happens, why the engineering is hard, who has a moat, and where the choke point sits. Then in Part II we will rank every name we cover in Tiers A, B, and C. By the time you finish the document, you should be able to look at any AI press release and immediately know which three or four companies actually benefit and which twenty are along for the ride.

The eight-layer stack

Layer 0 — Hyperscalers (the demand)

This is where the money originates. **Microsoft**, **Amazon Web Services**, **Google**, **Meta**, **Oracle**, and the newer pure-play GPU clouds like **CoreWeave** and **Nebius** are the buyers of essentially the entire downstream stack. Their 2026 capital expenditure budgets, when you add them up properly and net out real estate and non-AI spending, fall in the range of **$525 to $655 billion** depending on which sell-side analyst you trust and how you treat Oracle's debt-funded ramp. That number was $230 billion in 2023. The five-year compound is genuinely unprecedented for capital goods of this scale — there is no real analogue in industrial history. The closest comparable is the US interstate highway system, and that took thirty years.

What hyperscalers physically do is rent computation. A customer — OpenAI, Anthropic, an enterprise, a developer — sends a query or a training job to a data centre. Inside that data centre sit racks of GPU servers, each rack drawing somewhere between 30 and 130 kilowatts depending on vintage. The hyperscaler bills for time, for tokens, for GPU-hours, for stored bytes. The business model is straightforward: depreciate the hardware over five or six years, layer in power and real estate and labour, mark it up, and sell the capacity. Margins on AI inference are still negative for most operators at the workload level — the hardware is too new and too expensive — but the bet is that unit economics improve as Blackwell, then Rubin, then Rubin Ultra each deliver roughly 2-3x performance per dollar over their predecessors.

The bottleneck at Layer 0 is not money. The hyperscalers can raise capital essentially without limit at this point — Oracle just did $18 billion in a single bond deal — and Microsoft, Google, and Meta generate enough free cash flow to self-fund. The bottleneck is **everything else** in the stack downstream of them. They are demand-unconstrained and supply-constrained. That is the single most important sentence in this entire document. When a hyperscaler tells you their capex is "limited by GPU availability" or "limited by power" or "limited by substrate supply," they

are telling you exactly which sub-layer is the pinch point in that quarter. Listen to those calls carefully.

How Layer 0 connects: it buys everything below. It signs the cheques that flow to Nvidia, which flow to TSMC, which flow to ASML and Tokyo Electron and SK Hynix and Murata and Lasertec. Every dollar of hyperscaler capex propagates through six or seven layers of supply chain, and the further down the chain you go, the more concentrated the supplier base. Hyperscalers themselves are competitive — six or seven real players. By the time you reach Layer 5, you are looking at single suppliers for critical inputs.

Layer 1 — Compute fabric

This is the brains. Layer 1 is where the actual matrix multiplications happen — the transformer attention heads, the feedforward layers, the gradient updates. Physically, it is a collection of silicon dies, each die roughly the size of a postage stamp, each containing 80 to 200 billion transistors switching at gigahertz frequencies. **Nvidia** dominates the merchant GPU market with an estimated 85-90% share by revenue, and a software moat (CUDA, NCCL, TensorRT) that has now compounded for nearly two decades. **AMD** is the credible challenger, currently shipping MI300X and MI350 series with a roadmap into MI400. **Intel** is a distant third in the merchant market and is competing more meaningfully through its foundry arm than its products.

The custom ASIC side is structurally different and arguably more interesting from an investment standpoint. Each hyperscaler wants its own silicon for the same reason every car company wants its own engine: to escape the margin tax that Nvidia extracts and to tune the hardware to its specific workload. Google has TPU (designed in-house, manufactured at TSMC). Amazon has Trainium and Inferentia (designed with **Alchip** and **Marvell**). Meta has MTIA. Microsoft has Maia. OpenAI has a forthcoming chip co-designed with **Broadcom**. The merchant designers who service this market — Broadcom, Marvell, Alchip, and Taiwan's **GUC** — are in some ways better businesses than Nvidia, because their customer is locked into a multi-generation co-design relationship and switching costs are crushing.

The engineering problem at Layer 1 is the **memory wall**. A modern AI accelerator can do something like 10-20 petaflops of dense FP8 compute. To keep the arithmetic units fed, you need terabytes per second of memory bandwidth and microsecond-scale chip-to-chip interconnect. The chip itself is no longer the bottleneck — feeding the chip is. This is why HBM (Layer 3) and advanced packaging (Layer 2) and optical interconnect (Layer 4) all matter so much.

Bottleneck: design talent and customer relationships. There are perhaps two hundred engineers on Earth who can lead a frontier accelerator design from architecture to tape-out, and they all know each other. The economic moat is not the silicon itself; it is the software ecosystem and the multi-year customer co-design contracts. **ARM** sits adjacent to this layer as the CPU IP licensor — Grace, the CPU in Nvidia's superchip, runs ARM cores — and collects a small but high-margin royalty on essentially every AI server CPU shipped.

Connection: Layer 1 designs the chips that Layer 2 manufactures, Layer 3 feeds with memory, Layer 4 stitches together, Layer 5 supplies the sub-components for, Layer 6 tests, and Layer 7 powers. Every other layer in this document exists to serve Layer 1.

Layer 2 — Foundry and advanced packaging

If Layer 1 is the design, Layer 2 is the print shop — and it is one of the most consequential industrial bottlenecks in the world today. **TSMC** manufactures the overwhelming majority of leading-edge AI silicon. Nvidia, AMD, Broadcom, Marvell, Apple, MediaTek, and most ASIC startups all tape out at TSMC's N5, N4, N3, and increasingly N2 process nodes. **Samsung Foundry** is a credible second source, currently winning some Qualcomm and select hyperscaler ASIC work, but it remains behind TSMC on yield and density. **Intel Foundry** is a genuine wildcard — 18A is real, Microsoft has signed up, but the financial drag on Intel proper is severe and the next two years will determine whether IFS survives as a structurally separate business.

What happens physically: a 300mm silicon wafer is patterned through 80-100 lithography steps using extreme ultraviolet light at 13.5 nanometre wavelength. Each step deposits or etches a layer of metal, dielectric, or doped silicon. The leading nodes — N3, N2, A14 — require multi-patterning and increasingly high-NA EUV. A single ASML High-NA tool costs around $380 million and TSMC will operate dozens of them. The capital intensity is staggering: a single new fab module runs $20-25 billion and takes three to four years from green field to first wafers out.

The bottleneck at Layer 2 is not, surprisingly, the wafer fabrication itself. TSMC has aggressive capacity expansion in Arizona, Kumamoto, and Hsinchu, and the front-end is keeping pace with demand. The bottleneck is **CoWoS** — Chip-on-Wafer-on-Substrate — the advanced packaging step that bonds a logic die, multiple HBM stacks, and a silicon interposer onto a single substrate. Every Nvidia Blackwell, every AMD MI350, every Google TPU shipped today requires a CoWoS slot. TSMC's CoWoS capacity is the single tightest constraint in the entire AI hardware stack as of 2026, and it gates the whole industry's output. CoWoS-L (with local silicon interconnect bridges) and the move to glass substrates over the next 24 months will alleviate some of this but create new bottlenecks at Layer 4.

Connection: Layer 2 receives designs from Layer 1, consumes equipment from Applied Materials, ASML, Tokyo Electron, Lam, and KLA, takes memory from Layer 3 for stacking, and ships finished modules to Layer 6 for test. TSMC's quarterly CoWoS capacity number is, in my view, the single most important data point in AI capex forecasting.

Layer 3 — Memory

Memory is the bandwidth bottleneck. Every neuron in a transformer model has weights that must be loaded from memory into the compute fabric, used, and either written back or discarded. The math is brutal: a single trillion-parameter model in FP8 needs a terabyte of memory just to hold the weights, and inference requires reading those weights at rates measured in tens of terabytes per second. **HBM** — High Bandwidth Memory — is the solution. It stacks 8, 12, or 16 DRAM dies vertically using through-silicon vias, places the stack directly adjacent to the logic die on the same package, and delivers bandwidth that conventional DDR memory cannot approach.

Three companies make HBM at the scale required: **SK Hynix**, **Samsung**, and **Micron**. SK Hynix has been the technology leader for three consecutive generations — HBM3, HBM3E, and now HBM4 — and is Nvidia's primary supplier. Samsung has struggled with yield on the most recent nodes but remains a strategic dual-source. Micron is the smallest of the three but the only US-domiciled producer, which matters for CHIPS Act flows and US hyperscaler diversification. **SanDisk** and **Kioxia** sit adjacent to HBM in the NAND segment — flash memory for the storage tier — and are now positioning for **HBF**, High-Bandwidth Flash, which would extend the

bandwidth-storage hierarchy below HBM for inference workloads where capacity matters more than latency.

The engineering problem is twofold. First, stacking dies vertically with through-silicon vias is brutally hard — yield falls geometrically with stack height, so a 16-high HBM4 stack is much harder than two 8-highs. Second, the thermal management is severe: a 12-stack HBM3E module dissipates 15-20 watts in a volume smaller than a postage stamp, directly adjacent to a GPU dissipating 700-1000 watts. Heat has to come out without warping the stack or de-laminating the bonds. This is why hybrid bonding (Layer 4) and advanced thermal solutions (Layer 7) are so closely linked to HBM scaling.

The bottleneck is HBM4 capacity through 2027. SK Hynix has effectively sold out its 2026 production already. The pricing power is real — HBM ASPs are running 5-7x the equivalent commodity DRAM bit cost — and the margin pool for HBM3E and HBM4 is the single largest source of upside in the memory industry's history. Anyone who tells you memory is "commoditised" is fighting the last war.

Connection: Layer 3 sits on top of Layer 2's CoWoS substrate, is bonded to Layer 1's logic die, and is one of the three or four most fought-over inputs in the entire stack.

Layer 4 — Bonding, substrates, and photonics

This is the interconnect layer — how chips talk to other chips, how packages talk to other packages, and how racks talk to other racks. It is fragmenting into three sub-disciplines. **Hybrid bonding** — the precision die-to-die bonding required for HBM stacks and chiplet architectures — is dominated by **BE Semiconductor** (BESI) and **ASMPT**, two companies most retail investors have never heard of. Each shipped tool is essentially a custom-built machine, and the order books for both are running 12-18 months out. **Ibiden** and **Unimicron** dominate the ABF substrate market — the organic substrates that sit between the silicon package and the PCB — and Ibiden in particular has a near-monopoly on the highest-layer-count substrates required for Nvidia's largest packages.

The photonics half of Layer 4 is the rack-to-rack and intra-data-centre interconnect. Once you scale beyond a single GPU server, electrical signalling cannot carry the bandwidth required, and you have to convert to light. **Coherent** (COHR), **Lumentum** (LITE), and **Ciena** (CIEN) build the optical transceivers and DWDM systems that move bits between racks and between data centre buildings. **Corning** (GLW) sits adjacent here, both as the glass substrate supplier for next-generation packaging and as one of the world's primary optical fibre producers.

The engineering problem at Layer 4 is the bandwidth-power-distance trade-off. Copper interconnect maxes out at roughly 224 Gbps per lane over a few centimetres before signal integrity collapses. Above that, you need optical — but optical transceivers consume 15-30 watts each and a single rack might have 64 of them. Co-packaged optics (CPO), where the optical engine is integrated directly with the switch ASIC, is the next architectural step, and it changes the supplier mix meaningfully — Broadcom and Marvell are positioning aggressively here, as are Coherent and Lumentum.

Bottleneck: hybrid bonder lead times and ABF substrate capacity. BESI and ASMPT effectively split the hybrid bonding market between them with no credible third player, and Ibiden's highest-layer-count substrates run on six-month-plus lead times.

Connection: Layer 4 physically connects Layer 2's packaged chips and Layer 3's memory stacks, and electrically connects rack to rack inside Layer 0's data centres.

Layer 5 — Sub-component bottlenecks

This is the picks-and-shovels of the picks-and-shovels — the layer most investors miss and the layer with arguably the cleanest moats in the entire stack. A single AI server contains tens of thousands of **MLCCs** (multi-layer ceramic capacitors), and **Murata** of Kyoto controls roughly 40% of the global high-end MLCC market. There are perhaps three companies on Earth who can make the smallest, highest-capacitance MLCCs that the latest GPU power-delivery networks require, and Murata is the leader by a wide margin.

Corning, which we touched on at Layer 4, is also the dominant supplier of the **glass substrates** that will replace organic ABF substrates in the next generation of packaging. Glass has better dimensional stability, supports finer line widths, and handles thermal cycling better than organic — but it has to be flawlessly flat, defect-free, and laser-drillable, and Corning's Gorilla Glass heritage gives it a manufacturing lead measured in years over the next-best competitor.

Lasertec in Japan has a literal monopoly on **EUV mask inspection** — the tools that inspect the photomasks used in extreme-ultraviolet lithography. Without Lasertec's actinic inspection systems, you cannot reliably produce defect-free masks at 3nm and below. ASML cannot ship its EUV tools' output unless TSMC and Samsung can validate the masks, and they cannot validate the masks without Lasertec. There is no second supplier. This is a $5-8 billion business with one supplier, single-digit competitor share, and 50%+ operating margins.

Cadence and **Synopsys** are the **EDA duopoly** — the software platforms used to design every modern semiconductor. Every Nvidia chip, every AMD chip, every TPU, every Trainium — all designed in Cadence or Synopsys flows. The combined market is essentially a duopoly with high single-digit billion revenue each, mid-30s operating margins, and switching costs that are functionally infinite once a design team has trained on a flow. **Amphenol** completes the picture with high-speed connectors and cabling — the physical interconnect that links GPU trays to backplanes to rack switches. Inside every Nvidia NVL72 rack are tens of thousands of Amphenol contacts.

Bottleneck: each of these is a single-supplier or duopoly market. The economic structure of Layer 5 is the cleanest in the entire stack because the unit economics are tiny but the volumes are enormous and the customers have nowhere else to go.

Connection: Layer 5 sells into Layers 1, 2, 4, and 6, and the components touch every chip and every server shipped.

Layer 6 — Test and assembly

Every die that leaves Layer 2 must be tested. Every HBM stack from Layer 3 must be tested. Every assembled module must be tested again. Bad chips that ship into production are catastrophically expensive — a single failed GPU in a training cluster can cost hours of compute across an entire fabric — so the test step is non-negotiable, and the equipment is correspondingly expensive and consolidated.

Advantest of Japan is the dominant supplier of **SoC test equipment** for high-end logic, and effectively the sole supplier of high-end test for Nvidia GPUs and the leading custom ASICs. **Teradyne** is the global number two, stronger in memory test and in legacy SoC, and is repositioning

into AI test with the V93000 EXA Scale platform. **FormFactor** sits adjacent as the **probe card** supplier — the consumable interface between the test machine and the silicon wafer. **KLA** dominates **process control** — the inspection and metrology tools that catch defects during wafer fabrication, before the test step — and is the third leg of the semicap oligopoly alongside ASML and Applied Materials. **CAMT** and **Onto Innovation** (ONTO) play in the adjacent metrology and packaging-test segments, both with strong AI exposure.

Once tested, the packaging and final assembly is handled by the **OSAT** companies — outsourced semiconductor assembly and test. **ASE Technology** is the world leader, with **Amkor** as a close second. They handle the back-end packaging, final test, and shipment, and they are increasingly the operators of CoWoS-adjacent advanced packaging capacity in partnership with TSMC.

The engineering problem at Layer 6 is test time. Modern AI chips have so many transistors and so many functional units that exhaustive testing would take days per die. The art is in compressed test patterns — applying signals that exercise as many failure modes as possible in the shortest wall-clock time. Advantest's competitive advantage is software and pattern generation as much as it is hardware.

Bottleneck: high-end SoC tester capacity. Advantest has been capacity-constrained for two years and continues to add capacity, but the lead times on V93000 systems remain long. Probe card supply at FormFactor has similar dynamics — every new chip design requires a new probe card, and the design-to-delivery cycle is months.

Connection: Layer 6 sits between Layer 2 (manufacturing) and Layer 0 (deployment), and validates everything from Layers 1, 3, and 4.

Layer 7 — Power and thermal

This is where the physics gets brutal. A modern AI training cluster — say, a 100,000-GPU Blackwell deployment — draws roughly 150 megawatts continuous. That is the electrical demand of a mid-sized city, concentrated in a single building, running 24/7 at near-constant load. The grid was not designed for this. The next generation — Rubin Ultra clusters with a million-GPU horizons — will draw a gigawatt or more per site. Behind every AI training run is a chain of transformers, switchgear, and power-conditioning equipment that steps **765kV from the high-voltage grid down to 0.65V at the transistor**, with the final voltage regulation happening inches from the GPU die.

Constellation Energy (CEG) and **Vistra** (VST) are the primary nuclear and gas baseload generators positioned to sell directly into hyperscaler PPAs. CEG signed the Three Mile Island restart with Microsoft; Vistra has assembled a comparable portfolio. **GE Vernova** (GEV) builds the gas turbines and grid equipment that the next wave of merchant generation will require. **Quanta Services** (PWR) builds the high-voltage transmission lines that connect generation to data centres — and is now booked out years in advance. **Bloom Energy** (BE) sells solid-oxide fuel cells for behind-the-meter generation at the data centre site itself, an increasingly common workaround when grid interconnection queues stretch beyond five years.

Inside the building, **Vertiv** (VRT) is the leader in **liquid cooling** and the thermal management of high-density racks. Direct-to-chip and immersion cooling are no longer optional above roughly 50 kW per rack, and Blackwell racks already exceed that. **Vicor** (VICR) makes high-density DC-DC converters that handle the final voltage step-down at the rack and board level. **Modine** and **Powell Industries** play in the broader thermal and electrical infrastructure markets.

The engineering problem is power density. A Blackwell GPU dissipates close to a kilowatt in a few square centimetres of silicon. The heat flux is comparable to a nuclear reactor core, in volume terms. Conventional air cooling is dead at this point — liquid is mandatory — and the next architectural step is two-phase immersion, where the GPU is submerged in a dielectric fluid that boils on contact with the die. The companies positioned for this transition will earn outsized returns over the next five years.

Bottleneck: grid interconnection queues, gas turbine production capacity (GE Vernova is sold out into 2029), and high-voltage transmission construction. Power is the single tightest binding constraint on AI build-out for the next 24-36 months. Nvidia can ship chips. TSMC can package them. SK Hynix can supply memory. But if there is no power at the data centre site, none of it matters.

Connection: Layer 7 ultimately powers everything in Layers 0-6, and its constraints feed back upward as deployment delays at Layer 0.

The two cascades

There are two ways to think about how value flows through this stack, and you need both to make money. The first is the **demand cascade**, top-down. Dollars originate at Layer 0 — hyperscaler capex commitments — and propagate downward. A Microsoft commitment to spend $90 billion in 2026 translates into Nvidia orders at Layer 1, which translate into TSMC tape-outs at Layer 2, which translate into HBM allocations at SK Hynix at Layer 3, into BESI hybrid bonder orders at Layer 4, into Murata MLCC and Lasertec mask-inspection consumption at Layer 5, into Advantest tester slots at Layer 6, and into Vertiv liquid cooling and Constellation PPA revenue at Layer 7. The demand cascade is generally smooth and predictable on a 6-12 month forward basis once hyperscaler guidance is in. It is the framing that sell-side models use, and it is the framing that explains why every layer of the stack is up materially over the last 18 months.

The second cascade — and the one that actually drives stock-level returns — is the **bottleneck cascade**, bottom-up. The binding constraint determines the marginal price across the entire chain. If Layer 7 power is the choke point, then GPU sales are gated by deployment readiness, and the marginal dollar of upside accrues to Constellation, Vistra, GE Vernova, and Quanta. If Layer 5 EUV mask inspection is the choke point, the marginal dollar accrues to Lasertec — and you can have $600 billion of hyperscaler capex chasing it, but the cap on output is set by how many Lasertec tools ship that year. If Layer 3 HBM4 yield is the choke point, the marginal dollar accrues to SK Hynix. If Layer 2 CoWoS is the choke point, it accrues to TSMC and the advanced packaging equipment vendors who serve it. **Who limits whom** is the question, and it rotates every six to nine months as new capacity comes online in one layer and pushes the binding constraint to the next.

The interplay matters. In 2024, the binding constraint was HBM3E — SK Hynix was the bottleneck and the stock was the cleanest expression of AI scarcity. In 2025, it rotated to CoWoS — TSMC and its packaging vendors took the baton. In 2026, the binding constraint is splitting between HBM4 (still Layer 3), advanced bonding (Layer 4), and power (Layer 7). By 2027, on current trajectory, power will be the dominant constraint and the trade rotates decisively to Layer 7. The most important question we ask in the deep dives that follow is always: *where is the binding constraint in this layer over the next 12-18 months, and which name has the cleanest exposure to it?*

How to read the rest of this document

Part II ranks every name we cover into three tiers. **Tier A** is highest conviction — names where the moat, the bottleneck exposure, and the valuation align, and where we would be comfortable taking 3-5% portfolio positions. **Tier B** is solid second-line exposure — credible businesses with real AI leverage but either a less defensible moat, a more crowded competitive set, or richer valuation that requires more patience. **Tier C** is optionality and watch-list — names with genuine AI exposure but where the case is earlier, more speculative, or contingent on a specific catalyst we are tracking. Within each tier, we organise by layer, working downward from Layer 0 hyperscalers through Layer 7 power and thermal, so that you can read the document either by conviction (start at Tier A and stop wherever you run out of capital) or by layer (read the entire compute fabric section if that is the box you have to fill). Either way works. The deep dives are written in the Photoncap institutional template — ten sections per name, all the same structure — so you can compare across names in a glance once you know the format. Welcome to the syllabus. The exam is the P&L.

*See **Figure 1 — Picks and Shovels by Layer** (below) for the concentration / moat-depth visualisation of each layer.*

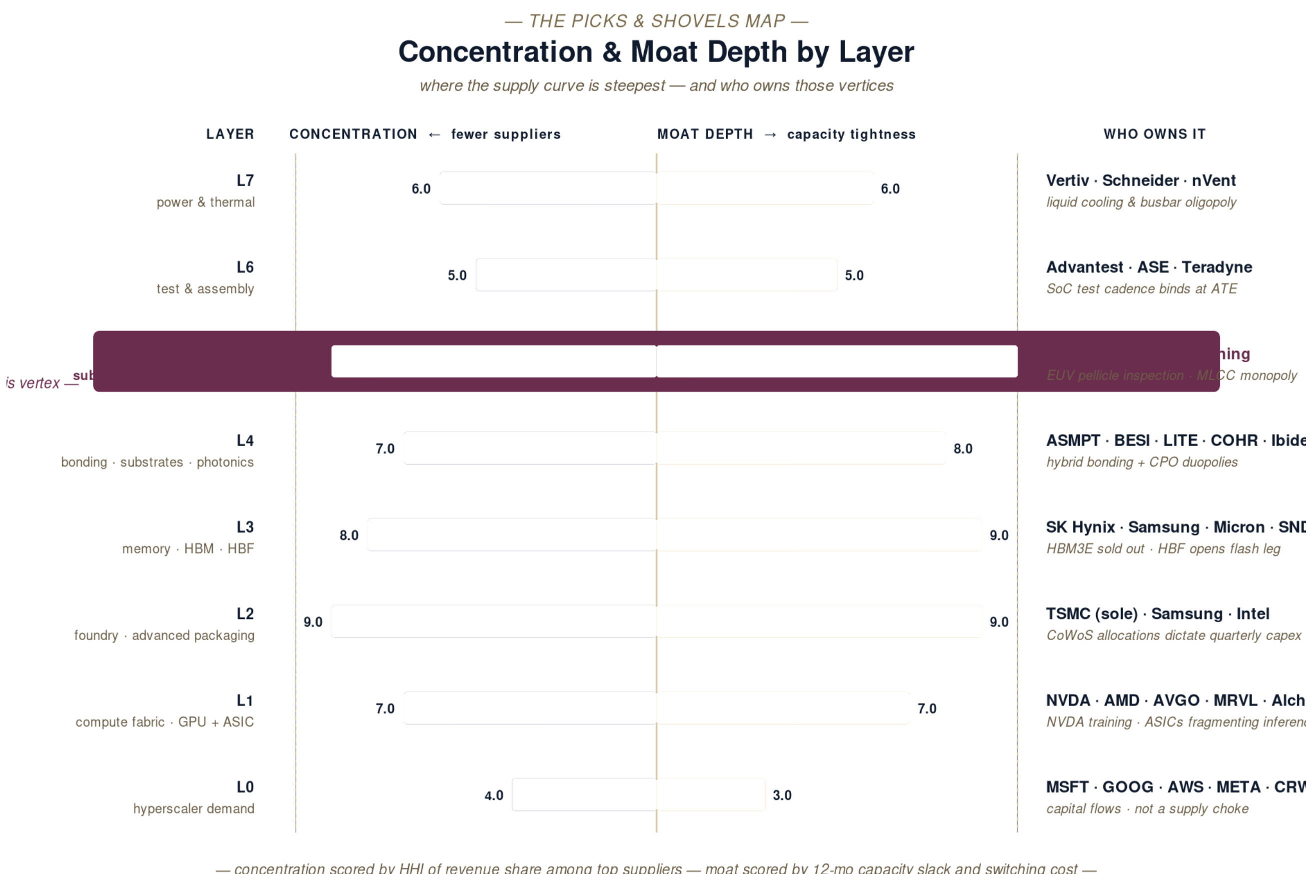

Figure 1: Figure 1 — Picks and Shovels by Layer

Figure 1 — Picks and Shovels by Layer. Concentration and moat-depth across L0–L7, with the named winners on the right.

Part 2 — Three Transitions Reshaping the Stack

The eight-layer stack we just walked through is not a static picture. It is a moving target — and the names that win at each layer change as the *workload itself* changes. The mistake most generalist investors are making in 2026 is mentally freezing the 2023 trade: long NVIDIA, long HBM3E, long anything-with-a-GPU-attached. That trade worked. It is also, increasingly, the past.

What follows is the framework we use internally to think about *which* picks-and-shovels matter at which point in the cycle. There are three transitions in motion right now — each on a different clock, each reshuffling the winners and losers within the same eight-layer stack. Get the transitions right and you get the rotations right. Get them wrong and you end up holding the last war's ammunition.

Transition 1 — Training to Inference (already underway, 2024-25)

The first transition is the one most investors have at least *heard* about, though far fewer have actually repositioned for it. Foundation-model training — the part of the AI workload that built NVIDIA into a four-trillion-dollar company — peaked in 2024. That is not a controversial statement anymore; it is increasingly the consensus of the hyperscaler capex committees themselves. Three forces converged. First, the absolute dollar cost of frontier training runs broke through the ten-billion-dollar barrier per model and CFOs started asking uncomfortable questions about ROI. Second, the scaling laws began to bend — Chinchilla-optimal, then post-Chinchilla, then the GPT-4.5 plateau, and finally the public DeepSeek moment in early 2025 when a Chinese lab demonstrated that frontier capability could be reached with roughly a tenth of the compute budget through algorithmic efficiency rather than brute force. Third, and most importantly, the *output* of training — the trained model — started being used. Heavily. By hundreds of millions of users running billions of inference calls per day.

The result is a workload mix that looked roughly 70/30 training-to-inference at the hyperscalers in 2023, is roughly 40/60 today, and will be 20/80 by 2027. Inference workloads are now growing five to ten times faster than training workloads. This is not a forecast — it is what the hyperscaler capex disclosures are already saying when you read them closely.

The investment consequence is that the two workloads have *fundamentally different hardware requirements*, and the market has been slow to internalize this. Training is latency-tolerant — a run takes weeks, nobody cares about a millisecond. Training is batch-friendly — you feed enormous batches through the network and amortize fixed costs. Training is throughput-king — what matters is FLOPS-per-dollar at the cluster level, which is why mega-clusters of H100s and H200s with exotic HBM3E bandwidth and InfiniBand fabric won the 2023-24 cycle. NVIDIA's moat in training is real and not going anywhere soon.

Inference is the opposite animal. It is latency-critical — a chatbot that takes four seconds to respond is a dead chatbot. It is request-by-request — you cannot wait around to batch a thousand queries because the user is waiting. It is energy-cost-sensitive — when you are serving a billion requests a day, every joule per token compounds into the operating-cost line of every hyperscaler

P&L. And it rewards a completely different silicon profile: not raw FLOPS but FLOPS-per-watt, not maximum bandwidth but bandwidth-at-the-right-tier, not flexibility but specialization.

This is why the custom-ASIC story matters so much in 2026 and why we have been front-running it through the design-house ecosystem. Amazon's Trainium 2 — designed in partnership with Alchip in Taiwan — is now serving a meaningful slice of Anthropic's inference workload. Google's TPU v5 and v6 — designed with Broadcom — are doing the same for Gemini. Microsoft's MAIA — co-designed with GUC and others — is ramping for OpenAI inference. Meta's MTIA is in production for recommendation inference. Each of these chips is *worse* than an H100 at training. Each of them is *better* than an H100 at inference for the specific workload it was designed for, on a cost-per-token basis. That is the entire point.

The picks-and-shovels reshuffle from this transition runs as follows. On the losing side: pure training-spec accelerators with no inference roadmap, marginal HBM3E suppliers whose only pitch is raw bandwidth, and the second-tier GPU cloud rentals that built businesses on training arbitrage. On the winning side: ASIC design-services houses where we have core positions — Alchip (3661 TT) and GUC (3443 TT) in Taiwan, Astera Labs (ALAB) for the connectivity layer between custom silicon and memory. Inference-optimized memory architectures — High-Bandwidth Flash, SLC NAND with much higher endurance and lower latency than standard TLC, which is precisely what makes SanDisk (SNDK), Western Digital (WDC), and Kioxia interesting in a way they have not been in a decade. Power-efficiency specialists, because inference cost is dominated by energy at scale — Navitas (NVTS) in gallium nitride, Vicor (VICR) in 48V-to-core power delivery, Power Integrations (POWI) in efficient AC-DC conversion. Advanced packaging for the high-batch inference racks where chiplet integration matters — BESI (BESI NA) in hybrid bonding, ASMPT (522 HK) in thermo-compression bonding.

And the survivors that adapt — NVIDIA itself with Blackwell and Rubin both having strong inference profiles, SK Hynix pivoting HBM3E volume toward HBM4 inference variants, TSMC fabricating every single one of these new ASICs. The point is not that the old leaders die. The point is that the *marginal dollar* of capex shifts decisively, and the names with the highest marginal-dollar leverage outperform.

*See **Figure 2 — Training vs Inference Bottleneck Map** for where the heat moves between eras.*

Figure 2 — Training vs Inference Bottleneck Map. Layer-by-layer heat-shift between the Training Era (2023-25) and the Inference Era (2026+).

Transition 2 — Inference to Agentic AI (happening now, 2026)

The second transition is the one we are living through right this minute, and it is the one that most investors are *still* mispricing because it does not have a clean tickertape moment the way ChatGPT did in late 2022. Agentic AI is not a product launch. It is a workload pattern shift — and like all workload pattern shifts, it shows up first in the infrastructure capex line and only later in the consumer narrative.

The shift in plain language: a 2024-vintage AI call was *single-shot*. You typed a question, the model produced an answer, transaction complete — maybe a thousand tokens of output, maybe two seconds of compute time. A 2026-vintage AI call is increasingly an *agent task*. You give

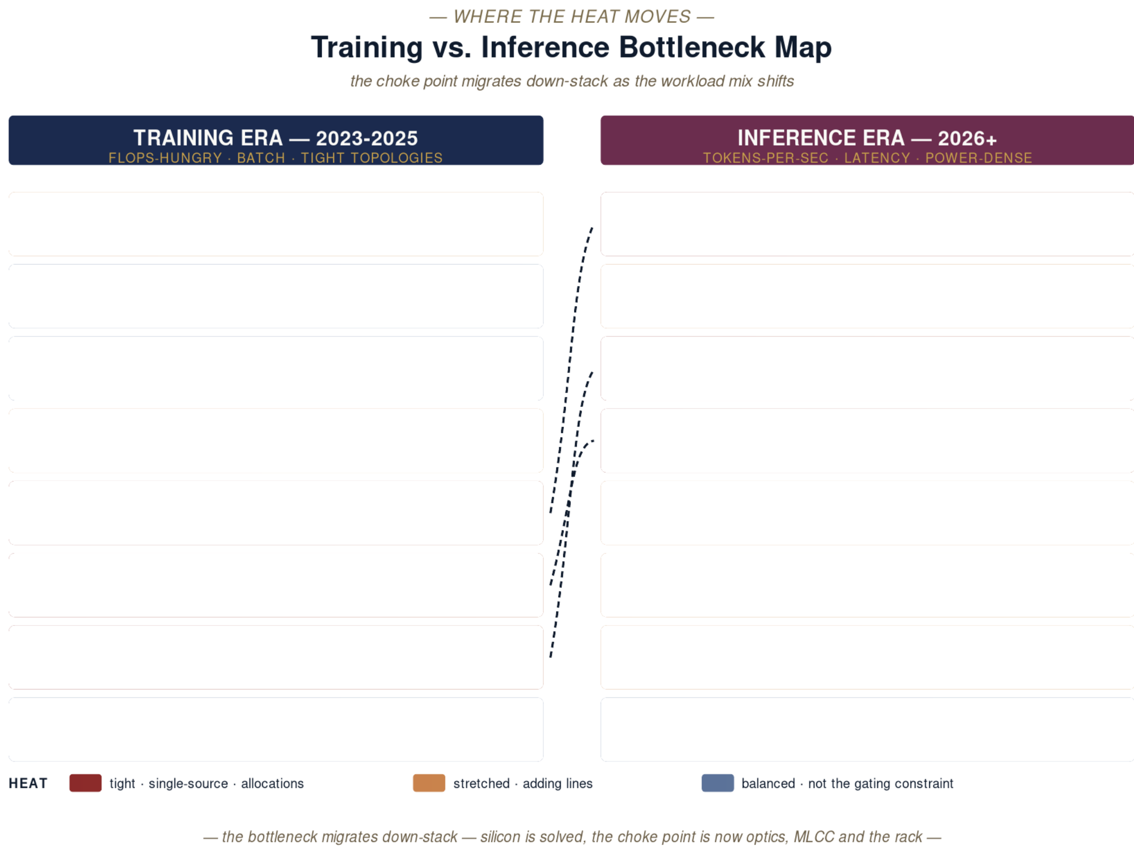

Figure 2: Figure 2 — Training vs Inference Bottleneck Map

the agent a goal — "research this company, build me a model, send the draft to my email" — and the agent decomposes the goal into sub-tasks, calls the model dozens of times to reason about each, calls external tools (search, code execution, file I/O) between each reasoning step, evaluates intermediate results, replans, retries, and eventually completes. A single user-facing "task" that takes five seconds of wall-clock time may consume one hundred times more tokens than the equivalent ChatGPT response from 2024. Anthropic's own published data on Claude Code workloads bears this out — agent sessions routinely consume hundreds of thousands to millions of tokens per task.

The infrastructure implication is brutal in its simplicity: ten-to-one-hundred times more inference compute per user-facing task. That is not a forecast either — it is what the token-volume disclosures from OpenAI, Anthropic, and Google are now showing in their developer-API logs, and it is what the hyperscaler inference capex lines for 2026 are sized for. The market has priced *more* inference. The market has not yet priced *one hundred times* more inference.

But the economics shift in a more subtle way too. In the single-shot world, cost-per-token was the only metric that mattered. In the agentic world, *cost-per-completed-task* is what matters — and that depends not just on per-token cost but on how efficiently the inference fabric can chain hundreds of sequential and parallel calls without thrashing on memory or interconnect latency. An agent that takes ten seconds to complete is twice as profitable as one that takes twenty seconds, because the underlying compute is amortized over half the wall-clock time and the user is twice as likely to come back. *Latency at the rack scale becomes a P&L line item.*

This is what is driving the photonics moment that we have been positioned heavily into. When you have to move terabytes per second between thousands of accelerators inside a single rack, with sub-microsecond tail latency, copper does not work anymore. Optical interconnect at the rack scale — co-packaged optics, linear-drive pluggable optics, near-package optical engines — moves from niche to mandatory. Lumentum (LITE) and Coherent (COHR) supply the laser engines and transceivers. Ciena (CIEN) supplies the metro and DCI optical layer that connects the racks to each other. Applied Optoelectronics (AAOI) plays in the 800G and 1.6T transceiver tier where the hyperscale ramp is steepest. Fabrinet (FN) does the precision assembly. This is a multi-year structural growth wave that is independent of the GPU cycle — even if Blackwell volumes plateau, the interconnect tier keeps growing because *the chips are being asked to talk to each other more*.

Distributed memory tells the same story. When an agent's working context — its scratchpad of intermediate results across a long task — exceeds what fits on a single GPU's HBM, you need a memory tier that lives between HBM and main DRAM. That is what Compute Express Link is for, and that is also why High-Bandwidth Flash from SanDisk-Kioxia is suddenly interesting — it provides a persistent, addressable, much-cheaper-per-bit tier that sits *just below* HBM in the hierarchy. Astera Labs (ALAB) sits in the middle of the CXL switching layer; SanDisk (SNDK) and Kioxia provide the flash.

Power density at the rack scale is the other binding constraint. Agentic workloads push average rack power from 30-40kW (2023) toward 130-200kW (2026 Blackwell racks) and 500kW+ (2027 Rubin and beyond). Air cooling is dead at those densities — physics, not preference. Vertiv (VRT) is the single largest pure-play in liquid cooling and rack-scale power; Modine (MOD) plays the heat-rejection side with coolant distribution units. Eaton (ETN) and Schneider provide the switchgear. And upstream of all of it, the actual electrons have to come from somewhere — which is why Constellation Energy (CEG) and Vistra (VST) and the rest of the regulated-nuclear-and-gas complex have been re-rated. The hyperscalers are now signing twenty-year PPAs at premium

prices because the binding constraint on their 2027-28 capacity build is not GPUs, not memory, not optics — it is megawatts at the substation.

And memory itself moves up a generation. HBM3E was the training-era memory. HBM4 — with stacked compute logic and through-silicon-via density that approaches the limits of physics — is the agentic-era memory. SK Hynix is the volume leader; Samsung is fighting to catch up; Micron (MU) holds the third slot. The HBM4 winner — and right now it is plainly SK Hynix — captures a disproportionate share of the agentic-inference value chain because every accelerator in every Rubin-class rack will be carrying its silicon.

The losers in this transition are easier to name than they are to short: dense GPU clusters built without rack-scale optical interconnect look like stranded assets the moment Rubin-class fabric becomes the standard. Air-cooled data-center capacity that cannot retrofit to liquid is similarly impaired. Pure-play DRAM exposure without an HBM4 pipeline is the wrong place to be. And the second-tier neocloud GPU rentals that built businesses on raw H100 throughput — many of them are about to discover their economics do not work at agentic-era cost-per-task.

*See **Figure 3 — The Agentic AI Stack** — the new orchestration layer and which existing layers feel the largest workload impact.*

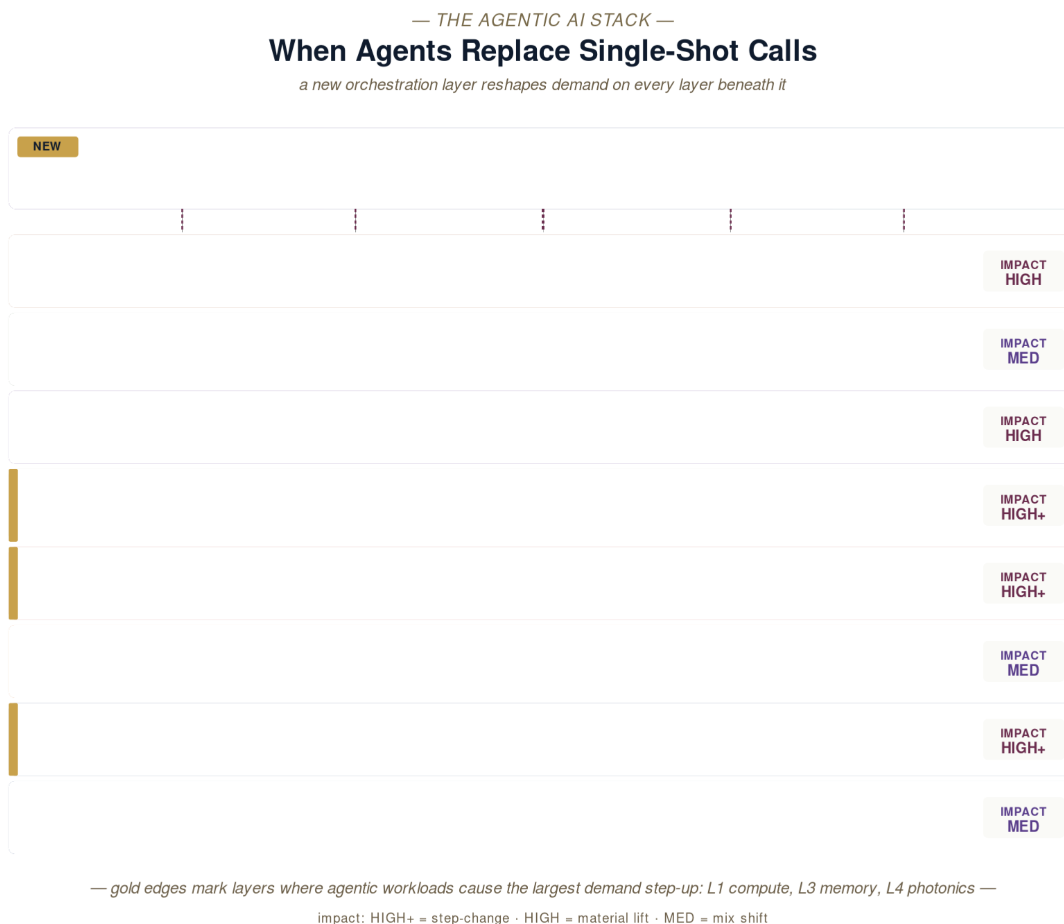

Figure 3: Figure 3 — The Agentic AI Stack

Figure 3 — The Agentic AI Stack. The new orchestration tier above the 8-layer stack and how

agentic workloads reshape demand at each layer beneath.

Transition 3 — Agentic to Embodied / Physical AI (early innings, 2027-2030)

The third transition is the one nobody can size accurately and everybody should be paying attention to anyway, because it is where the next ten-bagger names get made. Embodied AI — physical AI — is the move out of the data center and into the world. Humanoid robotics: Tesla's Optimus, Figure AI, 1X, China's UBTech with the Walker S2 and the 9880-series industrial humanoid. Autonomous vehicles: next-generation systems built on much larger neural networks running at the edge in real time, with Waymo, Tesla FSD, and the Chinese L4 fleet operators (Pony, WeRide, Baidu Apollo) all converging on similar architectures. Industrial automation that is not just programmed but *trained*. Drones. Surgical robots. The full physical-world deployment of capabilities that until recently lived inside a chat window.

The right way to think about this for the portfolio is not as a 2026 trade. It is a *call option* with a long runway — 2027 ramp, 2028-29 commercial volume, 2030+ for anything resembling mass deployment. We are positioned in size in a couple of names and tactically in the rest. But the *framework* matters now, because the early infrastructure orders are being placed now, and the supply chains that win embodied AI are being chosen now.

The defining feature of embodied AI from an infrastructure standpoint is that *inference moves to the edge*. A humanoid robot cannot wait two-hundred milliseconds for a cloud round-trip to decide whether to put its foot down. An autonomous vehicle absolutely cannot. The decision has to happen on-device, in real time, with safety-critical reliability — meaning the silicon, memory, sensors, and power conversion that make embodied AI work are *not* the data-center stack scaled down. They are a fundamentally different stack: ruggedized, miniaturized, power-budgeted, often custom.

And the volumes are different. There are perhaps fifty thousand AI training accelerators sold per quarter in the data-center world. There are *hundreds of millions* of edge-AI endpoints in the embodied-AI world if you take a five-to-ten-year view — every robot, every car, every drone, every smart industrial actuator. Even at a tenth the per-unit margin, the unit volumes are a thousand times larger. The TAM math is enormous; the timing math is the part that requires patience.

The picks-and-shovels in the embodied-AI stack break into several distinct buckets. Edge-AI silicon: Qualcomm (QCOM) is the obvious volume leader given its smartphone-era manufacturing scale and existing automotive ramp via the Snapdragon Ride platform; ARM Holdings (ARM) earns a royalty on essentially every embedded AI inference chip shipped globally, including the custom edge ASICs that the robotics OEMs are designing in-house; NVIDIA's Jetson and Drive platforms hold the high-end edge for robotics development. Power conversion miniaturization — when your AI processor lives inside a humanoid's chest cavity, power density becomes everything — Power Integrations (POWI) in efficient AC-DC and DC-DC at small form factor, Vicor (VICR) in 48V power architectures that have already won in the data center and now port directly into robotics, Navitas (NVTS) in gallium-nitride that delivers higher switching frequency at lower thermal cost, exactly what a battery-powered humanoid needs. Specialized sensors — LiDAR, vision, IMU, tactile — where the supply chain is more fragmented but Sony (image sensors), Aeva and Hesai (LiDAR), and the legacy automotive Tier-1s all have roles. Ruggedized memory — embodied systems run in vibration, in heat, in cold, in dust — SLC NAND from SanDisk and Kioxia

is the workhorse here because it withstands the duty cycles that consumer TLC cannot.

And then there are the *physical-input* names — the compound semiconductors and specialty materials that the embodied-AI build-out actually consumes in volume. Silicon carbide and gallium nitride for the high-voltage motor drives and onboard power — Wolfspeed (WOLF) if it survives its current restructuring, ON Semi (ON), Infineon, Soitec (SOI FP) which supplies the SiC and SOI substrates upstream of the device makers, AXT Inc (AXT) in indium-phosphide and gallium-arsenide for photonics and RF. Rare-earth permanent magnets for the actuators — every humanoid joint, every EV motor, every drone propulsion system consumes neodymium-iron-boron magnets, and the supply chain outside China is a multi-year strategic priority for both the US and the EU, which is part of the rare-earths trade we have been building. Precision optics for vision systems. Industrial-grade connectors — Amphenol (APH) and TE Connectivity (TEL) are not just data-center names; they are humanoid-and-EV names too.

The caveat is the timeline. Optimus volume in 2026 is small. Figure AI is pre-commercial. UBTech is selling industrial humanoids in low single-digit thousands. Waymo is operating, scaling, but still small. Putting size into embodied-AI pure-plays today is paying a 2030 valuation for a 2027 revenue line, and that requires patience that not every portfolio has. We treat the embodied-AI bucket the way a venture investor treats a call-option portfolio: many small bets, a few core convictions, and the discipline not to confuse them with the cash-flowing names in Transitions 1 and 2.

Cross-cutting observation — who wins in *every* transition

Here is where the framework yields its most useful insight, and the one we lean on most heavily in portfolio construction. There is a small group of names that *win in every transition* — names where it does not matter whether the dominant workload is training, inference, agentic, or embodied, because they sit at sub-component bottlenecks that every workload type still needs.

Corning (GLW) makes the glass substrate that goes into the optical-fibre interconnect for hyperscale data centers, the cover glass and precision optics for embodied-AI vision systems, the substrate for next-generation semiconductor packaging via its EXTREME Ultra-Low Loss product line. Every transition uses Corning glass.

Amphenol (APH) makes the high-speed interconnect — copper, optical, RF — that goes into every server, every accelerator board, every robotics platform, every EV powertrain. Every transition uses Amphenol connectors.

Cadence (CDNS) and Synopsys (SNPS) provide the electronic-design-automation software without which *no* AI chip — training, inference, edge, custom, merchant — can be designed. Every transition uses their EDA. The pricing power is structural; the customer concentration risk is essentially zero because the customer is the entire global semiconductor industry.

ASML (ASML) supplies the extreme-ultraviolet lithography tools that fabricate every leading-edge node. There is one supplier. There is no second supplier. Every transition uses ASML.

TSMC (TSM) fabricates essentially every AI chip worth fabricating at the leading edge — NVIDIA's GPUs, AMD's MI series, every hyperscaler's custom ASIC, the major edge-AI SoCs. The Taiwan concentration risk is real and is the only thing that keeps the position from being even larger. Every transition uses TSMC.

KLA Corporation (KLAC) provides the process-control and inspection equipment that every fab — TSMC, Samsung, Intel, SK Hynix, Micron — needs to run leading-edge nodes at yield. Every transition uses KLA.

These are the *infrastructure compounders*, and they share a common structural feature: they sit at points in the stack where the bottleneck is not workload-specific. The workload mix can shift completely — training to inference to agentic to embodied — and the bottleneck stays the same. Glass is still glass. EDA is still EDA. Lithography is still lithography. The substrate of compute does not care which application layer is running on top of it.

This is the deepest moat in the entire AI stack, and it is where the highest-conviction, longest-duration positions in the portfolio sit. These are the names we size aggressively and rebalance rarely. Everything above them in the stack is a rotation. They are the *substrate*.

Implications — how to position the book

The portfolio prescription that falls out of this framework is, we think, straightforward, though it cuts against a lot of consensus exposure.

First, *trim* aggressive training-spec exposure. NVIDIA stays a core position because its inference roadmap is strong and its CUDA moat is real — but the second-derivative names that are *only* about training-cluster build-out should be reduced. Marginal HBM3E specialists. Pure InfiniBand-fabric plays without a photonics roadmap. Single-product GPU cloud rentals.

Second, *lean into* the inference and agentic stack. That means: ASIC design houses (Alchip, GUC, Astera Labs), inference-optimized memory (SanDisk, Western Digital, Kioxia, plus SK Hynix on HBM4), rack-scale optical (Lumentum, Coherent, Ciena, Applied Optoelectronics, Fabrinet), liquid cooling and rack power (Vertiv, Modine), and the power-supply tier upstream (Constellation, Vistra, plus regulated utilities where the PPAs make sense). This is where the marginal hyperscaler dollar is going in 2026 and 2027.

Third, *build* the embodied-AI optionality patiently. Edge-AI silicon (Qualcomm, ARM), power miniaturization (Power Integrations, Vicor, Navitas), compound semis (Soitec, AXT, Wolfspeed selectively), specialized sensors. Position sizes should be smaller than the agentic-stack names; runway is longer; conviction is real but timing is the binding uncertainty.

Fourth, *size and sleep* on the cross-cutting compounders. Corning, Amphenol, Cadence, Synopsys, ASML, TSMC, KLA. These are not trading names. They are the underlying bet that AI as a category — whichever workload dominates in any given quarter — continues to need silicon, glass, light, and software. Re-rating risk is the only meaningful risk on any of them, and that risk is symmetric.

Fifth and finally — watch for the next transition. Our working hypothesis is that the 2028-30 window brings a *fourth* phase that the current framework only hints at: federated agentic AI, where embodied endpoints and cloud agents coordinate in real time through edge-cloud hybrid inference fabrics; or possibly the merger of agentic and embodied into a single hybrid workload pattern where a humanoid in a warehouse and a cloud agent in a data center are running pieces of the same task. Whichever way it breaks, it will reshuffle the picks-and-shovels again, and the analysts who are already thinking about it in 2026 will be ahead of the rotation when it comes.

Train on the past. Position for the future. The transitions are not optional; they are the entire shape of the next decade of AI infrastructure investment. The job is to be early enough to capture them and disciplined enough not to confuse them with each other.

*See **Figure 4 — LLM → Agentic → Embodied Timeline** — the three transitions sequenced with bottleneck heatmaps per phase.*

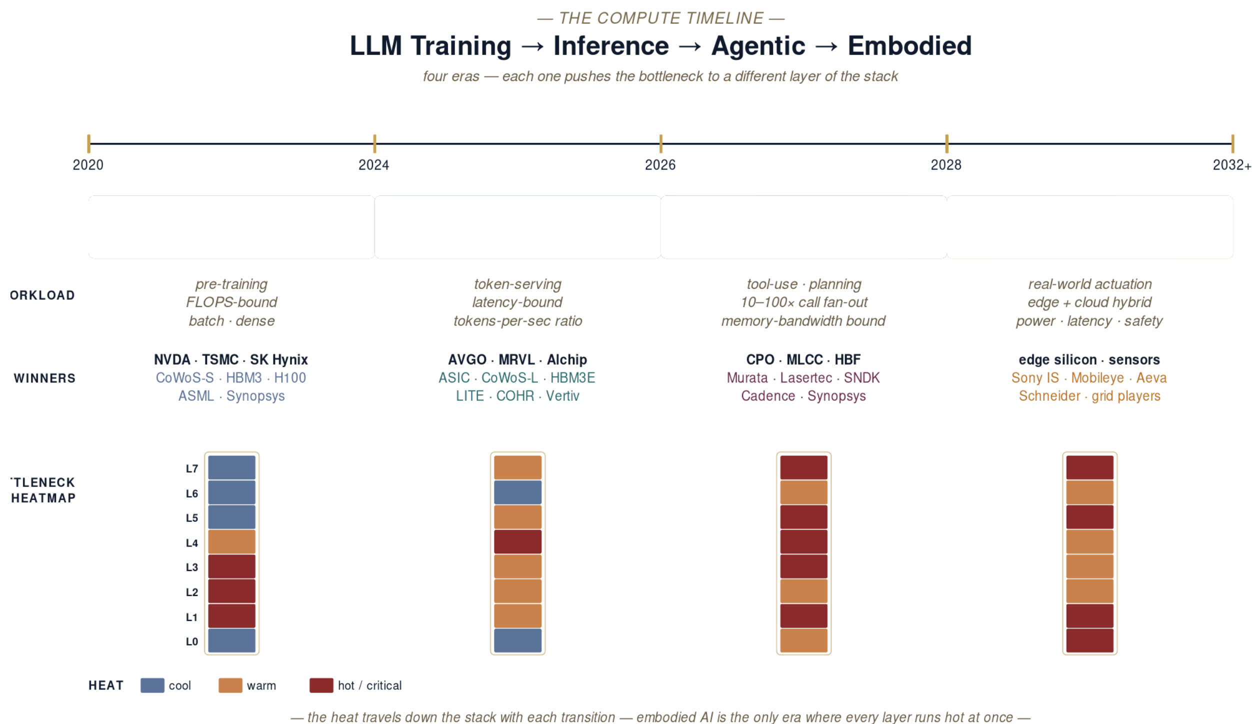

Figure 4: Figure 4 — LLM to Agentic to Embodied Timeline

Figure 4 — LLM → Agentic → Embodied Timeline. The three transitions sequenced across 2024-2030 with per-phase bottleneck heat overlays.

Part 2b — NVIDIA Q1 FY27 · Physical AI as the Next Wave

Empirical proof point for Transition 3. Six days after the Robotics-MASTER deep-dive was filed, NVIDIA quantified the physical-AI wave on its Q1 FY27 print. The numbers and the language below are what changed.

Headline

NVIDIA printed $81.6B of revenue (+85% YoY, +20% QoQ), beat consensus by ~$2.4B, guided Q2 to $91B vs $86B consensus, raised the dividend 25x and added $80B of buyback. Data Center hit $75.2B (+92% YoY), with networking alone at $14.8B (+199% YoY). The clean print is not the story. The story is the second sentence of Jensen's closing remarks: **"Physical AI is the next wave."** For the first time, NVIDIA quantified that wave on the call — physical-AI revenue

exceeded $9B over the trailing twelve months — and introduced a new Edge Computing segment ($6.4B in Q1, +29% YoY) that bundles PCs, consoles, robotics and automotive into a single number investors can finally track.

Macro summary

The agentic-AI inflection that powered this quarter is now treated as the established baseline. Huang explicitly defined two adjacent waves powering NVIDIA forward: **(1) agentic AI** (“demand has gone parabolic … agentic AI has arrived”) and **(2) robotic / physical AI** (“the world is rebuilding computing for agentic AI and robotic physical AI. Nvidia sits at the center of these transitions”). Blackwell GB300 / NVL72 demand was described as “particularly strong.” Vera Rubin entered full production in the quarter, with partner availability scheduled 2H 2026. The $1T Blackwell + Rubin revenue envelope through CY27 was reiterated. China remains excluded from the guide. The Vera CPU was flagged as a “$200B opportunity.” But the strategic centre of gravity of the call shifted: physical AI is no longer the post-credits scene — it has a revenue line, a segment, a product stack (Cosmos, Isaac GR00T N / N1.5, IGX Thor, Alpamayo, Omniverse NuRec) and a flagship customer (Uber, ~30 cities, 4 continents, by 2028).

The five-bullet Physical-AI thesis

1. **It’s a real revenue line now, not just a slide.** $9B TTM physical-AI revenue. That is bigger than the entire historical NVIDIA Automotive segment annualised and growing inside the new Edge Computing reporting block.
2. **The full software stack is in market.** Cosmos (world model) → Isaac GR00T N / N1.5 (humanoid foundation model) → Isaac Sim (training) → IGX Thor / Jetson Thor (on-robot compute) → Drive Thor (AV compute). Mirrors the CUDA → DGX → Blackwell vertical play.
3. **There is a dated commercial proof point.** Uber robotaxi: ~30 cities, 4 continents, by 2028. This converts physical AI from “decade-long option” to “scheduled deployment with an end date attached.”
4. **NVIDIA is going horizontal, not vertical.** Industrial + surgical + humanoid + AV. They are not picking a robot — they are arming everyone who builds one. Same playbook as data center.
5. **Edge Computing is now the segment investors should watch.** $6.4B / +29% YoY / +10% QoQ. Still small (~8% of total) but reframed as the physical-AI revenue proxy. Acceleration in Edge sequentials is the leading indicator that the wave has crested into the P&L.

What this means for picks-and-shovels

The cleanest direct read-through is to **HBM + advanced packaging**. The $1T Blackwell + Rubin envelope through CY27 already implied the largest HBM pull in industry history. Physical AI extends the cycle by another node — every Isaac/GR00T training run, every Cosmos world model, every Omniverse NuRec AV simulation lives on the same Blackwell/Rubin systems. SK Hynix, Micron and Samsung remain the only three vendors who can deliver HBM3E/HBM4 at this volume. TSMC + CoWoS remain the only place these systems get built. None of these names need to invent a physical-AI thesis — they ride the same Blackwell/Rubin curve, just for longer.

The second read-through is **AI networking and optics**. $14.8B of networking revenue at +199% YoY tells you the cycle bottleneck has moved from compute to interconnect. Spectrum-X at $10B

annualised, InfiniBand revenue roughly doubling QoQ. Physical AI makes this worse — robotaxi fleets, factory-floor humanoids and edge inferencing all push for smaller, lower-latency, more distributed networking fabrics. Arista, Coherent, Lumentum, Marvell and the Chinese transceiver pair (Eoptolink, Innolight) are direct beneficiaries. NVIDIA's own switching and optics roadmap (Spectrum-X, NVLink Switch, future co-packaged optics) becomes a quasi-monopoly inside the AI factory.

The third read-through — and the most interesting, because it is the most underweighted in the existing master file — is **physical-AI deployment infrastructure**: robotics simulation tools, sensors, motion, and the OEMs themselves. NVIDIA has explicitly horizontalised the foundation model (GR00T is open) and the compute (Jetson/IGX Thor). That means value capture in the next 24 months sits with the OEMs who win the form factor (Figure, Agility, 1X, Boston Dynamics, Apptronik on the humanoid side; Hyundai, Volvo, Mercedes on the AV side) and with the picks-and-shovels they need to build (Cognex/Keyence vision, Teradyne/Universal Robots arms, Rockwell automation, SiC power management for actuators, lidar). For the master file, this is the seam where we should be adding coverage — not by adding more NVIDIA exposure, which the portfolio already has, but by mapping the next two layers down.

What's NOT in this call but should be

Three things were conspicuously absent and should be flagged as known unknowns. **First**, Huang gave no hard physical-AI TAM number — no "$1 trillion humanoid market by 2030," no "labour-replacement multiple of GDP." The $50T / "65% of world GDP" framing investors associate with him is from a May 2025 interview, NOT this call. Anyone modelling a discrete physical-AI TAM is reaching beyond what management actually said. **Second**, no named humanoid OEM partners (Figure, 1X, Sanctuary, Apptronik) appeared in any recap of the prepared remarks. Boston Dynamics, Agility and Skild are referenced in adjacent NVIDIA ecosystem material around the print but not confirmed as named on the call itself. The implication: NVIDIA is keeping humanoid partner exclusivity options open and is not yet picking a winner. **Third**, no margin or unit-economics commentary on Edge Computing was offered — investors do not yet know whether physical AI is a 75%-gross-margin business like Data Center or a sub-50% margin pull on the model. Until they tell us, treat the $9B TTM as a top-line marker only, not a profit pool.

Part 2c — Physical AI · Cross-Layer Theme Linking AI Semis + Robotics

Where Part 2b describes the print, this section integrates physical AI into the 8-layer stack we built in Part 1 and reconciles it with the existing Robotics-MASTER coverage. The conclusion: physical AI does not displace the stack — it overlays a second demand curve, layer by layer, with most of the mix-shift work concentrated at L5.

Why this matters now

For three years, "robotics" has been the conference-slide layer of the AI bull case — the part Jensen Huang gestures toward in the last five minutes of a keynote, the part sell-side models

treat as a free option. On May 20, that posture changed. In the Q1 FY27 call, NVIDIA put a number on it: **physical-AI revenue exceeded $9 billion over the trailing twelve months**, and the company restructured its segment reporting around it. Two reportable segments now — Data Center and a newly broken-out Edge Computing — and Edge (PCs, consoles, robotics, automotive) printed **$6.4B in Q1, +29% YoY, +10% QoQ**. That is bigger than the entire historical NVIDIA Automotive segment annualised many times over. It is also the first time anyone outside management has a quarterly proxy to track the physical-AI ramp on a P&L cadence.

Jensen's framing was deliberate. Closing remarks: *"The world is rebuilding computing for agentic AI and robotic physical AI. Nvidia sits at the centre of these transitions."* And the bridge line: *"Physical AI is the next wave."* The implication is that NVIDIA management is no longer treating physical AI as a 2030 option — it is treating it as the adjacent S-curve to agentic AI, scheduled to scale through the back half of this decade. The flagship commercial proof point: an **Uber robotaxi rollout across roughly 30 cities and 4 continents by 2028**, powered by NVIDIA Drive Thor. Concrete deployment, concrete dates, concrete partner. That moves robotaxi from "decade-long maybe" to "scheduled deployment with an end date attached."

The product cadence reinforces it. NVIDIA shipped a full physical-AI stack alongside the print: **Cosmos** (world foundation model), **Isaac GR00T N / N1.5** (open humanoid foundation models), new **Isaac simulation frameworks**, **IGX Thor** general availability (industrial/medical edge compute), **Alpamayo 1.5** (open AV model), and **Omniverse NuRec** (neural reconstruction for AV simulation at scale). Mirrors the CUDA-to-Blackwell vertical integration playbook — world model → robot foundation model → sim → on-robot compute. NVIDIA is not picking a vertical; it is arming everyone who builds a robot. That is what makes physical AI a **cross-layer theme** rather than a new layer. It rides through all 8 layers of the existing AI Dashboard stack with different intensities. The point of this integration note is to map exactly where it lands and what it changes.

How Physical AI maps onto the 8-layer AI stack

The existing Dashboard reads top-down from L0 (hyperscalers, the demand) to L7 (power and thermal, the binding constraint). Physical AI does not replace any of those — it overlays a second, parallel demand curve on top of the existing agentic-AI curve, with different mix shifts at each layer. Here is what changes layer by layer.

Layer 0 — Hyperscalers (demand)

What changes: the buyer set widens. The traditional L0 buyer was the hyperscaler-plus-frontier-lab pair (Microsoft, AWS, Google, Meta, Oracle, CoreWeave, Nebius — plus OpenAI, Anthropic, xAI as renters). Physical AI adds a second buyer set that overlaps but is structurally different: **automotive OEMs** (Tesla, Hyundai, Mercedes, Volvo) build their own AV/robot training fleets; **industrial conglomerates** (Foxconn, Siemens, ABB) build factory-floor simulation farms; **robotaxi operators** (Uber via the NVIDIA partnership) run fleet inference infrastructure. The aggregate L0 wallet expands — not by a multiple of itself, but by a meaningful additional ramp of capex from a different industry base whose 2026-28 capex cycle is unrelated to the hyperscaler cycle. That is a diversification of the demand curve, not a duplication.

AI Dashboard names with upside: same hyperscaler list, plus the existing **TSLA** (already in coverage) as a new-style L0 buyer for its own training cluster (Dojo + outsourced) feeding Optimus

and FSD.

Layer 1 — Compute fabric

What changes: NVIDIA's grip widens, not narrows. The same CUDA software stack and the same Blackwell/Rubin silicon run the training side of physical AI. **GR00T N1.5 training runs**, **Cosmos world-model training**, **Omniverse NuRec scene generation** all consume top-of-stack GPU systems. On the inference side, three new SKUs become material: **Drive Thor** (in-vehicle compute for robotaxis), **Jetson Thor** (on-robot compute for humanoids and industrial), and **IGX Thor** (industrial/medical edge — now generally available). These are not data-centre Blackwells; they are smaller-die, ruggedised, lower-power variants — but they ship the CUDA ecosystem with them. **Qualcomm (QCOM)** is the credible #2 at the edge tier through Dragonwing IQ, with named design wins at Figure and NEURA Robotics. Custom-ASIC pressure (Broadcom, Marvell, Alchip ASICs designed for hyperscalers) is irrelevant to physical-AI workloads in the near term — too fragmented an end market to justify custom silicon.

AI Dashboard names with upside: **NVDA** (Tier A) — incremental ramp on top of agentic; **AMD** (Tier A) and **AVGO** (Tier A) — minimal direct physical-AI lift in the near term, mostly training spillover. **QCOM** — currently not Tier A in the AI Dashboard, deserves a relook for edge robotics. New names: NVIDIA-adjacent edge-compute pure-plays — **Lattice Semiconductor (LSCC)** — see below.

Layer 2 — Foundry and advanced packaging

What changes: less than you might think in 2026; more in 2027-28. Robot SoCs (Drive Thor, Jetson Thor, IGX Thor) are mostly TSMC N5/N4 today — not the cutting-edge N3/N2 that Blackwell and Rubin consume. So the marginal CoWoS slot impact in 2026 is modest. The longer-term read-through is different. Each physical-AI deployment cycle adds two-to-three new SoC types into the TSMC mix per OEM (one for AV, one for humanoid, one for the simulation cluster), which widens the customer base and extends the CoWoS-and-N3 demand curve a node past the agentic-AI peak that consensus models are coalescing on. **TSMC's CoWoS capacity** remains the single tightest constraint in the entire stack; physical AI is a tailwind, not a saviour.

AI Dashboard names with upside: **TSM** (Tier A) — same thesis, extended duration; **BESI.AS** and **ASMPT (0522.HK)** in hybrid bonding — modest near-term lift, real long-term lift as humanoid and AV SoCs adopt chiplet architectures. New names: none required; the existing coverage is correct.

Layer 3 — Memory

What changes: physical AI is a **training-side HBM** story, not an on-robot memory story. Cosmos world-model training, Isaac Sim renders, Omniverse NuRec scene generation all live on top of Blackwell/Rubin systems with HBM3E/HBM4. Every additional terabyte of synthetic training data is an additional HBM-bit pull. **MRAM (Everspin)** has a small but real on-robot memory niche — instant-state-save when power is lost mid-task — but it is a tens-of-millions-of-dollars story, not a Hynix-scale story. The dominant memory read-through to physical AI is the same as the dominant memory read-through to agentic AI: **HBM4 capacity through 2027**, **SK Hynix** (Tier A), **Micron (MU, Tier B / Conviction 9)** and **Samsung**. Physical AI extends the HBM cycle past 2027.

AI Dashboard names with upside: **000660 (SK Hynix)**, **MU**, **4062 (Samsung)** — same thesis, longer duration. **SNDK** — NAND adjacent, HBF (High-Bandwidth Flash) optionality holds for inference-heavy physical workloads. New names: Everspin **MRAM** for the niche on-robot persistent-memory case.

Layer 4 — Bonding, substrates, photonics

What changes: physical AI shifts more workload to the edge, which means **smaller-form-factor packaging** (Drive Thor, Jetson Thor) gets the marginal advanced-packaging slot in 2027-28. Hybrid bonding from **BESI** and **ASMPT** scales accordingly. On the photonics side, robotaxi fleets and factory-floor humanoids do not need 1.6T transceivers (they are not racks talking to racks at data-centre scale), but they do need the **wireless and short-reach optical** infrastructure that supports a distributed inference fabric. Co-packaged optics and chiplet-style optical interconnect from **Coherent (COHR)** and **Lumentum (LITE)** retain their training-cluster thesis; new SoC integrations from Broadcom and Marvell extend the curve.

AI Dashboard names with upside: **COHR (Tier B)**, **LITE (Tier B)**, **BESI.AS**, **ASMPT (0522.HK)**, **4062 / 4063 (Ibiden / Unimicron)** — same thesis, modest extension. New names: none required.

Layer 5 — Sub-component bottlenecks

This is where physical AI does its biggest mix-shift work, and where Robotics-MASTER's analysis becomes load-bearing. Physical AI introduces a **second sub-component layer** that doesn't exist for pure-agentic workloads: the **mechanical and sensing layer of the robot itself**. That layer is concentrated, oligopolistic, and not yet in the AI Dashboard's L5 coverage. The robotics master file calls out the key sockets:

- **Joint layer (motors, reducers, bearings, encoders)** — Japanese oligopoly. **Harmonic Drive (6324.T)** holds ~60-70% of strain-wave gears; **Nabtesco (6268.T)** holds ~60% of RV cycloidal gears; **MinebeaMitsumi (6479.T)** for miniature precision bearings; **THK (6481.T)** for linear motion. None currently in the AI Dashboard.
- **Sensing layer** — **Cognex (CGNX)** for industrial machine vision; **Keyence (6861.T)** for sensor breadth; **Allegro MicroSystems (ALGM)** for magnetic position sensing (already adjacent to AI Dashboard via VPG); **VPG (already Tier in coverage)** for force and tactile sensing — the single most concrete humanoid design-win pipeline among public US semis.
- **Materials** — **MP Materials (MP)** — already Tier A in AI Dashboard via rare-earths thesis; the NdFeB magnet bottleneck is one of the cleanest cross-overs.

AI Dashboard names with upside: **VPG (already in)** — explicit humanoid order book; **MP (Tier A, already in)** — every BLDC motor in a robot needs NdFeB; **Murata (6981.T)** — same MLCC bottleneck applies to robot power-delivery networks, not just data-centre boards. New names: **6324 (Harmonic Drive)**, **6268 (Nabtesco)**, **6479 (MinebeaMitsumi)**, **CGNX (Cognex)**, **ALGM (Allegro)**, **LSCC (Lattice Semiconductor)** — see new-names table below.

Layer 6 — Test and assembly

What changes: the test problem extends. **Advantest (6857.T, Tier A)** is the dominant SoC tester at high-end logic and effectively the sole supplier for NVIDIA GPU test. Robot SoCs add a second tester-loading vector — Drive Thor, Jetson Thor, IGX Thor all require automated wafer-level

test and probe-card consumables, and each new edge-compute SKU is a new probe-card design. **FormFactor**, **Teradyne (TER, in coverage)** and **CAMT** all see incremental, not transformative, lift from physical AI. The marginal pricing power stays at Advantest because the proportion of high-end test in the physical-AI BOM remains elevated.

AI Dashboard names with upside: **6857 (Advantest, Tier A)**, **TER (in coverage)**, **CAMT**, **ONTO**, **AMKR** — same thesis, slight mix tilt toward smaller-die test. New names: none required.

Layer 7 — Power and thermal

What changes: less than at first glance. Robotaxi and humanoid deployments do not move data-centre power budgets meaningfully (the L0 hyperscaler cluster is still where the megawatts go). But **edge inference power** at the rack-scale (factory floors with hundreds of GR00T-running humanoids, robotaxi fleet depots with overnight inference workloads) shows up as a new mid-density data-centre archetype: 30-50 kW racks dispersed across more sites, rather than the centralised 130-kW racks of Blackwell training clusters. Liquid cooling and modular power infrastructure from **Vertiv (VRT, Tier A)** is the cleanest read-through. **Constellation (CEG)**, **Vistra**, **GE Vernova (GEV)** all retain the same thesis — power is the binding constraint regardless of whether the workload is training or inference.

AI Dashboard names with upside: **VRT (Tier A)**, **CEG (Tier A)**, **GEV** — same thesis, extended duration into 2028-30. New names: none required.

Linked picks-and-shovels — names already in AI Dashboard with Physical AI tailwind

These are existing AI Dashboard coverage names where physical AI is a **net incremental** tailwind on top of the agentic-AI thesis they already carry. The framework is: do not re-rate purely on physical AI; do recognise that the duration of the thesis just extended.

Ticker	Current AI Dashboard tier	Why Physical AI matters for them	Linked layer
NVDA	Tier A	GR00T/Isaac/Cosmos = the physical-AI compute platform; $9B TTM physical-AI revenue already in the print	L1
TSM	Tier A	Drive Thor / Jetson Thor / IGX Thor add new SKU families to TSMC's customer mix; CoWoS demand curve extends past agentic peak	L2

Ticker	Current AI Dashboard tier	Why Physical AI matters for them	Linked layer
000660 (SK Hynix)	Tier A	Every Cosmos training run + Isaac Sim render = additional HBM-bit pull; physical AI extends the HBM4 cycle	L3
MU (Micron)	Tier B (Conv. 9)	Same HBM read-through; the only US-domiciled HBM producer is structurally advantaged for sovereign / industrial customers	L3
6857 (Advantest)	Tier A	Each new robot SoC is a new test-pattern problem; Advantest’s V93000 EXA Scale captures the marginal high-end test slot	L6
VRT (Vertiv)	Tier A	Liquid cooling extends into mid-density edge sites supporting fleet inference and factory-floor humanoids	L7
MP Materials	Tier A	NdFeB magnets — every BLDC motor in every robot needs them; ~2-4kg of NdFeB per humanoid; only Western pure-play	L5
VPG	In coverage	Force, torque, and fingertip tactile sensors for humanoids — explicit design wins, ramping to hundreds of robots/week by end-2026	L5

Ticker	Current AI Dashboard tier	Why Physical AI matters for them	Linked layer
COHR / LITE	Tier B	Photonics — modest near-term physical-AI lift; same training-cluster thesis extended	L4
CEG / GEV	Tier A / coverage	Power baseload thesis applies equally to physical-AI inference loads; gas turbines sold out into 2029	L7
AVGO / MRVL	Tier A / B	Custom-ASIC pressure mostly hyperscaler; minimal direct physical-AI lift in 2026, more in 2027-28	L1
TER (Teradyne)	In coverage	Universal Robots (cobots) + MiR (AMRs) is ~7% of revenue but the highest-growth slice; physical AI is the dedicated tailwind	L6
TSLA	In coverage	Optimus humanoid + FSD — a captive physical-AI L0 buyer + a vertical humanoid OEM	L0 / L5
9880 (UBTECH)	In coverage	The only listed pure-play actually shipping full-size humanoids into customer factories	L5
277810 (Rainbow Robotics)	In coverage	Samsung-controlled humanoid programme inside a Korean industrial wrapper	L5

The pattern: physical AI adds tailwind to existing Tier A names on the data-centre side (NVDA, TSM, SK Hynix, Advantest, Vertiv) and validates existing Tier-A side bets on the materials and sensing side (MP, VPG). It does not require re-ranking — it extends the duration.

New names — Physical AI pure-plays to add to the watchlist

The following names sit in the existing Robotics-MASTER basket but are NOT in the current AI Dashboard 98-ticker universe. Each has revenue today, a public-market listing investable through IBKR, and a clear bottleneck position in the physical-AI stack. No vaporware. Suggested tiers below are calibrated against the existing AI Dashboard ranking convention (Tier A = 3-5% sizing, Tier B = real thesis with valuation/timing caveats, Tier C = optionality).

Ticker	Name	Layer	Function	Suggested tier	Why now
6324	Harmonic Drive Systems (TSE)	L5	~60-70% global share of strain-wave gear reducers; the single largest single-supplier bottleneck in the humanoid joint BOM. A humanoid uses 24-30 of them	Tier A	Jensen's "industrial + surgical + humanoid" framing pulls every public humanoid programme through this name's order book. Currently extended on tape — leg in on weakness, but the thesis is the cleanest mechanical bottleneck in the basket.
6268	Nabtesco (TSE)	L5	~60% global share of RV cycloidal gear reducers; high-load humanoid joint sockets (hips, knees, shoulders)	Tier A	Same Japanese oligopoly story. Confirmed humanoid customer expansion. Same "extended" timing flag as 6324 — entry discipline matters.

Ticker	Name	Layer	Function	Suggested tier	Why now
6479	MinebeaMitsumi (TSE)	L5	Miniature ball bearings + integrated micro-actuators; co-developed humanoid micro-actuator with Harmonic Drive	Tier B	Less concentrated than 6324/6268 but a real humanoid order book; bearings are 100+ per humanoid. Cleaner valuation than 6324.
CGNX	Cognex	L5 / L2 of robot	Global #1 industrial machine vision; "the robot's eyes" in factory automation	Tier B	Industrial-automation buy cycle aligns with Foxconn / Siemens NVIDIA partnerships. Real fundamentals, modest tape extension.
ALGM	Allegro MicroSystems	L5 / L2 of robot	#1 magnetic position-sensing chip company; named in at least two public humanoid BOMs; robotics revenue line doubled YoY	Tier B	Currently the calmest chart in the sensing layer (RSI 55, +19% vs 50-DMA). Robotics is a fast-doubling option on top of a recovering auto base. Cleanest risk/reward in the sensing layer.

Ticker	Name	Layer	Function	Suggested tier	Why now
LSCC	Lattice Semiconductor	L1 of robot (mid-tier compute)	Low-power FPGAs as the sensor-fusion glue layer in humanoids; explicit NVIDIA Halos and TI robotics partnerships	Tier A	Strongest pure-robotics thesis in compute. Calm tape (RSI 58). AI mix already in the print; humanoid is the next leg.
005380	Hyundai Motor (KRX)	L5 OEM / wrapper	Owns ~80% of Boston Dynamics (Atlas, Spot); Mobis actuators; 30k-unit robot factory — wrapped inside a 14x-earnings automaker	Tier A	Cheapest listed exposure to a frontier humanoid programme. The robotics optionality is close to unpriced. Jensen named Boston Dynamics directly in adjacent ecosystem coverage.

Ticker	Name	Layer	Function	Suggested tier	Why now
SYM	Symbotic	L5 OEM	GAAP-profitable warehouse-robotics integrator with a $22B+ backlog; the only humanoid-adjacent OEM that makes money	Tier B	Oversold (RSI 35) on Walmart-concentration fears. Concentration is the real risk; the tape is the opportunity. Direct read-through to physical AI as the “warehouse” vertical Jensen called out.
6954	Fanuc (TSE)	L5 OEM (industrial)	World’s largest industrial robot-arm maker; named by Google/Fanuc physical-AI partnership coverage	Tier A	Pure-play industrial robotics incumbent with profitable cash flow; physical-AI overlay extends the customer cycle.
6506	Yaskawa Electric (TSE)	L5 OEM + L1 of robot	Industrial robots + servo/motion controllers; sits in both the OEM and the component-supplier role	Tier A	Cleanest dual-exposure name in the basket. Both buyer and supplier of physical-AI parts.

Top 5 to add to watchlist first (priority order): 6324 (Harmonic Drive), **MP Materials (already Tier A)**, **LSCC (Lattice)**, **005380 (Hyundai Motor)**, **ALGM (Allegro)**. These are the cleanest entries — three pure bottleneck names, one cheap optionality wrapper, one calm-tape sensing leader. Note: MP is already covered; the four genuinely new tickers to onboard into the

dashboard data pipeline are 6324, LSCC, 005380, ALGM. Add **6268 (Nabtesco)**, **CGNX**, **6954 (Fanuc)**, **6506 (Yaskawa)**, **SYM** and **6479 (Minebea)** in a second wave.

Names from the Robotics-MASTER basket that should NOT be elevated to the AI Dashboard universe at this stage: **AMBQ (Ambiq)** — extended +116% vs 50-DMA, Bucket D, deep-pullback-only; **MRAM (Everspin)** — extended +170% vs 50-DMA, also Bucket D; **6594 (Nidec)** — uninvestable on the accounting scandal disclosed February 2026; **454910 (Doosan)**, **ARBE**, **KITT**, **RR** — speculative or pre-revenue.

Validation table — our prior Robotics-MASTER report vs Jensen's May 20 call

The Robotics-MASTER report was published May 14, 2026 — six days before the NVIDIA Q1 FY27 call. The cross-reference below tests whether the report's thesis survived the print.

Robotics-MASTER claim	Jensen / Q1 FY27 confirmed?	Discrepancy / new info
NVIDIA is the de facto compute platform for humanoids via Jetson Thor + Isaac + GR00T + Omniverse	**Confirmed**	Q1 FY27 added Cosmos (world foundation model), GR00T N1.5 (open humanoid model line), IGX Thor general availability, Alpamayo 1.5 (open AV model), Omniverse NuRec (neural reconstruction for AV sim). Full vertical stack now in market.
Physical AI is a 2027-28 P&L story, not 2026	**Partial — accelerated**	Q1 FY27 print shows physical AI is ALREADY a $9B TTM revenue line. Faster than the report implied. Edge Computing segment now broken out so analysts can track quarterly.
Humanoid OEM layer has 10+ credible programmes — Tesla, Figure, Apptronik, Unitree, Agibot, UBTECH, 1X, Boston Dynamics, Sanctuary, Mentee	**Partial**	Jensen named Boston Dynamics + Agility + Foxconn + Hyundai in adjacent NVIDIA ecosystem material around the print; Figure, 1X, Sanctuary, Apptronik did NOT appear in the call recap. NVIDIA is keeping humanoid partner exclusivity open.
Robotaxi is a future deployment story	**Confirmed with dates attached**	Uber partnership: ~30 cities, 4 continents, by 2028 on NVIDIA Drive. Most concrete deployment timeline given on the call.

Robotics-MASTER claim	Jensen / Q1 FY27 confirmed?	Discrepancy / new info
Joint layer is a Japanese oligopoly (Harmonic, Nabtesco, Minebea, THK)	**Not directly addressed by Jensen**	The call did not name the component layer specifically, but the “industrial + surgical + humanoid” framing pulls every public humanoid programme through the same supplier base. Thesis intact.
NdFeB rare-earth magnets are the materials bottleneck; MP Materials is the only Western pure-play	**Not directly addressed by Jensen**	Not named on the call. Independent thesis — supported by ongoing US-China rare-earth tension, unchanged by the print.
Sensing layer — Cognex (vision), Allegro (position), VPG (force/tactile)	**Indirectly confirmed**	“Industrial + surgical + humanoid” coverage pulls every sensor type. VPG specifically has a public-disclosed humanoid order book that aligns with NVIDIA’s industrial robot ecosystem.
Compute layer — NVDA + QCOM + Lattice + Ambarella	**Confirmed for NVDA + QCOM**	NVIDIA + Qualcomm combined share at the high-level reasoning tier ~80%+. Lattice + Ambarella are mid-tier perception, not named by Jensen but consistent with the architecture.
Boston Dynamics is the marquee humanoid platform	**Confirmed**	Cited in adjacent NVIDIA ecosystem coverage as a flagship physical-AI partner. Hyundai (the 80% owner) is named directly.
Foxconn is an industrial automation partner	**Confirmed**	Cited as both customer (using Blackwell production lines) and as physical-AI fab-automation reference.
China A-share humanoid names (Unitree, Agibot) are out-of-universe	**No change**	Jensen did not name Chinese partners. Geopolitical posture unchanged.

Robotics-MASTER claim	Jensen / Q1 FY27 confirmed?	Discrepancy / new info
Tesla is a captive physical-AI vertical integrator	**Confirmed indirectly**	Tesla not named on the NVIDIA call, but the Optimus + FSD + Dojo programme is consistent with the framing.
Surgical robotics is a credible adjacency	**Confirmed**	Jensen explicitly listed “surgical” alongside “industrial” and “humanoid” in the breadth statement.
Picks-and-shovels beats OEMs in humanoid wave	**Implicitly confirmed**	NVIDIA’s strategy is explicitly horizontal — “arm everyone who builds a robot.” Confirms the picks-and-shovels framework in Robotics-MASTER.

Gaps — Jensen named, Robotics-MASTER did not cover: - **Skild AI** — robotics foundation model partner named in adjacent ecosystem coverage; private, not investable directly. - **Disney** — physical-AI/Omniverse partner; entertainment-robotics adjacency; not in Robotics-MASTER as the report’s geography is “investable robotics names,” not “physical-AI ecosystem partners.” - **Siemens** — industrial physical-AI / Omniverse partner; not in Robotics-MASTER (industrial conglomerate, not robotics pure-play). - **TSMC fab-automation use case** — TSM is in AI Dashboard, but not framed as a physical-AI customer in Robotics-MASTER.

Gaps — Robotics-MASTER covered, Jensen did not directly validate: - The **joint layer Japanese oligopoly** (6324, 6268, 6479, 6481) — Jensen did not name them, but the physical-AI framing pulls them through. Thesis intact. - **MRAM (Everspin)** — Jensen did not name on-robot persistent memory as a category. Thesis is a niche bet, not invalidated. - **OUST, AEVA, ARBE** — lidar / radar perception names. Jensen did not name them; Mobileye-style perception was not a focus of the call. Thesis is structurally intact but lacks the direct namecheck.

Bottom line: the Robotics-MASTER thesis survived the print intact and was, if anything, accelerated. The report’s “physical AI is a 2027-28 P&L story” framing was conservative — physical AI is already a $9B TTM revenue line, two years ahead of the implied curve. The report’s picks-and-shovels framework is exactly the same playbook NVIDIA management described in its prepared remarks. The two layers Jensen extended beyond the Robotics-MASTER scope — Skild AI, Siemens, Disney — are mostly private or industrial-conglomerate adjacencies that do not change the listed-equity coverage.

Updated thesis statement

Physical AI does not displace the existing AI Dashboard framework — it widens it. The 8-layer stack remains the right way to read the AI buildout, but every layer now carries a second demand curve on top of the agentic-AI curve, with most of the mix-shift concentrated in L1 (incremental NVIDIA edge SKUs + Qualcomm + Lattice), L3 (extended HBM cycle from training-side world-

model and sim workloads), L5 (a new mechanical-and-sensing sub-component layer — the robot's body — that maps to the Japanese joint-layer oligopoly, Cognex/Allegro/VPG sensing, and MP Materials NdFeB), and L0 (a new buyer set of automotive OEMs, industrial conglomerates, and robotaxi operators sitting alongside the hyperscalers). The cleanest expression of the theme inside the existing 98-ticker universe is to keep current Tier-A weights on NVDA, TSM, SK Hynix, Advantest, Vertiv, MP and VPG — these names carry the physical-AI extension for free — and to onboard four to six new names from the Robotics-MASTER basket as the dedicated robotics tier (Harmonic Drive, Lattice, Hyundai Motor, Allegro as priority; Nabtesco, Cognex, Fanuc, Yaskawa, Symbotic and Minebea as second wave). Treat Edge Computing segment growth as the quarterly leading indicator — when sequentials accelerate, the wave has crested into the P&L. Until then, the theme is real, the bottlenecks are real, and the duration of the existing AI Dashboard thesis just got a year longer.

Part 3 — Tier A · Highest Conviction

Sorted by AI stack layer ascending; within each layer by conviction rank. 18 names. Target portfolio weight: 60-75% of AI-allocated NLV concentrated here.

Layer 1 · NVDA — NVIDIA Corporation

One-line thesis

NVIDIA prints the silicon, the system, and the software that every meaningful AI training cluster on Earth is built on top of — and the May 20 print combined with the Saudi HUMAIN / UAE 500K-Blackwell sovereign-AI tailwind is the highest-asymmetry single event of the FY27 cycle.

What NVIDIA physically does

At the atomic layer NVIDIA designs the GPU dies (Hopper, Blackwell, Blackwell Ultra, and the next-generation Rubin platform) that perform the matrix-multiply-and-accumulate operations underneath every transformer-style large language model. A single Blackwell B200 is two reticle-limited 4nm dies bonded together on a CoWoS-L interposer with 192 GB of HBM3e in eight stacks immediately adjacent. A GB200 NVL72 rack puts 72 such GPUs and 36 Grace ARM CPUs into a single liquid-cooled 1.4-ton chassis wired together with the fifth-generation NVLink switch fabric — 130 TB/s of all-to-all bandwidth, more bandwidth inside one rack than the entire global Internet backbone moved a decade ago.

But the silicon is only the entry ticket. NVIDIA's actual monopoly sits one layer up — CUDA, cuDNN, NCCL, TensorRT, the Megatron / NeMo / Dynamo software stacks, the Spectrum-X and Quantum InfiniBand fabrics, and the Mission Control / Run:ai orchestration layer that turns a 100,000-GPU campus into something a researcher can actually train on without the cluster falling over every 17 minutes. Replacing the chip is possible (AMD MI355X exists, Google TPU v6 exists, AWS Trainium 3 exists). Replacing the entire vertically integrated stack from die to data-center operating system is what every credible competitor has been trying and failing to do for ten years.

For an AI-stack investor NVIDIA is Layer 1 silicon and Layer 6 frameworks simultaneously — the most vertically integrated position in the cycle, and the principal reason gross margin has

held above 70% for seven straight quarters in a business that consensus told you in 2022 would mean-revert toward 50%.

The financial print

Q4 FY26 (reported Feb 26 2026) printed $39.3B of revenue against $38.1B consensus and $0.89 of non-GAAP EPS against $0.84 — beat across the board, with the Blackwell Ultra ramp accelerating into the quarter. Gross margin held at 73.5% non-GAAP despite the well-flagged Blackwell yield drag, and data-center revenue at $35.6B (+91% YoY) confirmed that the platform transition from Hopper has not produced even a single quarter of sequential air-pocket — the unique characteristic of this cycle versus every prior NVIDIA product transition.

Q1 FY27 prints May 20 (after the close). Consensus sits at approximately $79.2B of revenue and $1.78 of EPS. The setup is asymmetric in a way that rarely happens at $5T market cap: the Saudi HUMAIN partnership announced May 13 — 18,000 GB300s in phase one, scaling to 500 MW of NVIDIA-supplied capacity — and the May 15 US-government greenlight on the UAE deal worth up to 500,000 Blackwells per year, both landed inside the quoted quarter. Neither is in consensus rev, neither is in consensus FY27 EPS, both are confirmed deals with named customers and signed agreements. The stock sits at $225 versus the all-time high of $236.54 set May 14 — 1m +8%, 3m +19% — meaning the print can re-rate the name on guide alone without needing the headline number to surprise.

Customer mix

Hyperscalers (Microsoft, Meta, Amazon, Google, Oracle Cloud) plus Anthropic, OpenAI, xAI account for roughly 55-60% of FY26 data-center revenue. The new leg is sovereign AI — Saudi PIF (HUMAIN), UAE G42, Singapore EDB, Japanese METI-backed clusters, and the rumoured Indian sovereign program — moving from <5% of revenue in FY25 to a credible 15-20% by FY27 if HUMAIN + UAE deliver on stated capacity. Enterprise (DGX Cloud, NIM microservices) is the third leg, still single-digit percent of revenue but growing 100%+ YoY. The concentration risk is real but materially less than it was 18 months ago when four hyperscalers were 65%+ of the print.

The disaggregation that matters for cycle analysis: the top-four hyperscaler customers (Microsoft, Meta, Amazon, Google) together represent approximately 45-50% of data-center revenue, down from 65%+ in early 2024. The frontier-model labs (OpenAI, Anthropic, xAI) represent roughly 10-15% of revenue and are growing the fastest in percentage terms. The sovereign-AI customer set — currently 5-8% of revenue but with HUMAIN and UAE contracts that lift this to 15-20% by FY27 — operates on multi-year contracted basis that materially reduces the cycle-sensitivity of the order book. The "tail" — neoclouds (CoreWeave, Lambda, Nebius), enterprise (Oracle Cloud, Tesla, internal corporate deployments), and the broader long-tail of regional and university clusters — collectively represents 15-20% of revenue and is structurally less concentrated than the hyperscaler cohort.

The Saudi HUMAIN announcement on May 13 — 18,000 GB300s in phase one, scaling to 500 MW of NVIDIA-supplied capacity — is the largest single contracted commitment in the sovereign-AI category to date, and the UAE 500K Blackwells per year approval on May 15 is comparable. Both deals were announced inside the May 20 print quarter, neither is in consensus FY27 revenue, and both are confirmed agreements with PIF-level and government-level counterparty quality.

Competitive context

AMD's MI355X has CUDA-compatible ROCm 7 software, competitive HBM3e bandwidth, and serious commitments from Microsoft and Meta — but it is a single-chip story without the rack-scale fabric. Google's TPU v6 (Trillium) and the rumoured v7 are internal-only, never offered as merchant silicon, so they constrain Google capex toward NVIDIA but don't compete in the broader market. AWS Trainium 3 (Marvell-co-designed) is at 50% of Anthropic's training load per AWS commentary — meaningful, but Anthropic continues to take NVIDIA capacity in parallel. Intel Gaudi 3 has effectively conceded. The MI355X is the only credible cross-vendor competitor in 2026, and it takes share at the edges of the cluster — primarily inference and smaller training jobs — not the frontier model build.

NVIDIA's competitive moat is the vertical stack. The 1m-GPU OpenAI Stargate buildout, the xAI Memphis Colossus 2 expansion to 1m GPUs, the Microsoft Fairwater campus in Wisconsin — these are not chip purchases, they are platform deployments. The cost of switching the platform at this scale, measured in research-team-years of CUDA rewrites and validated software stacks, is the actual moat.

Terminal risk

The terminal risk is the Jevons-inverted scenario: DeepSeek-style algorithmic efficiency gains compound faster than capability gains, frontier model training quietly becomes a 10x-smaller compute problem, and hyperscaler capex collapses from $400B in 2026 to $200B in 2028 — not because demand died, but because the per-token cost of intelligence dropped faster than demand expanded. This is the "bubble warning" that gets recycled every six months in sell-side decks. It is not the base case but it is also not zero.

Bull case

The base case is FY27 revenue of $310-340B (consensus $295B), FY27 EPS of $7.00-7.80 (consensus $6.50), gross margin holding 72-74%, and the Rubin platform launching into shipping volumes in early 2027 with the same kind of two-quarter air-gap-free transition Blackwell delivered. At a 35x forward earnings multiple — well below current and below the 5-year average — that prints a stock at $250-275, +10-22% from spot.

The upside case is sovereign AI moving from 15% of revenue to 25-30% by 2028 as Saudi, UAE, Indian, and Japanese national programs ramp; Rubin pricing 30%+ above Blackwell on a per-rack basis (consistent with TSMC capacity costs and HBM4 pricing); and the inference market — currently 40% of data-center revenue — compounding 80%+ for two more years as the agentic-AI transition pulls inference compute from one-shot LLM serving to multi-step tool-using agents. That case prints a stock at $330-380, +47-69% from spot, on FY28 EPS of $10-12 at 30-32x.

The agentic-AI transition is the most consequential and least-modelled tailwind. Frontier-model inference today serves principally one-shot or short-turn LLM chat workloads. The 2026-27 transition to agentic AI (multi-step tool use, plan-and-execute frameworks, persistent agents) compounds inference compute by 10-100x per user interaction — every user query becomes a chain of model calls rather than a single forward pass. The consensus FY27 model attributes essentially zero of this scaling to NVIDIA's inference compute line. The bull case is that this transition lifts hyperscaler inference compute spend by 50-80% YoY for two consecutive years through 2028, and NVIDIA captures the majority share given the CUDA-native inference stack.

Gap / bear case

Three things the market may be missing the wrong way. First, Blackwell Ultra is shipping into a customer base that absorbed Hopper, GB200, and Blackwell Ultra inside 18 months — the question is not demand, but whether even hyperscaler balance sheets can keep the +35-40% YoY data-center capex pace through 2027 without something breaking on the financing side. The aggregate hyperscaler capex pace in 2026 is approximately $400B (Microsoft, Meta, Amazon, Google, Oracle combined), versus aggregate operating cash flow of approximately $550B. The capex-to-OCF ratio is at a 25-year high; one bad quarter of cloud-services growth could force capex moderation.

Second, the China revenue line that briefly returned with H20 has gone again under the latest export tightening, and the consensus FY27 model assumes ~$5B of China contribution that may not materialize. Third, gross margin guidance has crept down quarter-on-quarter — 75.0% to 73.5% to 73.0% guide for Q1 — and at some point the Rubin yield ramp will pressure that line harder than consensus expects.

The DeepSeek / algorithmic-efficiency thread is the recurring tail concern. DeepSeek demonstrated in early 2025 that frontier model training could plausibly be done at 10-15% of the GPU-hours that the consensus had assumed, and while the equity market shrugged off the implications, the underlying technical reality is that algorithmic compounding compounds. If 2026-27 sees a step-change in inference-time scaling efficiency (e.g., a new model architecture that requires materially fewer parameters or runs at materially lower compute cost per token), the aggregate GPU demand curve bends downward even as AI adoption accelerates — the Jevons-inverted scenario. This is not the base case but it is the largest tail risk in the name.

Optionality

Four options not in consensus. First, Rubin platform launch ahead of schedule (Q1 27 versus consensus Q2 27) on the back of TSMC N2 ramp running clean and the CoWoS-L capacity ramp arriving ahead of plan. Second, NVIDIA Drive autonomous-vehicle revenue inflecting in 2027 as Mercedes, Toyota, and BYD ramp NVIDIA-Orin-based L3+ systems — a $5-10B line currently lost in the corporate “other” bucket. Third, the long-tail sovereign-AI option: if Saudi HUMAIN moves from 500 MW to 1.5 GW of contracted capacity by 2028 (per PIF stated ambitions), the marginal revenue from that single account exceeds NVIDIA’s entire FY24 data-center business. Fourth, the inference-software monetisation layer: NIM microservices, the Dynamo inference orchestration platform, and the Run:ai enterprise scheduler are positioned to monetise the inference cycle separately from hardware revenue — an “AWS for AI” software monetisation step that the consensus model attributes essentially zero value to today.

The trade

NVDA — BUY 9/10. Entry*: starter at $220-225 spot, add on $200-210 if Q1 FY27 print disappoints and the stock pulls back into the 50-day.* ***Position size****: 4-5% of NLV as the core AI cycle exposure — larger than any other single name in the stack.* ***Stop****: daily close below $195 (technical break of the 200-day moving average and the December 2025 base).* ***Catalyst date****: May 20 print after-hours; Rubin tape-out commentary; Aug 27 Q2 FY27 print; CES 2027 keynote.* ***Trim/exit triggers****: data-center YoY growth slowing below 50% on a single quarter; gross margin guide below 71%; any verifiable cancellation of Stargate or HUMAIN.* ***Conviction****: this is the*

highest-quality compounder in the cycle and the cleanest expression of every Layer-1-through-Layer-6 thesis simultaneously. The reason the conviction is 9/10 not 10/10 is the bubble-warning tail risk — at 35-40x forward you cannot ignore that the consensus is on the same side of the trade as you are.

Layer 1 · AVGO — Broadcom Inc.

One-line thesis

Broadcom is the only credible non-NVIDIA design partner for hyperscaler custom AI silicon at scale — Google TPU, Meta MTIA, ByteDance, the rumoured OpenAI XPU — and a confirmed $73B AI backlog into a $100B 2027 AI revenue target makes the June 3 print a structurally bullish setup.

What Broadcom physically does

Broadcom sits in two physically distinct semiconductor businesses that together form one of the cleanest AI bottlenecks in the stack. The first is AI custom silicon — the co-design of hyperscaler-specific accelerators (Google's TPU v5e, v5p, v6 Trillium; Meta's MTIA v1 and v2; ByteDance's internal accelerator; and the strongly rumoured OpenAI Titan / XPU due 2026-27) where Broadcom owns the SerDes, the packaging integration, and the back-end physical implementation while the customer owns the architecture and the compiler. Roughly 50% of the world's hyperscaler custom-accelerator volume rolls through Broadcom's Avago / LSI design teams in San Jose, Singapore, and Bangalore. The second business is AI networking — the Tomahawk 5 and Tomahawk 6 Ethernet switch silicon (51.2T and 102.4T per ASIC respectively), the Jericho 3 routing silicon, the Bailly 51.2T co-packaged optics module, and the 800G / 1.6T SerDes / PAM4 DSPs that move data inside and between AI clusters. This is the "second silicon" that NVIDIA Spectrum-X competes with — and where Cisco / Arista / Marvell / Broadcom compete with each other for the merchant share that hyperscalers refuse to single-source to NVIDIA InfiniBand.

In plain English: every hyperscaler AI rack that is not pure NVIDIA-NVLink contains Broadcom silicon, usually multiple Broadcom dies, in the switching layer. Every credible non-NVIDIA training cluster (Google's TPU pods, Meta's MTIA-based RSC, the AWS Trainium clusters at Anthropic) contains Broadcom-co-designed accelerators or Broadcom networking or both. The company is the only Tier-1 fab-less semiconductor vendor with a meaningful design-services franchise for AI silicon — Marvell is the only direct comp, and Broadcom's share of the custom-XPU market is materially larger.

The non-AI half of Broadcom — VMware, mainframe software, broadband, wireless — generates roughly half of corporate cash flow but is structurally low-growth. The AI half is what re-rates the stock and is the entirety of the equity thesis.

The financial print

Q1 FY26 (reported March 4 2026, quarter ended Feb) printed $14.92B of revenue against $14.62B consensus, with AI semiconductor revenue at $8.4B (+106% YoY) and total semiconductor revenue at $11.7B (+62% YoY). Non-GAAP EPS of $2.10 beat consensus by 9 cents. The headline that moved the stock was the disclosed $73B AI backlog and CEO Hock Tan's guidance that the

AI revenue line should hit a $100B annualized run-rate by end-2027. Software revenue at $5.4B grew 47% YoY (VMware re-pricing still flowing through).

Q2 FY26 prints June 3 after-hours. Consensus sits at approximately $15.05B of revenue and $2.40 of non-GAAP EPS, with whisper expectations as high as $2.55 on the back of continued AI semiconductor strength. The Goldman PT raise to $500 from $475 on May 12 and the broader Wall Street move toward $500-550 PTs ahead of the print reflect a tape that is set up for a beat — the risk is that the bar is now high enough that an in-line print causes consolidation.

The stock at $423 sits roughly 8% below the all-time high of $462 set in mid-April. 1m +12%, 3m +18%. Market cap above $2T puts AVGO firmly in the top-5 US semiconductor capitalisation tier alongside NVDA, TSM and the AMD/QCOM cohort below.

Customer mix

The AI custom-silicon revenue is heavily concentrated in three hyperscaler customers — Google (TPU), Meta (MTIA), and ByteDance — which together represent approximately 80% of the AI custom-silicon line. The rumoured fourth and fifth customers (OpenAI / Titan, and either Apple or Microsoft for inference silicon) move from rumour to confirmed during 2026-27 on the current sell-side consensus path. On the networking side, the customer mix is broader — every hyperscaler buys Tomahawk and Jericho silicon for some portion of their fleet, with Cisco / Arista / Juniper as the system-integration vendors in between. Concentration risk in AI silicon is genuine; concentration risk in networking is much lower because the customer set spans hyperscalers, enterprise, and the AI-native cloud build-outs (CoreWeave, Lambda, Oracle Cloud).

Google is likely Broadcom's largest single customer at roughly 25-30% of AI revenue, given the TPU v6 (Trillium) volume that ramped through 2025 and the v7 production now ramping through 2026. Meta MTIA v2 is the second-largest contributor at roughly 15-20% of AI revenue. ByteDance — the dark-horse customer that Western sell-side typically under-attributes — likely represents 10-15% of the AI line, tied to TikTok and Doubao internal AI build-out. The remaining ~30-40% of the AI custom-silicon line is distributed across smaller wins and the design-services engagements that have not yet shipped at production volume.

The networking franchise has a different structure: every hyperscaler buys Tomahawk silicon at multiple data-center generations, but the volume is mediated through systems vendors (Cisco using the Silicon One competing platform, Arista using Tomahawk as the primary platform, Juniper using mixed solutions). The software business (VMware, mainframe, broadband, wireless) goes to enterprise and operator customer sets that are entirely separate from the AI silicon book — this provides the cash-flow ballast that allows Broadcom to fund the AI custom-silicon R&D investment without the equity equivalent of the pure-AI-cycle volatility of Marvell.

Competitive context

Marvell is the primary competitor in custom AI silicon — AWS Trainium 3 is Marvell-designed, the Microsoft Maia speculation is Marvell-flavoured, and the broader custom-ASIC design-services market is essentially a Broadcom-Marvell duopoly. Broadcom's structural advantage is scale (Avago's back-end IP library is materially larger than Marvell's), packaging integration (Broadcom owns more of the 2.5D / CoWoS-S interface IP), and the bundled networking offering. In switching silicon, Cisco's Silicon One and Marvell's Teralynx are the direct competitors at the

merchant 51.2T / 102.4T tier; Broadcom Tomahawk maintains roughly 60% merchant share and is the default reference platform every hyperscaler validates against.

The VMware acquisition is now 24 months into integration and the once-controversial price-up strategy has delivered the cash-flow accretion Hock Tan promised, with very little customer churn beyond the long-tail enterprise base. That cash flow is what funds the AI design-services investment.

Terminal risk

The terminal risk is the hyperscaler-bypass scenario: if a Google or a Meta moves to fully in-house design teams and cuts Broadcom from the TPU / MTIA roadmap on a future generation, the AI custom-silicon line gets a 25-30% air-pocket on a single product transition. There is precedent — Apple did exactly this with its modem effort against Qualcomm — and the only reason it has not happened with Broadcom is the design complexity at 3nm and 2nm nodes, where hyperscaler internal teams genuinely cannot execute alone. If the AI-stack node moves to a model where the customer fully owns physical design (Google has hinted at this for v7), Broadcom's value-add compresses materially.

Bull case

The disclosed path is $73B of contracted AI backlog converting through 2026-27 plus net new design wins lifting the FY27 AI revenue run-rate to $100B, on company guidance. If that lands, total FY27 revenue prints near $80-85B (FY25 was $51.6B), non-GAAP EPS at $10.50-12.00, and at a 35x forward — below the 5-year average and consistent with the AI-quality re-rate that has happened across the cohort — the stock is a $370-420 print, +5-25% off spot but with the option value of an OpenAI Titan announcement.

The upside case is OpenAI Titan confirmed as a Broadcom design partnership in 2026 (rumoured but unconfirmed), Apple inference silicon design-win added in 2027, and the AI revenue line passing $120B run-rate by FY28. That scenario prints the stock toward $550-620 in 2027, +30-46% from spot. Goldman's $500 PT and the broader $500-550 sell-side cohort are tracking this case at a discount.

The structural reason the upside case is credible is that hyperscaler custom-silicon spend is converging on a multi-vendor procurement structure where every major operator validates designs at both Broadcom and Marvell. As the number of custom-XPU customers expands from three (Google, Meta, ByteDance) to a credible five-to-seven (adding OpenAI, Apple, Microsoft for inference, possibly Tesla for Dojo successor), the design-services TAM materially expands. Broadcom is positioned to capture 50-60% of this expanded TAM given the SerDes IP advantage and the depth of packaging-integration capability. The conservative-base-case versus upside-case differential is principally the question of whether one or two of these rumoured wins convert during 2026.

Gap / bear case

Two things the market may be missing the wrong way. First, the $73B AI backlog figure is a multi-year revenue book, not a one-year guide — converting that to a $100B FY27 run-rate requires net-new wins beyond what is already booked, and the OpenAI Titan / Apple inference contributions are in the consensus model but not yet contractually papered. Second, the VMware software

re-pricing tailwind is now in the base, and the FY27 software comp gets harder as VMware ARR has now substantially re-rated to subscription pricing.

The third concern is concentration: if any one of Google, Meta, or ByteDance pulls a design generation in-house, the AI custom-silicon line gets choppy in a way the bull case currently dismisses.

Optionality

OpenAI Titan / custom XPU disclosure during 2026 — this is the largest single discrete option in the name, and the rumours have moved from "speculative" to "concrete enough that Hock Tan is responding 'no comment' " on the earnings calls. The reported timing on Titan first silicon is late 2026, with mass-production volume in 2027 — a $4-6B annual revenue line at peak if the program scales to multi-100k-chip-per-year deployment. Apple inference silicon partnership (rumoured for Apple's data-center inference build-out around Apple Intelligence) is the second-largest option — Apple historically does not use external design partners, but the publicly disclosed pivot to data-center inference compute for the Apple Intelligence rollout requires custom silicon that Apple's internal team has not been resourced to deliver in time for the 2026-27 deployment.

The slower-burn option is hyperscaler co-packaged optics adoption ramping in 2027-28 (Bailly 51.2T CPO) — a Broadcom-led category that could add a $3-5B incremental annual revenue line within five years. CPO eliminates the pluggable optical transceiver and integrates the optical engine directly into the switch ASIC, dramatically reducing power per bit at the expense of higher manufacturing complexity. Broadcom is the share leader at the CPO development tier; first commercial deployments are expected at Microsoft / Google in 2027-28.

The fourth option is the broader networking displacement of NVIDIA InfiniBand. As hyperscalers move toward open-Ethernet AI clusters (the UALink / Ultra Ethernet Consortium effort, jointly backed by Broadcom, AMD, Cisco, Microsoft, Meta and others), the Tomahawk 6 silicon at 102.4T is positioned to capture share that NVIDIA Spectrum-X and Quantum InfiniBand currently hold. Even modest share displacement in the AI networking spend ($30B+ annually by 2027 on consensus) materially accretes to Broadcom.

The trade

AVGO — BUY 8.5/10. Entry*: starter at $420-425 spot; size up on $390-400 pullback into the 50-day moving average if the broader semiconductor cohort consolidates ahead of the June 3 print.* ***Position size****: 3-4% of NLV as the second-largest AI Layer-1 position behind NVDA.* ***Stop****: daily close below $370 (50-day MA breach plus reversal of the April highs).* ***Catalyst date****: June 3 Q2 FY26 print after-hours; any OpenAI Titan / Apple inference disclosure during 2026-27; Q3 FY26 print early September.* ***Trim/exit triggers****: AI semiconductor YoY growth decelerating below 70% on a single quarter; verifiable loss of Google TPU v7 design partnership; software margin contraction signalling VMware integration fatigue.* ***Conviction****: the cleanest non-NVIDIA Layer-1 AI exposure in the public market, with a software cash-flow ballast that makes the equity less GPU-cycle-sensitive than peers. The June 3 print is the binary.*

Layer 1 · MRVL — Marvell Technology

One-line thesis

Marvell is the second-largest custom AI ASIC design house, the principal silicon partner on AWS Trainium 3, and a Microsoft Maia design win speculatively priced in — and the May 27 print into a stock that is +135% in 10 weeks is the asymmetric binary of the entire AI semiconductor cohort.

What Marvell physically does

Marvell occupies the design-services niche immediately adjacent to Broadcom — co-designing custom AI accelerators for hyperscaler and large-enterprise customers, plus a substantial business in electro-optics (the silicon photonics and PAM4 DSP chips that drive 800G and 1.6T pluggable transceivers across every AI data-center optical link). The flagship AI program is AWS Trainium — Marvell is the principal design partner on Trainium 2 (in production) and Trainium 3 (ramping H2 2026 for the Anthropic Rainier supercluster build-out). The Microsoft Maia 100 / 200 / 300 series is rumoured to use Marvell on at least one generation of design service. The Google Axion ARM CPU also has Marvell-touched IP though that's a smaller line.

The custom-silicon design-services business at Marvell is structured around five technical pillars: high-speed SerDes (the IO links that move data on and off the chip — typically 112G or 224G per lane at AI generation), advanced packaging integration (CoWoS-S and CoWoS-L interface design at TSMC), high-bandwidth memory interface IP (the controllers that connect the accelerator die to its adjacent HBM stacks), networking-on-chip fabric design (the internal data-movement structure that connects compute cores to memory), and the physical implementation expertise (place-and-route at 3nm and 2nm nodes, signoff for advanced packaging). Of these, the SerDes IP and the advanced-packaging integration are the principal moats — both are multi-decade IP libraries that hyperscaler internal teams cannot easily replicate.

In addition to custom AI silicon, Marvell makes the optical DSPs that sit inside Lumentum and Coherent transceivers — the 100G / 200G PAM4 chips that drive 800G transceivers, and the next-generation 200G-per-lane silicon that enables 1.6T pluggables for AI cluster spine connectivity. The electro-optics franchise was built principally through the 2021 acquisition of Inphi (the merchant PAM4 DSP leader) and the 2022 acquisition of Innovium (cloud-grade switching silicon). Together these acquisitions gave Marvell a $2-3B annual revenue line in optical and networking silicon that operates structurally above the historical Marvell margin profile and is pulled by the same AI data-center capex cycle that drives the custom-silicon design-services revenue. This positions Marvell as a Layer-1 silicon company that simultaneously captures Layer-4 (optical connectivity) economics.

The data-center storage controller business — once Marvell's bread and butter — has compressed to a single-digit revenue contributor as the AI custom-silicon line has scaled. The enterprise networking line (the Prestera switching silicon for data-center top-of-rack switches) is the third-largest segment but is also shrinking as a percentage of revenue. The shift in the mix is the principal driver of the equity re-rate: Marvell five years ago was a storage / enterprise networking company; Marvell today is an AI custom-silicon and electro-optics company.

For an AI-stack investor: AWS Trainium volume is the swing variable. If Anthropic's Rainier cluster (1m+ Trainium 3 chips through 2027) builds out on AWS's current timeline, Marvell's AI revenue line approximately doubles between FY26 and FY28. The May 16 AMD cross-investment disclosure ($6.5M for an equity stake) is the second-order validation — AMD is signalling that Marvell's optical-DSP technology is part of the MI400-generation networking roadmap, and the

strategic interpretation matters more than the dollars.

The financial print

Q4 FY26 (reported March 6 2026, quarter ended Feb) printed a beat across rev and EPS with the AI custom silicon and electro-optics businesses both contributing, and the guide was raised on AWS Trainium 3 ramp visibility. The market reaction has been dramatic: the stock is +135% over the trailing three months, +22% in the trailing month, and just touched an all-time high in mid-May. Goldman raised its PT to $125 from $100 on May 15 — and the stock at $181 is now well above the average sell-side PT, which sits closer to $135-145. The buy-side is ahead of the sell-side by a wide margin.

The fundamental backdrop into Q1 FY27: data-center revenue at Q4 ran approximately $1.4B (+78% YoY), with custom AI silicon contributing roughly $700-800M of that line and electro-optics contributing approximately $400-500M. Storage controllers and enterprise networking were the residual contribution, both declining in absolute dollar terms as the AI lines have scaled. Gross margin came in at 62.5% non-GAAP — pressured from the prior-year 65% level by mix-shift toward higher-volume lower-margin custom-silicon products. Operating margin held at approximately 28% non-GAAP, with the company maintaining its long-term targets in the 32-35% range as scale economics improve through 2027.

Q1 FY27 prints May 27 after-hours. Consensus sits at approximately $2.1B revenue and roughly $0.85 of EPS, but with a +135% rally into the print the bar is clearly higher than consensus — anything less than a sizeable beat-and-raise risks a sell-the-news reaction. The May 21 / May 27 timing window matters because of the cross-vendor signal embedded in the May 16 disclosure that AMD invested $6.5M in Marvell — small dollars, but the optical-link cross-validation matters strategically.

The positioning into the print is the principal tactical risk. Marvell has been one of the top three best-performing US semiconductor names of 2026 year-to-date; institutional positioning is materially long versus historical baselines; the consensus PT cohort is well below spot. This is the classic crowded-long setup into an earnings event — the binary is between (a) a clean beat-and-raise that produces $200-210 and a fresh leg higher, or (b) an in-line print that produces a 15-25% pullback regardless of the underlying fundamentals.

Customer mix

AWS is the single largest customer at approximately 25-30% of revenue, with Trainium custom-silicon plus electro-optics revenue combined. Microsoft is rising — currently 10-15% of revenue, potentially the largest customer by FY28 if the rumoured Maia design wins land. The optical-DSP business is broadly distributed across the transceiver-module makers (Coherent, Lumentum, Innolight, Sumitomo Eudyna, Hisense Broadband) and is structurally less concentrated. Customer concentration in the AI custom-silicon line is currently the largest financial risk, and the principal reason Marvell trades at a discount to Broadcom on forward multiples.

Competitive context

Broadcom is the dominant competitor in custom AI silicon — larger team, longer design library, more packaging integration. Marvell's structural advantage is the electro-optics adjacency (Broadcom owns SerDes and DSP but Marvell owns more of the merchant optical-DSP market)

and the AWS Trainium incumbency that is now multi-generational. Alchip and GUC are the Taiwanese ASIC houses that compete on the back-end of custom-silicon design but typically lack the front-end architectural co-design depth Broadcom and Marvell offer. In optical DSPs, the competition is Inphi (which Marvell acquired in 2020), MaxLinear (smaller-scale), and Broadcom's own optical-DSP line — Marvell maintains roughly 40-50% merchant share at the leading 100G/200G PAM4 generation.

The AMD cross-investment disclosure on May 16 is significant. AMD does not invest in semiconductor companies frivolously; the most likely interpretation is that AMD is validating Marvell's optical-DSP technology for MI400-generation interconnect and is publicly signalling commitment to a non-NVIDIA networking option. That single data point is worth a leg of the bull case.

Terminal risk

The terminal risk is the same as Broadcom but more acute: a single major customer pulling a design generation in-house. AWS has the internal Annapurna design team that originated Graviton, and the long-tail question of whether Trainium 4 or 5 stays at Marvell or moves to AWS-internal is the multi-year variable. If AWS pulls Trainium 4 in-house, Marvell's AI custom-silicon revenue takes a multi-quarter air-pocket that cannot be filled by Microsoft Maia ramp alone. The current sell-side bull cases mostly assume Trainium remains at Marvell through 2028; that assumption is doing a lot of work.

Bull case

The base case is FY27 revenue of $9.0-9.8B (FY26 prints around $7.0B), AI custom-silicon revenue at $3.5-4.0B (up from $2.2B in FY26), electro-optics holding $2.0B+, and non-GAAP EPS at $4.20-4.80 versus FY26 around $3.10. At 35x forward — generous but not unreasonable for the cohort — that prints the stock at $150-170, roughly in line with spot but with multi-year compounding.

The upside case is Microsoft Maia 200 design win confirmed in 2026, Trainium 3 ramp accelerating on Anthropic Rainier build, and a third hyperscaler design win (Meta on a future MTIA generation, or Apple inference) lifting the AI custom-silicon line above $5B by FY28. That case prints the stock at $230-260 in 2027 on FY28 EPS of $6.00-7.00.

The AWS Trainium 3 ramp is the principal volume driver. Anthropic's Project Rainier — disclosed in 2024 as a 1m+ Trainium-chip supercluster build across multiple AWS regions — is the largest single Trainium 3 commitment and the principal reason Marvell's AI custom-silicon revenue line is doubling between FY26 and FY28. AWS volume commentary through 2025-26 has been consistent: Trainium 3 production capacity is sold out for 2026, deployments are running ahead of plan, and Trainium 4 design work is already underway with Marvell on a continuing-engagement basis. The continuity of the Marvell-AWS relationship through multiple generations is the structural advantage that distinguishes Marvell from a pure design-services vendor.

Gap / bear case

Three concerns the market may be glossing over. First, the +135% in 10 weeks has compressed the risk-reward — consensus PT of $135 sits 25% below spot, meaning the marginal sell-side analyst would need to raise PTs above $200 for the stock to look reasonable on technicals. Second, the Microsoft Maia design win is rumour, not confirmed disclosure; if Marvell does not announce the win during 2026, the stock loses a third of its rally driver. Third, the gross margin trajectory

— historically 65-66% non-GAAP — has been compressing as the lower-margin custom-silicon revenue grows; FY27 GM consensus is closer to 62-63%, which is a real headwind even if revenue compounds at 35%+.

Optionality

Microsoft Maia design-win confirmation during 2026 — the single biggest discrete option. Apple inference-silicon design-services partnership (rumoured for late 2026 / 2027). And the slow-burn option: 1.6T-class optical DSP volume ramp through 2027-28 as the 800G generation gives way and Marvell maintains share leadership at the leading-edge generation. Each of these is a separate $300-600M annual revenue layer that compounds the AI thesis.

The trade

MRVL — BUY 7.5/10. Entry*: this is the tactically toughest entry in Tier A. Do not chase above $185. Wait for either (a) a clean beat-and-raise on May 27 with the stock breaking out above $200, in which case add on the breakout, or (b) a sell-the-news pullback to $150-160 post-print, which is the better R/R.* ***Position size****: 1.5-2.5% of NLV — half of NVDA / AVGO sizing because of the customer concentration and the recent rally.* ***Stop****: daily close below $145 (the early-March base before the rip).* ***Catalyst date****: May 27 print after-hours; any Microsoft Maia disclosure; Q2 FY27 print late August.* ***Trim/exit triggers****: AWS Trainium 3 ramp slipping into 2027 instead of H2 26; gross margin guide below 62%; AMD investment narrative reversed.* ***Conviction****: real AI custom-silicon franchise, real momentum, but a positioned-long set-up into the print. Treat this as the asymmetric-but-mispositioned trade — own it small, scale on better entry, do not anchor to the +135% rally as the new baseline.*

Layer 2 · TSM — TSMC ADR

One-line thesis

TSMC owns the only sub-5nm logic fab capacity that NVIDIA, AMD, Broadcom, Marvell, Apple, Qualcomm, MediaTek and the AI custom-silicon cohort can physically use to build product through 2027 — the cleanest, deepest moat in the cycle, with capex guided to the high end of $52-56B and the eighth straight beat-and-raise just confirmed.

What TSMC physically does

TSMC manufactures the leading-edge logic wafers — N5, N4P, N3E, N3P, and starting H2 2026 the N2 node with gate-all-around (GAA) nanosheet transistors — that every AI silicon design house ultimately depends on for production. The Hsinchu, Tainan, Kaohsiung, and increasingly Arizona and Kumamoto fabs collectively process roughly 90% of the world's sub-7nm wafers. NVIDIA Blackwell is N4P. NVIDIA Rubin is N3P moving to N2. AMD MI355X is N3E. Apple M5 / M6 and A19 / A20 are N3E moving to N2. Every meaningful AI silicon program — Google TPU, Meta MTIA, AWS Trainium, Microsoft Maia, the OpenAI Titan if/when it ships — is fabricated at TSMC. There is no alternative at the leading edge that can absorb material volume — Samsung Foundry's 3nm process has yield and customer-traction problems that have persisted for three years; Intel 18A is just beginning to ship pilot product; nobody else is close.

Beyond wafer manufacture, TSMC also owns the dominant advanced-packaging franchise — CoWoS-S (current AI workhorse for GPU + HBM stacking), CoWoS-L (Blackwell's 2-die reticle-stitched package), CoWoS-R (rumoured for Rubin), the InFO line for mobile, and the SoIC (System-on-Integrated-Chip) hybrid-bonding platform that will enable 3D logic-on-logic stacking from 2027 onward. CoWoS capacity has been the binding constraint on AI silicon supply for two years running; TSMC's stated capex ramp is meant to lift CoWoS-L monthly wafers from ~50k in 2025 to ~120k by end-2026 and approximately 200k by 2027.

For an AI-stack investor: TSMC is the upstream choke point. Every NVIDIA dollar, every Broadcom dollar, every Marvell dollar of AI silicon revenue first becomes a TSMC wafer-equivalent. The economics of the cycle ultimately settle into TSMC's ASP per wafer, capacity utilisation, and packaging mix.

The financial print

Q1 2026 (reported April 16 2026) printed $35.9B of revenue against $35.5B consensus — beat — at +40.6% YoY in USD terms. Non-GAAP EPS of $3.49 beat $3.26 consensus by 7.1%. This was the eighth consecutive quarter of revenue and EPS beats — a track record matched by essentially no other large-cap semiconductor company. Q1 gross margin printed 58.3%; AI HPC revenue dominated the wafer mix; the company maintained FY26 USD revenue growth guidance above 30%, and pushed full-year capex toward the high end of the $52-56B band — implying $14B+ of quarterly capex through the back half.

The April monthly revenue release (May 10 disclosure) showed NT$410.7B in April, +17.5% YoY, with YTD running +29.9%. Monthly data is the highest-frequency confirmation in the sector; the trend remains intact. The stock at $404 sits roughly 8% below the all-time high. 1m +9%, 3m +13%.

Q2 prints July 17. Consensus is $39-40B revenue and $3.85-4.00 EPS. The setup is similar to the AVGO setup — high consensus expectations into a print with strong monthly data behind it. The most likely surprise is on capex commentary: any uplift in the FY26 capex band toward $58-60B would be read positively by the equipment supply chain (KLAC, LRCX, AMAT, ASML, BESI, ONTO, CAMT) and would confirm the back-half AI ramp visibility.

Customer mix

Apple remains the single largest customer at approximately 22-24% of revenue (M-series Mac, A-series iPhone, M-series iPad). NVIDIA is now the second-largest customer and likely the fastest-growing, at approximately 18-20% of revenue and rising toward 25% on consensus FY27 modelling. AMD, Broadcom, Marvell, Qualcomm, MediaTek collectively contribute another 20-25%. The Chinese fabless cohort (HiSilicon residual, the new Huawei design teams) is now restricted under the US export regime and contributes <5%. The customer base is the most diversified in foundry — Apple's iPhone-cycle volatility, NVIDIA's AI cycle, AMD's PC and AI, and the broader mobile / consumer book all rotate through the same wafer mix.

The technology-node mix is what determines current margin. Q1 26 wafer revenue mix: 5nm-and-below (N5, N4, N3) at approximately 60% of revenue and rising; 7nm at approximately 15% (declining); 16nm and 28nm at the balance for legacy product. The high-performance computing platform — which captures all AI silicon — represented approximately 59% of Q1 26 wafer revenue, with smartphone at 28% (Apple, MediaTek, Qualcomm baseband), IoT at roughly 5%,

automotive at 5%, and consumer at 3%. The HPC mix is unprecedented in TSMC's history (smartphone was the largest segment as recently as 2022) and is the principal driver of the gross margin uplift from 53% in 2023 to 58%+ in 2026.

Competitive context

There is no head-on competitor at the leading edge. Samsung Foundry's 3nm has not closed the yield gap and the principal large customers (Qualcomm Snapdragon, NVIDIA in any meaningful program) have all returned to TSMC after pilot runs at Samsung in 2023-25. Intel Foundry Services (18A) is now shipping pilot Panther Lake silicon and is taking external design wins (rumoured Microsoft, rumoured a defence customer) but is at least three years behind TSMC on volume and at least five on the packaging franchise. Chinese foundries (SMIC) are stuck at 7nm equivalent and barred from EUV imports.

The structural moat is the combination of process technology, packaging integration, and customer-design ecosystem (the OIP — Open Innovation Platform — which has standardised IP libraries across the leading-edge cohort). Even if Samsung or Intel close the process gap, the packaging franchise (CoWoS, SoIC) and the customer-design lock-in are 2-3 year minimum moves to displace. The customer-design ecosystem in particular is under-appreciated: TSMC ships hundreds of process-design-kits (PDKs), IP libraries, EDA tool integrations, and reference-design flows that every fabless customer integrates into its design methodology. Switching foundries means re-validating every IP block, every EDA tool flow, every signoff methodology — a 12-24 month engineering effort that even the largest customers do not undertake casually.

The CoWoS-L packaging franchise — the 2.5D interposer technology that bonds GPU dies to HBM stacks — is the additional moat. TSMC's CoWoS capacity has been the binding constraint on AI silicon supply for two consecutive years; the announced capacity ramp from ~50k wafers/month in 2025 to ~120k wafers/month by end-2026 to approximately 200k wafers/month by 2027 is itself a major industry story (driving equipment orders at KLAC, LRCX, AMAT, BESI, Onto, Camtek). Samsung's I-Cube competing packaging platform has won materially less customer commitment; Intel's Foveros 3D packaging is technically promising but lacks the customer ecosystem to compete at AI scale.

Terminal risk

The terminal risk is geopolitical, not commercial. A Chinese kinetic action against Taiwan would close the leading-edge logic supply chain for the duration of the conflict and force a 2-3 year disorderly re-organization of the global semiconductor map. This is the well-understood tail risk that supports an embedded geopolitical discount in the TSM multiple — typically 15-20% below where a peer-quality monopoly franchise would trade. The Arizona fabs (Phoenix Fab 1 in production, Fab 2 starting H2 26, Fab 3 announced for 2028) are the principal mitigant but are not yet at meaningful percentage of capacity. A 2027-28 Taiwan crisis scenario remains the principal risk-management variable for any TSM position.

Bull case

The base case is FY26 revenue of $130-138B (+30%+ YoY in USD), FY27 revenue of $160-175B (+25%), gross margin holding 56-58%, and the N2 ramp in H2 26 contributing materially through

2027. At a 26x forward — discount to NVDA, broadly in line with the high-quality compounder cohort — that prints the stock at $470-510 in 2027, +16-26% from spot.

The upside case is two-fold: capex guide raises through 2026 toward $60-65B annual run-rate (the equipment supply chain prices this in immediately); and the Arizona Fab 2 ramp accelerating in 2027, enabling US-domestic AI silicon volumes that allow TSMC to capture re-shoring premium pricing. That case prints the stock at $560-620 in 2027.

The pricing-power discussion is the underappreciated bull driver. TSMC has historically raised wafer prices 5-7% per year at the leading edge to recover the rising capex per wafer. The 2024-26 cycle has seen accelerated pricing — N3 wafer ASPs moved from ~$15k initially to closer to $22-25k by Q1 26, with reports of Apple and NVIDIA paying premiums above standard pricing for guaranteed capacity allocation. If the N2 generation prices at 15-20% above N3 (consistent with capex per wafer), and if AI HPC customers continue to pay incremental premium for CoWoS-L capacity allocation, the gross margin trajectory through 2027 is materially above current consensus 56-58%.

The Arizona fab is the secondary leg of the upside case. Fab 1 is in production with limited volume; Fab 2 starting H2 26 is expected to reach 100k wafers/month at N3 by end-2027; Fab 3 announced for 2028 will produce N2 silicon. Once US-domestic AI silicon volumes ramp materially in 2027-28, TSMC is positioned to capture re-shoring premium pricing from NVIDIA, Apple, Broadcom, AMD, and the DoD-adjacent customer set. The Arizona capacity is also the principal mitigant against the Taiwan-strait geopolitical risk, and as that capacity grows the embedded geopolitical discount in TSM's multiple should compress.

Gap / bear case

Two concerns the market may be missing. First, Q2 26 capex commentary may include a moderation signal if Apple iPhone demand softens — Apple is still the largest customer and Q2 typically reflects iPhone build commitments. Second, the FY27 gross margin trajectory depends on N2 ramp yields holding above the 50% bring-up threshold by H2 27 — if N2 yield trajectory slips, gross margin compresses sequentially through 2027 and the multi-year compounding case takes a haircut.

The third concern is the consensus PT setup: TSM PTs are now broadly $480-520, meaning the stock has roughly 20% upside to consensus — generous but not extreme — and a Taiwan-strait headline event remains the principal tail risk.

Optionality

Three options under-priced in consensus. CoWoS-L capacity expansion above guided run-rate in late 2026 (every AI silicon house is asking for more), which directly accretes to the equipment cohort but also lifts TSMC ASP per wafer. SoIC adoption ramping in 2027-28 for hybrid-bonded 3D logic — the first credible volume case is Apple M-series with stacked DRAM around 2027, plus the rumoured NVIDIA Rubin Ultra SoIC variant. And the long-tail option: a Saudi or UAE sovereign-AI capacity buildout that commits to a dedicated TSMC fab outside Taiwan, sometime in 2027-28.

The trade

***TSM — BUY 9/10. Entry**: starter at $400-410 spot; full-size on $370-385 pullback if a Taiwan-strait headline or broader semi correction creates entry. **Position size**: 4-5% of NLV — pair-trade tier with NVDA as the two-name core of any AI cycle exposure. **Stop**: daily close below $345 (200-day MA breach plus reversal of the post-Q1 rally). **Catalyst date**: July 17 Q2 print; June monthly rev release in early July; ongoing monthly data; N2 ramp commentary in H2 26. **Trim/exit triggers**: monthly revenue YoY growth below +15% on a single month; capex guide cut below $52B; any verifiable Apple or NVIDIA leakage to Samsung or Intel. **Conviction**: the highest-quality compounder in the semiconductor universe, with the most defensible moat. The reason this is 9/10 not 10/10 is the geopolitical tail — at $405 the equity already prices that risk at roughly a 20% discount to a hypothetical Texas-located analogue, but the tail is fat enough that conviction cannot exceed 9/10.*

Layer 2 · KLAC — KLA Corporation

One-line thesis

KLA is the semiconductor process-control monopoly — 85% global share in optical-pattern wafer inspection — and the combination of HBM advanced-packaging inspection pull, the 10-for-1 stock split, and a 21% dividend raise alongside a clean Q3 FY26 beat-and-raise makes this the highest-quality “boring monopoly” trade in the cycle.

What KLA physically does

Every leading-edge fab — TSMC N3 / N2, Samsung 3nm, Intel 18A, SK Hynix HBM lines, Micron Boise HBM — runs hundreds of KLA inspection tools at every step of the wafer manufacturing process. The flagship products are optical-pattern wafer inspectors (the 29xx, 39xx, and 89xx series for unpatterned and patterned defect inspection), e-beam inspection and review tools (eDR-7xxx series), reticle-inspection systems for EUV mask qualification, and the suite of overlay and CD-SEM metrology tools that measure whether the silicon is being printed within sub-nanometer tolerances. In rough numbers: a leading-edge logic fab will install $1-1.5B of KLA equipment per 100k-wafer-month capacity layer, and a leading-edge HBM fab installs roughly $300-500M per HBM ramp step.

The economic logic of inspection is asymmetric. A leading-edge wafer at 3nm carries roughly 100-200 die per wafer, with each die selling for hundreds to thousands of dollars at NVIDIA / Apple / AMD ASPs. A single yield-killing defect — a particle, a pattern collapse, an overlay mis-print — that escapes inspection in early process steps destroys the wafer downstream after another $5-10k of process value has been added. KLA tools sit upstream of every process step that costs more than $1k per wafer to perform, with the explicit purpose of catching defects before that value is added. The throughput requirement is high (a single tool scans hundreds of wafers per hour) and the false-positive rate matters (every false flag stops a wafer for re-inspection, costing throughput).

KLA’s near-monopoly is on the optical-pattern wafer inspection segment specifically, where TechInsights and Gartner estimate KLA holds roughly 85% global share — competitors like Hitachi High-Tech, Applied Materials’ VeritySEM line, and Onto Innovation operate in adjacent

metrology niches but do not directly contest the high-volume in-line inspection franchise. The structural reason for the share is the combined cost of the optical illumination system (deep-UV at 193nm wavelength, multi-pass differential interference contrast imaging), the defect-recognition software trained on 25 years of fab-floor data, and the customer-co-development relationships that lock in the next-generation inspection recipes a node before that node ships.

In advanced packaging specifically — the HBM stacking process, CoWoS through-silicon-via inspection, hybrid-bonding alignment metrology — KLA has been adding direct exposure faster than any other equipment vendor. The acquisition of SPTS in 2018 plus the organic build-out of the ICOS bump-inspection line plus the metrology tools that ship into TSMC CoWoS-L lines collectively add another $1.5-2B of advanced-packaging-specific TAM through 2027. The reticle-inspection franchise (the Teron family for EUV mask qualification) is a related growth line — every high-NA EUV scanner from ASML requires a stack of reticle-inspection tools at the fab to verify mask integrity at the sub-nanometer pattern level, and KLA is the dominant supplier with effectively no merchant competitor at the leading edge.

The financial print

Q3 FY26 (reported April 30 2026, quarter ended March) printed $3.42B of revenue against $3.32B consensus — beat by $100M (+11% YoY) — and non-GAAP EPS of $9.40 also beat. The company raised the FY26 guidance bracket to “high-teens” revenue growth on the strength of AI, advanced packaging, and memory pulls. The Board also announced a 10-for-1 stock split (effective late Q2 FY26) and a 21% dividend hike — twin signals that management views the current revenue trajectory as durable. The stock at $1,892 sits roughly 2% below the May 12 all-time high of $1,939; 1m +12%, 3m +28%, +155% off the 52-week low.

The 10-for-1 split is more than cosmetic. KLA’s nominal share price near $1,900 had become a meaningful friction for retail flow, options activity, and inclusion in equal-weight indices. Post-split, the $190-equivalent stock price unlocks retail participation that historically corresponded to a one-time multiple expansion of 100-200 basis points across the comparable semiconductor cohort (LRCX, AVGO did the same in 2024). The dividend raise is the conservative complementary signal — management is committing to a payout ratio increase rather than aggressive buybacks, telegraphing capital-return discipline rather than financial engineering.

Q4 FY26 prints July 30. Consensus sits at approximately $3.55B of revenue and $9.80-10.00 of non-GAAP EPS. The setup is constructive — KLAC has rerated to the high-quality compounder multiple tier (35-37x forward) and any beat-raise extends that.

Customer mix

TSMC is the largest single customer at approximately 25-30% of revenue, with Samsung Foundry, SK Hynix, Micron, Intel, and the Chinese foundries (SMIC, YMTC pre-restrictions) collectively contributing another 50%. The remaining 20-25% is the broader memory cohort (Kioxia, Western Digital, Samsung Memory) plus the analog / specialty foundries (Tower, Vanguard, GlobalFoundries). The China exposure is real but declining — KLAC has been the most affected by export controls of any non-EUV equipment vendor — but the offset is the advanced-packaging-specific pull at TSMC and the HBM expansion at SK Hynix / Micron, both of which are running ahead of plan.

The geography split is the more useful disaggregation. Roughly 38-42% of revenue is Taiwan

(principally TSMC), 18-22% is Korea (Samsung Foundry, Samsung Memory, SK Hynix), 12-15% is China (declining from over 25% pre-restrictions), 12-15% is North America (Intel, Micron, GlobalFoundries), and the balance is Japan, Europe, and Southeast Asia. The Taiwan and Korea weighting puts KLAC squarely in the centre of the AI silicon manufacturing geography — every leading-edge AI silicon die ultimately passes through KLA inspection at TSMC, and every HBM stack passes through KLA inspection at SK Hynix or Samsung Memory or Micron Boise.

Competitive context

There is essentially no head-on competitor at the optical-pattern wafer inspection tier. Applied Materials' inspection portfolio (the VeritySEM / e-beam lines) competes in adjacent segments but does not directly contest the high-volume optical inspection tier. Hitachi High-Tech is meaningfully present in Japan-based customers (Kioxia, Renesas) but has not won materially at TSMC, Samsung, or SK Hynix. Onto Innovation, which is also in this Tier A list, is in the bump-inspection / advanced-packaging adjacent niche and is more an ecosystem partner than a competitor at the leading-edge wafer-inspection tier.

The structural defence is the volume-of-defect-recognition-data moat. KLA tools generate hundreds of millions of defect images per quarter; the recipes shipped to new fabs are pre-trained on this aggregated data and ship at production-ready maturity from day one. A competitor with even a technically superior point-tool cannot replicate that data moat in less than five years.

Terminal risk

The terminal risk is twofold. First, China export restriction tightening — KLAC has historically derived roughly 25% of revenue from China and the marginal restrictions on advanced-process equipment have already shaved that line; further tightening (which is broadly the policy direction under the current US administration) cuts a few hundred million of revenue per round. Second, the long-tail risk that as AI logic moves to 2nm and below, a fundamentally different inspection paradigm (multi-beam e-beam, or X-ray) displaces optical-pattern inspection — Hitachi and TEL have been investing in multi-beam e-beam for a decade, and a credible technology transition is the only scenario where KLAC's monopoly compresses.

Bull case

The base case is FY26 revenue of $13.5-14.0B (high-teens growth on $11.7B FY25), FY27 revenue of $15.5-16.5B, non-GAAP EPS rising from $36 in FY26 to $42-44 in FY27. At 38x forward — premium to the equipment peers but consistent with KLA's monopoly franchise — that prints the stock at $1,600-1,670 pre-split, $160-167 post-split — broadly flat to spot but with capital return underpinning. The base case is not where the alpha is.

The upside case is HBM packaging inspection ramping faster than the FY27 base assumes — the SK Hynix and Micron HBM4 lines collectively add roughly $500-700M of KLA TAM through 2026-27, and if KLA captures full share (likely, given the customer relationships) that revenue is upside to consensus. CoWoS-L and CoWoS-R inspection at TSMC adds another $300-500M. The split-driven retail flow plus broader equal-weight index buying creates a 5-10% multiple expansion. Aggregate upside case: stock at $2,200-2,400 in 2027 (or $220-240 post-split), +16-27% from spot.

The structural argument for KLAC carrying a premium multiple to LRCX, AMAT, and the broader equipment cohort is the recurring-service-revenue mix. KLAC's installed-base service revenue runs at approximately 30%+ of total revenue and grows in the high single-digits annually regardless of capex cycle — software updates, defect-recipe libraries, tool calibration services, and the broader customer-support business. This is structurally the closest thing in the equipment sector to a software-like recurring revenue model, and the equity has historically traded at a 20-30% multiple premium to LRCX and AMAT to reflect this. The 38x forward multiple at current levels sits at the high end of the historical premium range but is not at extremes.

Gap / bear case

Three concerns. First, the +155% off 52-week low has stretched the multiple — at 38x forward, KLAC is at the upper end of its 5-year range and a multiple compression back toward 32x on any beat-miss or guide-down prints the stock at $1,500 pre-split. Second, the DRAM cost dynamics that benefit SK Hynix and Micron equity also pressure DRAM-maker capex discipline — if the memory cycle softens in 2027 (which is the consensus path now that it has run hot for two years), the memory equipment book of KLAC takes a haircut. Third, the China-export-restriction overhang remains live — the latest round of export-controls in late 2025 specifically called out inspection equipment as a target category, and continued restrictions through 2026-27 compress KLAC's China book at the same time the broader equipment cohort is benefiting from the AI-cycle re-rate.

The +155% off lows also raises the question of what the marginal incremental buyer's thesis is. The stock has run from the 52-week low principally on AI-cycle multiple expansion plus consensus EPS revisions; a continued re-rate from current levels requires either (a) consensus EPS revisions higher than what the company is currently guiding, or (b) further multiple expansion from already-elevated levels. The 10-for-1 split-driven retail flow is the principal mechanism for option (b); without it, the asymmetric upside from current levels is more modest than the trailing 12-month return would suggest.

Optionality

Four options. Reticle-inspection demand from EUV mask qualification at TSMC, Samsung, and Intel scaling with the high-NA EUV (NXE:5000) installed base — each high-NA tool requires multiple reticle-inspection passes per wafer cycle, and the high-NA install base is expected to grow from approximately 10 systems in 2025 to 40+ by 2027. Hybrid-bonding alignment metrology — KLA has not loudly disclosed its hybrid-bonding franchise but the IP base sitting at the intersection of overlay + CD-SEM + bond-pad inspection is the natural extension. The split-driven retail-flow option: post-10:1 split, the equal-weight index inclusion plus retail derivatives liquidity historically add a one-time 100-200bp multiple expansion that has not yet fully arrived in the stock.

The fourth option is the CoWoS-L and CoWoS-R inspection ramp at TSMC. CoWoS capacity expansion through 2026-27 is the single largest equipment-pull line at TSMC, and KLA captures roughly $300-500M of incremental TAM per CoWoS capacity step. This is upside to consensus FY27 models that have been calibrated to the H2 25 TSMC capex announcement; subsequent capex revisions higher (which the supply chain has been signalling) directly flow to KLA revenue.

The trade

KLAC — BUY 8.5/10. ***Entry***: *starter at $1,890-1,900 spot; size up on $1,750-1,800 pullback. Post-split entry (post-late-Q2-FY26) at $180 equivalent.* ***Position size***: *2.5-3.5% of NLV as the highest-quality equipment monopoly position.* ***Stop***: *daily close below $1,650 pre-split / $165 post-split (200-day MA breach).* ***Catalyst date***: *July 30 Q4 FY26 print; the 10-for-1 split execution date in Q2 FY26; September Apple iPhone launch as a sentiment driver for the broader semi cohort.* ***Trim/exit triggers***: *high-teens FY26 guide cut to mid-teens; any China-specific tightening removing >$500M annual revenue; multiple expansion above 40x forward without earnings revision.* ***Conviction***: *the cleanest "boring monopoly" trade in the equipment cohort. The 10:1 split + 21% dividend hike is management telling you the next 24 months of revenue trajectory is durable. Own it in size.*

Layer 2 · LRCX — Lam Research

One-line thesis

Lam Research owns the etch and deposition franchise that physically makes HBM through-silicon vias and 3D NAND stacks — the two highest-growth equipment lines in the cycle — and the Q3 FY26 record print plus the SK Hynix / Samsung HBM TSV capacity ramp through 2027 makes this the equipment-pair-trade alongside KLAC.

What Lam Research physically does

Lam makes the plasma-etch tools (the Kiyo, Flex, and Versys lines for conductor and dielectric etch) and the deposition tools (Striker ALD, Vector PECVD, ALTUS tungsten CVD) that perform two specific operations in semiconductor manufacture: removing precisely-defined volumes of material from the wafer surface (etch) and laying down precisely-controlled thin films of material on the wafer surface (deposition). At leading-edge nodes these steps run hundreds of times per wafer, and the binding-constraint volume for AI memory specifically comes from two product lines.

First, the deep-silicon etch tool — the Sense.i Plus Bosch-process etcher — is the principal piece of equipment that creates through-silicon vias (TSVs) for HBM. Every HBM die has thousands of TSVs running vertically through the silicon to connect adjacent layers in the stack; etching those vias to ~5µm diameter and ~50µm depth at 100,000+ vias per die per hour throughput is what Sense.i Plus does, and Lam holds approximately 80%+ market share in this specific equipment niche. The physics are not trivial. The Bosch process alternates between an SF_6-based etch step and a C_4F_8-based passivation step at sub-second cycle times, producing scallop-free vertical sidewalls at aspect ratios exceeding 10:1. Doing this at production throughput on 300mm wafers carrying tens of thousands of vias each is the operational moat that has kept Lam dominant against TEL (which competes in the broader silicon-etch market but has not won meaningfully at the HBM-specific TSV step) and Applied Materials (which competes at logic etch but has minimal HBM TSV presence).

Second, the high-aspect-ratio etcher for 3D NAND vertical channels — the Flex-D family — performs a similar function at higher aspect ratios for 3D NAND vertical channels, where Lam holds 60-70% share against TEL. Modern 3D NAND structures at Kioxia, Western Digital BiCS, and

Samsung V-NAND now stack 400+ layers vertically, with vertical channel aspect ratios exceeding 70:1 — physical requirements that have driven a multi-decade Lam-TEL technology race where Lam has held the share leadership since the 256-layer node.

The third critical line is atomic-layer deposition (ALD) for high-k metal-gate and capacitor films at advanced logic and DRAM nodes — the Striker family — where Lam competes with TEL and Applied Materials for share in a faster-growing TAM. ALD scales materially with both DRAM node migration (1z to 1a to 1b to 1c) and with leading-edge logic capacitor work, and Lam has been investing heavily in expanding the ALD portfolio.

The fourth segment is the broader conductor-etch (the Kiyo platform) and dielectric-etch (Flex platform) franchises, which serve every leading-edge logic node. As TSMC ramps N2 and Intel ramps 18A, these tools are pulled through capacity expansions at multi-billion-dollar order volumes per fab.

For an AI-stack investor: every HBM die ever made in production volume has passed through Lam Sense.i Plus etchers. The HBM cycle is mechanically pulling Lam revenue through every wafer increment, and the company's tool-utilisation data is one of the cleanest leading indicators of HBM capacity build-out in the entire equipment cohort.

The financial print

Q3 FY26 (reported April 22 2026, quarter ended March) printed record revenue of $5.84B (+24% YoY) and EPS of $1.47 against $1.36 consensus — beat by 8%. The company raised June-quarter guidance on AI memory etch demand. Gross margin held at 48.2% on a normalized basis. The Q3 result is meaningful because Lam's revenue base mix has flipped over 24 months — memory equipment now accounts for roughly 50%+ of the equipment line (versus historically 30-40%), with HBM-specific tool orders running well above pre-cycle peak.

The composition of the Q3 revenue is worth dissecting. System revenue grew 27% YoY to roughly $4.0B; customer support business operations (the installed-base service revenue) grew 18% YoY to approximately $1.85B. The service revenue — which historically runs at a 60%+ gross margin and is structurally non-cyclical — has been quietly compounding at mid-teens for several years and represents an increasing portion of the steady-state earnings power. Operating margin came in at 32.1% non-GAAP, the highest quarterly level since the 2022 cycle peak. The capital return announced alongside the print included continued aggressive buybacks (roughly $1.5B per quarter through 2026) plus a dividend hike that complements the operating leverage.

The stock at $285 is 1m +8.4%, 3m +24% (note: LRCX completed its 10-for-1 split in October 2024, so all share counts and historical comparison points are on the post-split basis). Q4 FY26 prints July 22. Consensus sits around $6.0-6.2B revenue and $1.60+ EPS. The same pattern as KLAC applies — the bar is high but the underlying demand visibility through 2027 is the strongest in two decades. The principal swing variable into the print is the China-revenue line: Lam's China revenue at Q3 ran approximately $1.5B (declining from the cycle peak of ~$2.2B in fiscal 2024) and the consensus FY27 model assumes that line stabilises rather than deteriorates further. Any further restriction tightening compresses that line on the next print.

Customer mix

SK Hynix is now one of the largest individual customers — likely 12-15% of revenue currently, driven by HBM3e and HBM4 line buildouts. Samsung Memory is comparable at 12-15% of rev-

enue, with HBM4 PyeongTaek line build-out as the principal driver. Micron is approximately 10-12% (Boise / Idaho HBM ramp). TSMC and Intel together represent another 15-20% (logic etch + ALD demand for N3 / N2 / 18A). Kioxia / Western Digital is the 3D NAND etch demand layer. The customer concentration is similar to KLAC but with a higher memory tilt — which means the equity has more memory-cycle beta than KLAC.

The memory-cycle beta needs to be understood in detail. Memory-equipment revenue at Lam swung from approximately 30% of total in 2023 (the trough) to 50%+ in 2026 (the current peak). The principal cycle question is whether the 2026-27 memory order book holds or compresses. SK Hynix and Micron HBM capex is contracted under LTAs and effectively non-cyclical through 2027; Samsung HBM4 capex is more discretionary but tied to NVIDIA qualification timing rather than memory ASPs; Kioxia / WD 3D NAND capex is the most cyclical layer and is the most exposed to a NAND ASP correction. Aggregate exposure: roughly 60-70% of the memory book is HBM-linked (low cycle-beta) and 30-40% is NAND-linked (higher cycle-beta).

Competitive context

Applied Materials is the largest competitor across the broad equipment portfolio but only meaningfully competes with Lam in select etch and deposition niches. TEL (Tokyo Electron, 8035.T) is the principal direct competitor in high-aspect-ratio etch for 3D NAND and in select ALD applications — TEL has been gaining share at Kioxia and at the Chinese memory makers. ASM International is a competitor in ALD specifically. The structural advantage Lam holds is the installed-base service revenue (roughly 35% of revenue, growing 15% YoY) and the multi-decade customer co-development partnerships at the HBM lines.

The HBM-specific etch franchise is the deepest moat — Sense.i Plus has been refined across HBM2 / HBM3 / HBM3e generations with customer-specific recipes, and the install base at SK Hynix M16 and Samsung PyeongTaek is too embedded to be displaced by a TEL or AMAT tool on a single product transition.

Terminal risk

Two terminal risks. First, the same China-export tightening that affects KLAC — Lam has roughly 20-25% China revenue exposure (declining post-restrictions) and further tightening compresses that line. Second, the longer-tail risk that as 3D NAND layer counts saturate around 400+ layers, the etch-tool replenishment cycle slows — though the offset is that HBM stack heights are simultaneously rising (12-Hi to 16-Hi to 20-Hi) and TSV density is growing per die, which sustains Lam volume even if NAND moderates.

Bull case

The base case is FY26 revenue of $22-23B (versus FY25 around $17B), FY27 revenue of $25-27B, gross margin holding 48-50%, and non-GAAP EPS rising from FY26 around $5.80 toward FY27 $6.80-7.20. At 28x forward — discount to KLAC, broadly in line with the equipment cohort — that prints the stock at $190-200 post the well-flagged 10-for-1 split that came earlier in 2024, equivalent to current levels with modest upside.

The upside case is HBM4 + HBM5 capacity additions running well above current sell-side modelling — Trendforce now forecasts HBM unit growth of +60-80% in 2026 with another step in 2027 — which lifts Lam HBM-specific revenue by an incremental $1-1.5B annually through 2027.

The 3D NAND layer-count ramp at Kioxia, Western Digital, and the Samsung NAND side adds another $400-700M. Combined, the upside case prints FY27 EPS toward $8.00 and the stock at $310-340 in 2027.

The HBM3e to HBM4 generational transition is the principal near-term driver. The HBM4 die ships in 12-Hi configuration mainstream and 16-Hi configuration premium, with each step requiring proportionally more TSV etches per die. The base-die migration from a DRAM process to a 5nm logic process at TSMC (an SK Hynix decision for HBM4 specifically) requires additional CD-SEM and metrology pull-through that benefits Lam's adjacent ALD and conductor-etch lines. The aggregate Lam revenue per HBM4 wafer is approximately 25-30% higher than the equivalent HBM3 wafer — a structural mix improvement that compounds the unit-volume growth.

Gap / bear case

Three concerns. First, +24% in 30 days has compressed the risk-reward; the stock pulled back 4.5% on May 15 on macro / yields concerns, which is a small reminder that even high-quality equipment names trade with cyclical equity beta. Second, the FY26 guidance has tightened to "growth on AI demand" rather than explicit quantification — consensus has moved up the EPS revisions but the company has been cautious about extrapolation, and any sequential softening in the memory order book (which is possible if SK Hynix or Samsung pulls in CapEx) would generate a multi-quarter air-gap. Third, the China-export-restriction overhang remains live in much the same way as KLAC — Lam has been the most affected of the broad equipment cohort over 2022-25, and continued restrictions through 2026-27 compress the China revenue line.

The fourth concern relates to share dynamics specifically. While Lam holds 80%+ share at HBM TSV etch, TEL has been investing in competing tools and has been gaining share at the 3D NAND high-aspect-ratio etch tier — Kioxia / Toshiba memory has been historically a TEL-leaning customer, and the Samsung NAND book is more competitively distributed than the HBM line. If TEL captures incremental NAND etch share through the 2027 generation transition, Lam's NAND-equipment revenue compresses against expectations.

Optionality

Three options. Hybrid-bonding integration alongside BESI / ASMPT — Lam has IP exposure here through the LRCX-BESI joint development on the HBM hybrid-bonding back-end. Continued China-domestic memory build-out (CXMT, YMTC) generating equipment demand at the tools that remain unrestricted. And the M&A optionality at the equipment-vendor consolidation level — both LRCX-AMAT and LRCX-BESI scenarios have been floated by sell-side over 18 months; while neither is base case, either becomes a one-time premium event.

The trade

LRCX — BUY 8/10. Entry*: starter at $285 spot; size up on $260-270 pullback into the 50-day moving average.* ***Position size****: 2.5-3% of NLV as the etch/deposition equipment leg, paired against KLAC as the inspection / metrology leg.* ***Stop****: daily close below $245 (200-day MA breach and prior structural support).* ***Catalyst date****: July 22 Q4 FY26 print; September SEMICON West commentary on memory order book; any Trendforce or SemiAnalysis HBM capacity revision.* ***Trim/exit triggers****: memory-equipment order book softening below 25% YoY; gross*

margin guide below 47%; any verifiable China-restriction tightening that removes >$1B of annual revenue. ***Conviction****: the cleanest HBM TSV etch monopoly and the second-best equipment franchise in the cycle after KLAC. Treat LRCX + KLAC as a pair — owning both at 5-6% combined gives you the equipment-cycle exposure without single-name concentration risk.*

Layer 2 · CAMT — Camtek Ltd.

One-line thesis

Camtek is the pure-play HBM bump-inspection franchise — a small-cap with the cleanest single-thesis HBM-cycle exposure in the equipment cohort, the post-Q1 -7% sell-off creates an entry, and the H2 26 ramp guided at >25% above H1 is the cleanest forward-visibility in the small-cap space.

What Camtek physically does

Camtek makes optical-inspection tools that examine the μbumps (microscopic solder bumps, typically 25-40μm in diameter) and the through-silicon-via openings on HBM die wafers and on the underlying logic + interposer wafers before, during, and after the stacking process. The flagship product is the **Eagle G6** automated optical inspection platform (and the newer **Eagle XT** variant) — a multi-camera, multi-illumination optical inspection tool that scans 300mm wafers at high throughput looking for bump-height anomalies, bump-coplanarity defects, via-opening blockages, residue contamination, and any defect that would cause a TC-bonded HBM stack to fail mechanically or electrically.

The physical inspection challenge is non-trivial. A 12-Hi HBM3e stack contains roughly 12,000 μbumps per layer, with each bump approximately 25-30μm in diameter and the entire bump array required to remain within ±1μm coplanarity tolerance across the full die surface. A single anomalous bump destroys the entire stack — and HBM stacks cost upwards of $300 each at HBM3e generation, rising toward $500-700 at HBM4. The yield economics make in-line bump inspection an obligatory step rather than a sampling step, which is why every HBM-producing fab runs Camtek (and Onto) tools at multiple inspection passes per wafer cycle.

In the HBM stacking flow specifically, Camtek tools sit at three inspection steps: pre-bond (inspecting the bumps on each die before stacking begins), post-bond (verifying the bond integrity after each layer), and incoming-wafer inspection at the OSAT for CoWoS-L underfill quality. In CoWoS-L specifically — the packaging platform for NVIDIA Blackwell and Rubin — Camtek tools also inspect the redistribution layer (RDL) bumps at the silicon interposer level, and the next-generation Eagle platform under development carries explicit hybrid-bonding-pad alignment metrology to position for the HBM5 / HBM6 hybrid-bonding transition.

In rough numbers, a single HBM4 capacity expansion step (e.g., a new 25k-wafer-month HBM4 line at SK Hynix M16) requires roughly $20-30M of Camtek equipment, plus several million of recurring service revenue per year for software updates, recipe development, and consumables. The Eagle platform is the dominant tool at this inspection tier and Camtek holds approximately 70-80% share among the HBM customer set, with a particularly deep installed-base position at SK Hynix that has been built across HBM2, HBM3, and HBM3e generations and is now extending into HBM4.

The product roadmap matters: the Eagle XT introduced in late 2025 added higher-throughput multi-camera scanning specifically designed for the 16-Hi HBM stacks coming on HBM4 generation, and the next-generation tool (rumoured Eagle G7 or under a new product name) will integrate the hybrid-bonding alignment metrology under development. Camtek's bet on the hybrid-bonding inspection paradigm is the principal R&D priority and the principal terminal-risk mitigation.

The financial print

Q1 2026 (reported May 12 2026) printed revenue of $121.7M against $119.7M consensus — beat — and EPS of $0.70 against $0.69 (slight beat). The guide for Q2 was raised to $129-131M and management commented that the H2 2026 revenue should run >25% above H1, on the back of HBM4 capacity additions at SK Hynix, Samsung, and Micron. The stock sold off -7.4% pre-market on the print, principally on operating margin compression to 22.4% (versus historical 28%+) — a reflection of mix-shift toward the lower-margin HBM-specific tools and increased R&D ahead of HBM5 / hybrid-bonding inspection tool development.

The margin compression deserves attention. Camtek's historical gross margin has run 49-51% non-GAAP with operating margin in the 28-32% range. The Q1 print at 22.4% operating margin is roughly 600 basis points below the prior-year quarter — driven by three factors: (1) a one-time mix-shift toward HBM-specific tools that carry lower margins than Camtek's historical advanced-packaging products, (2) elevated R&D spend related to the hybrid-bonding inspection roadmap and the next-generation Eagle XT platform, and (3) higher engineering deployment costs as Camtek scales its customer-support headcount into Korean and Taiwanese HBM lines. Management has guided that operating margin should recover toward 25-27% in H2 2026 as the volume ramp normalises and the R&D step-up moderates. Whether that guidance holds is the principal financial-risk variable.

The stock at $192 is 1m -7.4%, 3m +5% — meaningfully behind the broader semi equipment cohort over the same window. The underperformance is doing the entry-point work: the relative compression versus KLAC (+28% over the same window), LRCX (+24%), and ONTO (+18%) creates the asymmetric R/R that the bull case requires. Q2 prints August 11. Consensus is the guided $129-131M revenue range plus EPS around $0.85-0.95. The setup post-pullback is materially better — the stock has compressed back to a multiple closer to ONTO and BESI on a forward-earnings basis, and the H2 revenue ramp visibility is the most concrete in the equipment cohort, with the SK Hynix HBM4 line build-out, the Samsung PyeongTaek HBM4 ramp, and the Micron Idaho IPF1 capacity step all contributing through the August-October order book.

Customer mix

SK Hynix is the largest customer at approximately 35-40% of revenue (HBM3e and HBM4 line inspection), with Samsung Memory at 20-25% (PyeongTaek HBM4 build-out), Micron at 15-20% (Idaho HBM ramp), and TSMC + the OSAT cohort (ASE, Amkor, SPIL) at the balance — roughly 15-20% combined for CoWoS-L underfill and RDL inspection. The customer concentration in HBM is real but the underlying demand is structurally underpinned by the HBM unit growth forecast of 60-80% in 2026.

Competitive context

The direct competitor is **Onto Innovation** (also in this Tier A list) — Onto's bump-inspection tools (the Dragonfly platform inherited from Rudolph Technologies and the Atlas in-die metrology) compete head-on with Camtek's Eagle line. Onto is structurally larger and has the broader product portfolio (metrology + inspection across multiple application areas) while Camtek is the pure-play HBM bump-inspection specialist with the deeper customer-specific recipe library.

The competitive dynamic is closer to a duopoly than a single-vendor monopoly: Camtek + Onto collectively hold ~90%+ of the global bump-inspection market, with Hitachi High-Tech and a handful of Korean local vendors (Kornic, Park Systems) in single-digit share niches. The two companies divide the customer set: Camtek is dominant at SK Hynix, Onto is dominant at Micron, both compete at Samsung. The structural valuation question — discussed under Onto below — is whether the duopoly compresses to a single winner over the next 3-5 years or whether the HBM cycle is large enough to sustain both at premium economics.

Camtek's structural advantages within the duopoly are three. First, the SK Hynix relationship is multi-decade — Camtek's tools have been the reference inspection platform at the M14 / M16 HBM lines through HBM2 / HBM3 / HBM3e generations, and the recipe-development relationship is locked in at a level that an Onto displacement attempt cannot easily replicate inside a single product transition. Second, the Israeli engineering base (Camtek is headquartered in Migdal Haemek, Israel) operates with materially lower fixed costs than Onto's New York / Massachusetts cost structure, contributing to historical gross margin advantages. Third, the product focus is sharper — Camtek's R&D is entirely on the bump-inspection / advanced-packaging vertical, while Onto's spend is divided across bump-inspection, in-die metrology, lithography, and the broader portfolio.

The downside of the focus is precisely the customer concentration risk: in any scenario where SK Hynix HBM share compresses (per the Hanmi share-loss discussion, ASMPT is taking marginal HBM4 socket at SK Hynix), Camtek's customer exposure is more concentrated than Onto's. The H1 26 print suggested this is not yet impacting the order book — but the second-derivative read on SK Hynix capex through 2027 is the variable to monitor.

Terminal risk

The terminal risk is the same as Hanmi's at the equipment-cohort level: hybrid bonding adoption for HBM5 / HBM6 generations changes the inspection paradigm. If HBM moves to hybrid bonding (BESI-led, ASMPT-led) starting in 2027-28, the µbump inspection franchise compresses as fewer wafers carry traditional µbumps. Camtek's mitigation is the hybrid-bonding alignment metrology roadmap (under development); whether Camtek can win at the next-generation inspection tier is the binary question on a 3-5 year horizon.

Bull case

The base case is FY26 revenue of $530-550M (versus FY25 around $430M), FY27 revenue of $620-680M, gross margin holding 48-50% and operating margin recovering from the Q1 22% trough toward 28-30% as the H2 mix normalizes. At 30x forward EPS — premium to historical Camtek multiple but consistent with HBM-pure-play exposure — that prints the stock at $235-265 in 2027, +22-38% from spot.

The upside case is HBM unit growth landing at the +80% end of the Trendforce range and Camtek capturing full share at HBM4 + early HBM5 — which lifts FY27 revenue toward $750M and EPS toward $7.00. At 32x forward, the stock prints $260-290 in 2027.

The smaller-cap nature of Camtek versus its competitors creates additional upside leverage on share-take scenarios. A single 25k-wafer-month HBM4 capacity expansion at Samsung Pyeong-Taek would contribute roughly $20-30M of incremental Camtek tool revenue — meaningful at the ~$500M base. If Samsung's HBM4 program ramps materially faster than the consensus path (which assumes Samsung remains a quarter behind SK Hynix through 2026-27), Camtek's revenue line takes a step-function uplift that the small-cap valuation amplifies. Conversely, the small-cap nature also means a single program cancellation has 5-10% revenue impact — the asymmetric exposure cuts both ways.

Gap / bear case

Three concerns. First, operating margin compression to 22% in Q1 is the warning sign that the customer mix is shifting toward lower-margin tools — if this persists rather than normalising, the FY26 EPS print disappoints despite revenue beating. Second, the duopoly with Onto means pricing discipline is harder to maintain than at single-vendor franchises like KLAC. Third, valuation premium to ONTO has historically run 20-30% on forward multiples and is currently elevated — a re-rating toward Onto-level multiples would compress Camtek by 15-20%.

Optionality

Two real options. First, the SK Hynix Wide TC Bonder qualification cycle in H2 26 brings net new tool order volume that Camtek captures in inspection alongside Hanmi's bonder side. Second, the hybrid-bonding inspection roadmap — if Camtek can credibly position for the HBM5 generation alongside BESI hybrid bonding, the long-tail terminal risk diminishes and the multiple expands.

The trade

CAMT — BUY 8/10. Entry*: $190-195 post-Q1-pullback — a better entry than pre-print. Add on $175-180 if the broader equipment cohort consolidates further.* ***Position size****: 1.5-2.5% of NLV — smaller than the large-cap equipment names because of the single-customer-concentration risk and small-cap volatility.* ***Stop****: daily close below $158 (the December 2025 base and the 200-day moving average).* ***Catalyst date****: August 11 Q2 print; SK Hynix Q2 print July 24 (read-through on HBM4 capex); September SEMICON Taiwan commentary; HBM5 hybrid-bonding inspection product disclosure expected late 2026.* ***Trim/exit triggers****: operating margin compression below 22% on a second quarter; any verifiable hybrid-bonding inspection loss at SK Hynix or Samsung; revenue guide cut below the H2 +25% step.* ***Conviction****: the cleanest pure-play HBM bump-inspection name and the best small-cap equipment exposure. Pair with ONTO at 3-4% combined for the bump-inspection duopoly trade.*

Layer 2 · ONTO — Onto Innovation

One-line thesis

Onto Innovation is the bump-inspection / metrology duopoly partner to Camtek, structurally cheaper on forward multiples, and the Rigaku 27% stake announcement plus the Q1 beat and the average sell-side PT of $352 (versus $271 spot) signals a 30% PT upside cushion that Camtek lacks at current levels.

What Onto Innovation physically does

Onto operates across three semiconductor-equipment categories: bump and incoming-wafer inspection (the Dragonfly platform inherited from the 2019 Rudolph Technologies merger), in-die optical and X-ray metrology (the Atlas family of overlay and film-thickness measurement tools, plus the Iris XRD line for film stress and composition), and lithography systems for advanced packaging (the JetStep step-and-flash imprint lithography line for fan-out wafer-level packaging). The portfolio is broader than Camtek's pure HBM bump-inspection focus, with material exposure to fan-out packaging (Apple A-series, AMD chiplets), advanced photoresist metrology at TSMC and Samsung, and the front-end memory-process metrology business at Micron and SK Hynix.

The Dragonfly inspection platform — originally developed by Rudolph Technologies and significantly expanded by Onto post-merger — performs the same physical task as Camtek's Eagle: optical inspection of µbumps, redistribution layers, and incoming wafer surfaces for HBM and advanced-packaging applications. The Atlas in-die metrology family is a different product category — these are precision overlay and film-thickness measurement tools that ship into leading-edge logic and DRAM fabs and sit upstream of the assembly process at the wafer-manufacturing tier. The JetStep step-and-flash imprint lithography platform is a niche but strategically important line — it is the dominant tool for fan-out wafer-level packaging redistribution-layer patterning at OSAT customers (ASE, Amkor, SPIL), which is the packaging platform used for Apple's M-series and A-series chips plus AMD's chiplet products.

For HBM specifically, Onto's Dragonfly platform competes head-on with Camtek's Eagle for bump-inspection share — with Onto more dominant at Micron and Camtek more dominant at SK Hynix as discussed above. The Atlas in-die metrology family is the broader product line that gives Onto exposure beyond HBM into the leading-edge logic and advanced-packaging metrology pull. The customer relationship at Micron has been particularly important — Onto's Boise / Idaho installed base is the principal reason the company's revenue mix has shifted toward memory-equipment exposure over the past three years.

The pending Rigaku partnership (announced May 7 with the 27% stake purchase for $710M) brings Onto a strategic relationship to the leading Japanese X-ray analytical instrument vendor, which adds capability in X-ray fluorescence (XRF) and X-ray diffraction (XRD) for advanced HBM packaging — an emerging metrology need as HBM stack heights move to 16-Hi and 20-Hi. The strategic logic is the same as Camtek developing toward hybrid-bonding metrology: positioning for the next-generation inspection paradigm where traditional optical methods become insufficient. The Rigaku capabilities specifically enable measurement of bonded interfaces, film composition at stacked die boundaries, and stress / strain analysis at the through-silicon-via structure — all of which become materially more important as the µbump shrinks and the inspection demand shifts to direct-bond integrity measurements.

The CPO (co-packaged optics) opportunity layer is the additional adjacency: as Broadcom Bailly, Cisco silicon photonics products, and the broader CPO category ramp through 2027-28, the panel-

level packaging inspection requirements for these products fall squarely in Onto's JetStep + Dragonfly product overlap. This is a single-digit revenue contributor currently but is structurally one of the highest-growth adjacency lines in the equipment landscape.

The financial print

Q1 2026 (reported May 7 2026) printed revenue of $291.95M against $289M consensus — modest beat (+1%) — and EPS of $1.42 against $1.37 consensus (+3.6%). Guidance for Q2 was raised to greater than $300M, implying 12-14% core organic growth. The Rigaku 27% stake at $710M was the strategic-narrative headline alongside the print, with the JV expected to close in H2 26 and a $500M bridge-loan added to the balance sheet for the transaction.

The Q1 revenue mix matters for tracking the underlying franchise health. Specialty device and advanced packaging — the HBM bump-inspection and packaging-related book — ran approximately $180-190M (roughly 62-65% of revenue), with the front-end advanced-nodes business at approximately $80-90M and lithography (JetStep) at $20-25M. The packaging book grew approximately 30% YoY while the front-end book was approximately flat — indicating the HBM-driven tailwind is doing the lifting, and the leading-edge logic metrology line is operating at trough utilisation but is positioned for an N2 / 18A ramp inflection through 2027.

The stock at $271 is 1m +15%, 3m +18%. The average sell-side PT of $352 — published as recently as May 13 (Simply Wall St aggregation) — implies 30% upside to spot, the widest cushion of any name in this Tier A list. Q2 prints August 6. Consensus is the guided $300M+ revenue range plus EPS around $1.50-1.60.

The valuation differential versus Camtek is the principal observation: ONTO trades at roughly 28x forward versus Camtek's 33x forward despite broadly comparable HBM exposure. The discount reflects the broader product portfolio (more cyclical drag from non-HBM segments) but is wider than the structural difference justifies given Onto's stronger Micron position. The cushion to consensus PT (30%) is also wider than Camtek (~10%), creating the asymmetric R/R that argues for owning both names paired but with the Onto position carrying the heavier weighting on entry valuation.

Customer mix

Micron is the largest customer at approximately 25-30% of revenue (HBM and DRAM front-end metrology, plus bump-inspection on the Idaho HBM line). TSMC at 15-20% (advanced-packaging metrology, photoresist film-thickness inspection at N3 and N2). SK Hynix at 10-15% (front-end DRAM metrology, with smaller bump-inspection share than Camtek). Samsung at 10-12% (memory and foundry). The remaining 25-30% is split across Apple's OSAT supply chain (fan-out wafer-level packaging at ASE / Amkor), Intel, and the broader specialty foundries.

The customer mix is materially more diversified than Camtek's HBM-concentrated mix, which is the structural reason Onto carries a discount on the bump-inspection franchise but a premium on the broader metrology mix. Net result is a more balanced revenue profile with less HBM-cycle beta.

Competitive context

KLAC is the competitor at the front-end metrology tier — Onto's Atlas competes against KLA's overlay and CD-SEM products at TSMC, Samsung, Intel — and KLA holds the larger share at front-end logic. Camtek is the direct competitor in bump-inspection as discussed above. The structural defence is the multi-product portfolio: where any single competitor wins on one product line, Onto has typically positioned an adjacent product line that wins on another step. The Rigaku partnership extends this strategy into X-ray metrology.

Terminal risk

Same as Camtek on the bump-inspection franchise — hybrid-bonding paradigm shift in HBM5 / HBM6 changes the inspection mix. The mitigant is broader: Onto's broader portfolio (in-die metrology, lithography for advanced packaging, the Rigaku X-ray adjacency) is more diversified and less dependent on the µbump generation than Camtek. The Rigaku JV specifically positions Onto for the next-generation inspection paradigm where X-ray methods become primary.

Secondary terminal risk: the Rigaku partnership integration risk — at $710M stake plus $500M bridge loan, the balance sheet is more levered than Onto's historical conservatism. If the JV underperforms on synergies, the equity multiple compresses.

Bull case

The base case is FY26 revenue of $1.18-1.22B (versus FY25 around $1.05B), FY27 revenue of $1.35-1.45B, gross margin holding 53-55%, and non-GAAP EPS rising from FY26 around $6.20 toward FY27 $7.20-7.80. At 32x forward — premium to current 28x, reflecting the Rigaku synergy and HBM compounding — that prints the stock at $230-250 in 2027, broadly flat to spot but with capital return underpinning.

The upside case is the Rigaku JV delivering on synergy ahead of schedule and the HBM bump-inspection share at Micron expanding into the Idaho IPF2 ramp through 2027 — combined that lifts FY27 EPS toward $9.00 and the stock at $310-340 in 2027, +14-25% from spot, consistent with the sell-side $352 average PT.

The CPO panel-level packaging inspection adjacency is the upside-case leg most under-appreciated by current sell-side modelling. As Broadcom Bailly 51.2T CPO products ramp in 2027, as Cisco silicon photonics deployments scale at hyperscaler data centers, and as the broader CPO category transitions from pilot to production volume through 2028, the panel-level packaging metrology and inspection requirements scale proportionally. Onto's JetStep imprint lithography platform plus the Dragonfly inspection platform are positioned to capture roughly 60-70% of this addressable inspection TAM. Even at the lower end of CPO adoption assumptions, this is a $200-400M annual revenue line by 2028 — material at Onto's $1.2B revenue base.

The Rigaku synergy case is the second under-appreciated leg. The X-ray metrology capability that Rigaku brings is uniquely positioned for the HBM stacked-die inspection market, where traditional optical inspection methods reach physical limits as bond pitches shrink below 10 microns. If the JV delivers commercial integration ahead of schedule and Onto-Rigaku captures the X-ray-based stack-integrity inspection volume that the HBM4 16-Hi generation requires, the synergy line accretes meaningfully through 2027.

Gap / bear case

Three concerns. First, the Rigaku integration. The $710M stake + $500M bridge loan adds meaningful balance-sheet complexity to a company that has historically run net cash; if the synergies disappoint, the equity multiple compresses 100-200bp. Second, the front-end metrology business carries cyclical exposure to broader logic and DRAM capex — if 2027 sees a memory cycle softening, Onto's revenue takes a haircut alongside KLAC and LRCX. Third, the +18% three-month rally has compressed the entry-level discount versus consensus PT — but the cushion is still wider than peers, which is the principal R/R argument.

Optionality

Three options. Rigaku JV synergy realization through H2 26 / 2027 — every dollar of incremental X-ray metrology revenue at the HBM packaging tier accretes directly to Onto. The CPO (co-packaged optics) panel-level packaging inspection market — emerging in 2027 as Broadcom Bailly and Cisco silicon photonics products ramp, this is a $500M+ inspection TAM addressed by Onto's JetStep + Dragonfly portfolio. And the more speculative option: M&A from a larger equipment vendor (AMAT, LRCX, KLAC) consolidating the bump-inspection duopoly — neither base case nor near-term, but the strategic logic exists.

The trade

ONTO — BUY 8/10. Entry*: $270-275 spot; better entry on $250-260 pullback.* ***Position size****: 1.5-2.5% of NLV, paired with CAMT at 3-4% combined for the bump-inspection duopoly trade.* ***Stop****: daily close below $215 (the December 2025 base and prior breakdown level).* ***Catalyst date****: August 6 Q2 print; Rigaku JV close target H2 26; Micron Q3 print June 25 (read-through on Boise HBM capex); SEMICON West July 2026 commentary.* ***Trim/exit triggers****: Rigaku JV synergy timeline slipping into 2027; front-end metrology revenue softening below 5% YoY; any verifiable share loss at Micron bump-inspection to Camtek.* ***Conviction****: the cheaper of the bump-inspection duopoly with a Rigaku option leg on top. Treat as the value half of the pair-trade with CAMT — ONTO has more upside on PT cushion, CAMT has cleaner HBM-pure-play exposure.*

Layer 3 · 000660 — SK Hynix

One-line thesis

SK Hynix is the cleanest HBM monopoly trade in the global equity market — Q1 26 printed revenue +198% YoY and operating margin at 72% with management explicitly guiding "HBM demand for the next 3 years far exceeds capacity," and the NVDA Rubin design win plus the HBM4E samples in H2 26 makes this the highest-conviction Layer-3 memory position in the cycle.

What SK Hynix physically does

SK Hynix is the global leader in High Bandwidth Memory (HBM) — the stacked DRAM that sits adjacent to every AI GPU and is the single binding constraint on every AI training cluster and inference fleet. The Icheon M14 and M16 fabs in South Korea, plus the Cheongju M15 fab, plus the dedicated HBM advanced-packaging lines that surround them, collectively produce roughly

50-55% of the world's HBM volume in 2026 — and the share at the leading HBM3e and HBM4 generations is closer to 60%. SK Hynix is the sole HBM3e supplier to NVIDIA Blackwell and Blackwell Ultra; the company is the lead-customer launch partner on HBM4 for NVIDIA Rubin shipping H2 26 into 2027.

The physical process at the HBM stacking step: 12-Hi or 16-Hi stacks of DRAM dies (each die ~30μm thick after TSV grind) bonded vertically with MR-MUF (mass reflow molded underfill) or TC-NCF (thermal-compression non-conductive film) using Hanmi or ASMPT TC bonders, with thousands of through-silicon vias per die etched by Lam Sense.i Plus tools, inspected by Camtek Eagle / Onto Dragonfly tools, with film metrology done by KLAC, ONTO, and Hitachi tools. The HBM die itself is a 1z or 1a-process DRAM at roughly 10-15nm half-pitch; the next-generation HBM4 base die moves to a 5nm logic process (produced at TSMC for SK Hynix specifically) for additional buffer logic and I/O scaling.

The HBM4 generation marks a structural shift in the supplier-customer relationship. The base die — historically a low-margin DRAM-process buffer die produced internally — moves to a 5nm logic process produced at TSMC, with custom logic IP designed by SK Hynix and integrated I/O for customer-specific interface variants. NVIDIA Rubin gets an SK Hynix-specific HBM4 SKU with optimised base-die logic; AMD and Broadcom get different SKUs. This customisation increases HBM ASPs by roughly 25-35% versus HBM3e and structurally raises the gross margin on the HBM line — a key reason the FY26 operating margin printed at 72% and the company can sustain near that level through the cycle.

Outside HBM, SK Hynix is the second-largest DRAM maker globally (after Samsung) and a meaningful NAND player through the Solidigm subsidiary acquired from Intel. The DRAM standard product line (DDR5 server, LPDDR5x mobile) is the cycle-driven cash flow ballast; HBM is the structural growth and margin driver. The Solidigm acquisition — completed in 2021 for $9B — has gone from a problematic integration through 2022-23 into a contributor as AI-inference NAND demand has materialised in 2025-26. The data-center NAND pricing surge captured by SanDisk Q3 FY26 (+130% QoQ NAND ASP) and Kioxia FY26 (record operating profit) is also flowing through Solidigm at smaller scale.

The financial print

Q1 2026 (reported April 23 2026) printed revenue of KRW 52.6 trillion — +198% YoY, +60% QoQ — operating profit of KRW 37.6 trillion (operating margin 72%, +405% YoY), and net income of KRW 40.3 trillion (net margin 77%). These are the strongest semiconductor margins ever recorded by a memory maker in cycle history — and they are not a one-quarter anomaly. Management guided that HBM demand for the next three years "far exceeds capacity," and that HBM4 mass production is in qualification with HBM4E samples shipping in H2 26 ahead of mass production targeted for 2027.

The Q1 print also confirmed that 2026 HBM, DRAM, and NAND capacity is "essentially sold out" — a statement that has been used by SK Hynix's IR team in every quarterly call going back to Q3 2024 and that the market has now priced into the 12-month forward equity revisions. The stock at KRW 1,820,000 sits roughly 9% below the all-time high of KRW 1,995,000 set May 15. 1m +35%, 3m +77% — the strongest 90-day rally of any name in this Tier A list, and the principal reason the position-sizing discussion below is more nuanced than the conviction call.

Q2 2026 prints July 24. The consensus revenue is approximately KRW 50-53 trillion with operat-

ing profit around KRW 35-38 trillion. The setup is constructive but the bar is now extraordinarily high — any sequential softening in HBM3e ASPs or any HBM4 qualification slippage at NVIDIA Rubin can produce a multi-week sell-off even into a beat-and-raise print.

Customer mix

NVIDIA is the single largest end-customer for HBM and represents approximately 60-65% of SK Hynix HBM revenue through 2026 — sole-supplier on HBM3e for Blackwell, lead-customer launch partner on HBM4 for Rubin. AMD is the second-largest at 15-20% (MI355X and the rumoured MI400). Google (TPU v6 / v7) and Broadcom-designed custom silicon collectively account for another 10-15% of HBM through 2026. The customer concentration is real and the principal financial risk — NVIDIA's HBM share-of-wallet is the single largest determinant of SK Hynix's quarterly margin.

The contracting structure has shifted materially in this cycle. Historical DRAM commercial structure was quarterly spot pricing with limited forward visibility. The 2024-25 HBM cycle has moved to long-term-agreement (LTA) pricing — multi-quarter or multi-year fixed-price commitments at the wafer or chip-stack level, principally between SK Hynix and NVIDIA/AMD. SK Hynix has publicly disclosed that 2026 HBM capacity is sold out under LTA commitments and that 2027 HBM capacity is "essentially committed." This is a structural shift away from cyclical pricing and is the principal reason consensus equity models attribute lower cycle-beta to SK Hynix than would be implied by historical memory-cycle analogues.

Outside HBM, the DRAM standard-product customer base is the broader cloud / data-center / handset cohort, while NAND through Solidigm goes to the data-center storage market (a meaningful tailwind given the AI-driven storage demand discussed under WDC / SNDK / Kioxia in Part 2 of this Tier A series). The DRAM standard-product mix is approximately 60-65% server / data-center (DDR5 RDIMM, MRDIMM) and 30-35% mobile (LPDDR5x for premium smartphones), with the balance in client PC and graphics DRAM. The server mix tilt is the principal exposure to the hyperscaler capex cycle — which is the broader AI thesis but adds cyclicality.

Competitive context

Samsung Memory and Micron are the two direct competitors. Samsung lost the HBM3 / HBM3e qualification cycle at NVIDIA in 2024-25 and is now publicly committed to closing the gap on HBM4 — the PyeongTaek HBM4 line is in qualification with NVIDIA targeting H2 26, but the trade press through Q2 26 (Korea Economic Daily, Trendforce, Korea Herald) consistently suggests Samsung is still 1-2 quarters behind SK Hynix on HBM4 yield. Micron has fully qualified HBM3e at NVIDIA Blackwell (a 2025 win) and is ramping the Boise / Idaho IPF1 / IPF2 capacity through 2027 — Micron's share of NVIDIA HBM has expanded from <5% in 2024 to roughly 20-25% in 2026, and is forecast to expand further on the Idaho ramp.

The structural moat for SK Hynix is the multi-decade HBM-specific R&D base (SK Hynix invented commercial HBM in 2013 with AMD), the customer-co-development lock-in at NVIDIA, the advanced-packaging-line capacity that surrounds the M14/M16 fabs, and the cost-position advantage at the Korean operating base. The moat is real but not absolute — Samsung is closing the HBM4 gap, Micron is taking marginal share, and the duopoly+ economic structure is more competitive than the SK-Hynix-monopoly narrative often presented in sell-side material.

Terminal risk

Two terminal risks. First, the cycle: HBM is still a memory product, and memory products cycle. The current cycle has been extraordinary in both peak ASPs and capacity tightness, but every prior memory super-cycle (1995, 1999, 2007, 2017, 2021) ended with a sharp ASP correction and a 12-18 month equity drawdown of 40-60%. The bull case requires that the 2026-27 HBM cycle is structurally different — driven by long-term-agreement (LTA) pricing visibility through 2027 and the binding-constraint nature of HBM supply — but a Samsung HBM4 qualification breakthrough or a Micron capacity ramp surprise can compress HBM ASPs faster than the consensus 2027 EPS model anticipates.

Second, the AI-capex tail: if hyperscaler AI capex moderates in 2027-28 (the bubble warning recycled every six months), HBM unit volumes compress as the marginal NVIDIA Rubin / AMD MI400 build slows. SK Hynix's equity exposure to AI capex is the highest in the memory cohort.

Bull case

The base case is FY26 revenue of KRW 195-210 trillion (versus FY25 KRW ~85T), operating profit at KRW 140-155 trillion, operating margin holding above 70%, and FY26 net income at KRW 130-145 trillion (~$95-105B equivalent). At 13-15x forward earnings — historically generous for a memory maker but appropriate for the HBM-monopoly tilt — that prints the stock at KRW 2,200,000-2,400,000 in 2027, +21-32% from spot.

The upside case is HBM4 qualification at NVIDIA Rubin landing clean H2 26 and Samsung HBM4 qualification slipping into 2027, which preserves SK Hynix's HBM monopoly economics for a full additional generation. Combined with NAND pricing surge through Solidigm (read-through from SanDisk Q3 NAND +130% QoQ, see Part 2) and the HBM4E samples landing on schedule, FY27 net income prints toward KRW 170-180 trillion. At 15x forward, the stock prints KRW 2,800,000-3,000,000 in late 2027, +54-65% from spot.

Gap / bear case

Three concerns. First, the +77% three-month rally has compressed the entry-level R/R — the stock has rerated to peak HBM-monopoly multiple and a Samsung qualification breakthrough produces an immediate 15-25% pullback. Second, the operating margin at 72% is mathematically near-impossible to maintain through a full cycle — even at peak DRAM 2017 the operating margin topped out at ~50%, and the consensus FY27 model assumes margin moderation to 55-60% which is already a meaningful step down. Third, the equity is approaching a market cap of KRW 1,300 trillion (roughly $950B equivalent), which puts SK Hynix in the global top-25 by market cap — at that scale the institutional bid is broader but the marginal flow is harder to sustain.

Optionality

Four options. The NVIDIA Rubin design-win on HBM4 — currently the base case — confirmed publicly during H2 26 removes the principal qualification-risk overhang and triggers the next equity re-rating leg. HBM4E mass production targeting 2027 with a one-generation lead on Samsung extends the monopoly economics for another year. The NAND optionality through Solidigm: if NAND ASPs continue to surge on AI-inference storage demand (the SanDisk / Kioxia / WDC theme), Solidigm contributes incremental margin that is currently modelled as roughly break-even.

The fourth option is hybrid-bonding leadership through the HBM5 / HBM6 transition. SK Hynix has been the most aggressive HBM maker in committing to hybrid-bonding integration (with Hanwha Semitech as the local equipment partner and BESI as the global partner). If SK Hynix delivers production hybrid-bonded HBM5 ahead of Samsung and Micron — which is the consensus path on 2027-28 timelines — the monopoly economics extend through another full generation. This is the long-tail terminal-value preservation that distinguishes SK Hynix from its competitors and that is not fully reflected in the current FY27-28 consensus.

The trade

000660 (SK Hynix) — BUY 9/10. Entry*: trim into spot at KRW 1,820,000; full size on pull-back to KRW 1,600,000-1,700,000 if the cohort consolidates. The +77% three-month rally has compressed the entry — do not chase here.* ***Position size****: 4-5% of NLV — paired with Micron in Part 2 as the HBM bilateral exposure. The Korean equity carries a 10-15% structural geopolitical / governance discount versus a US-listed analogue (this is the reason MU is sized comparably even at smaller fundamental margins).* ***Stop****: daily close below KRW 1,350,000 (the early-March base and the 200-day moving average).* ***Catalyst date****: July 24 Q2 print; NVIDIA Q1 FY27 print May 20 (HBM commentary read-through); September SEMICON Korea (HBM4 yield disclosure); Samsung Q2 earnings late July (HBM4 qualification status).* ***Trim/exit triggers****: operating margin guide below 65%; any verifiable Samsung HBM4 qualification breakthrough at NVIDIA; FY27 HBM ASPs softening on Trendforce monthly data.* ***Conviction****: the cleanest HBM monopoly trade in the cycle, with the strongest single-quarter print of any name in this Tier A list. The reason this is 9/10 not 10/10 is the cycle-peak margin question — HBM is the best memory product ever invented, but it is still memory, and memory cycles. Size at 4-5% with full awareness of the +77% three-month context.*

Layer 3 · MU — Micron Technology

One-line thesis

Best US-listed HBM pure-play running at 3x prior-year revenue with HBM sold out through 2027 — own it for the operating-leverage print into a memory super-cycle that consensus is still 12 months behind on.

What Micron physically does

Micron is one of three companies on the planet that can manufacture leading-edge DRAM at scale — alongside SK Hynix and Samsung — and one of the four that still matter in NAND. Inside the AI stack the relevant business is High Bandwidth Memory: 3D stacks of 8 to 16 DRAM dies bonded vertically onto a logic base die, threaded with thousands of through-silicon vias that pipe data into a GPU at multi-TB/s bandwidths. Every Blackwell, every B300, every MI355X, every Trainium chip ships with HBM bolted onto its package. You cannot ship more AI silicon than you can ship HBM, and Micron is the only American name in that triumvirate.

Physically, Micron fabricates the DRAM die in Boise, Hiroshima, Taichung and now Idaho — runs the wafers through TSV etch, thinning and stacking — and then ships finished HBM3E 8-Hi / 12-Hi modules to Nvidia, AMD, Broadcom and the hyperscalers' captive ASIC programs. Micron sits at

Layer 3 of the AI stack — the memory layer — where the binding constraint of the entire compute economy currently lives. The Idaho IPF1/IPF2 greenfield expansion, supported by US CHIPS Act funding and several billion of customer prepayments, is the multi-year capacity vehicle. The Taichung HBM2E/3 expansion is the near-term volume engine. The Hiroshima 1-gamma node ramp is the technology lever for HBM4.

The thing to internalise about Micron is that it is not a memory cyclical anymore — at least not in the way the last twenty years taught investors to model it. HBM is sold under multi-year long-term agreements with fixed pricing into 2027. The commodity DRAM line remains cyclical, but the HBM line is structurally contracted, and HBM is now the majority of incremental revenue and almost all of incremental gross profit. The structural break from the historical Micron pattern — the four-year boom-bust DRAM cycle that returned 40% to shareholders one year and -30% the next — is that HBM contracts are now written more like aerospace supplier agreements than commodity-memory POs: take-or-pay, fixed-price escalators, multi-year volume commitments with non-cancellable deposits.

The HBM3E ramp through 2024-25 and the HBM4 production qualification cycle through 2026 are running ahead of consensus modelling. Micron's HBM3E 12-Hi is qualified at Nvidia Blackwell, and the HBM4 8-Hi/12-Hi samples have been delivered to lead customers for production qualification ahead of the Rubin platform launch. The technology roadmap to HBM4E and HBM5 is intact, with the 1-gamma node providing the bit-density step required to keep pace with Hynix on per-stack capacity. The Idaho fab is the geographic insurance against any single-site disruption risk, which is itself a hyperscaler procurement priority.

The financial print

Q2 FY26 (March quarter, reported March 20) printed **$23.86 billion of revenue — roughly three times the prior-year quarter** — and beat consensus on every line. Q3 guidance came at ~$33.5 billion, which is itself above the high end of Street consensus and was raised again in mid-cycle commentary. Gross margins are running in the high-50s%, operating margins comfortably above 40%. FY26 EPS consensus has now reset to **$58 against $8.29 in FY25** — a seven-fold step-up in earnings power that the sell-side modelled to $20 just six months ago and which the buy-side is still adjusting position size against.

The stock closed **$803.63 on May 14** and is up roughly **+180% year-to-date**. The Motley Fool $2,000 twelve-month target circulated May 11 is not as outlandish as the headline reads — it bakes in the consensus HBM TAM expansion from $35 billion in 2025 to roughly $100 billion in 2028, applies a normalised peer multiple, and works backward. The base case at consensus FY27 EPS of ~$70-80 and a 25x multiple lands you at $1,750-2,000 even without the optionality. Q3 FY26 reports **June 25** — the next discrete catalyst.

The earnings-power math deserves an additional dimension. Micron's revenue per wafer-out on HBM is roughly 4-5x the revenue per wafer-out on commodity DRAM, and the gross profit per wafer is 7-9x higher because the HBM contribution margin is structurally higher. As the HBM mix moves from ~30% of revenue in FY26 to a consensus ~50% in FY27 and a stretch case 60%+ in FY28, the consolidated gross-margin trajectory is mechanically upward irrespective of commodity DRAM pricing. That is the lever that turns Micron from a cyclical $8-25 EPS business through the historical cycle into a $50-100 EPS business through the AI cycle.

Customer mix

Hyperscaler concentration is high and tightening rather than diversifying. Nvidia is the largest single buyer of Micron HBM3E by a wide margin — the B200/B300/Rubin pipeline runs on Micron and SK Hynix as the qualified vendors, with Samsung still working through HBM3E qualification — and that single customer is probably 25-35% of incremental HBM revenue in FY26. AMD's MI355X and MI400 pulls represent the next-largest tranche. AWS Trainium, Microsoft Maia and Google's next-generation TPU programs each take meaningful HBM volume directly or through their custom-silicon partners.

The customer book is heavily contracted into 2027 via LTAs that lock both price and volume. Management commentary from the March print explicitly said Micron can only meet **50-66% of medium-term demand** — translation: the order book is bigger than the wafer plan, customers are paying upfront to secure allocation, and pricing is being negotiated at fixed escalators rather than spot. That is the kind of customer mix that produces operating margin expansion through a cycle, not compression at the peak.

Competitive context

The competitive set is two names: SK Hynix and Samsung. SK Hynix is the share leader in HBM3E and the first to qualify HBM4 with Nvidia — the structural #1 in the memory bottleneck. Micron is the credible #2 with the cleanest US-listed exposure, the fastest HBM3E qualification cycle Nvidia has ever run, and the cleanest balance-sheet improvement. Samsung is the laggard — HBM3E Nvidia qualification has been pending for over eighteen months, foundry losses persist, and the captive-fab structure makes it the slow mover.

Micron's moat is three-pronged. First, the technology gap on 1-beta/1-gamma DRAM nodes is narrow but real — Micron is roughly one node behind Hynix, ahead of Samsung. Second, the customer relationships are sticky because HBM qualification cycles are 12-18 months, which means a customer who has qualified Micron does not lightly switch back. Third, the US-listed wrapper plus CHIPS Act geography gives Micron a structural advantage for any hyperscaler that wants supply-chain diversification away from Korean concentration risk. The competitive context is a duopoly-plus-laggard rather than a true three-way race, and Micron sits inside the duopoly.

The pricing-power signature inside the competitive set is captured in the recent contract structure. Through 2023-24, HBM pricing was negotiated quarterly and pricing power belonged primarily to the customer. Through 2025, HBM pricing migrated to annual contracts. Through 2026, HBM is being sold on multi-year LTAs with non-cancellable structure and pricing escalators that favour the supplier. Micron has captured this transition fully — the visibility into 2027 is unprecedented for a DRAM business — and the marginal pricing leverage now sits with Micron, Hynix and Samsung rather than with the GPU vendors. That is a structural shift in the memory food chain that is not yet reflected in consensus through-cycle margin assumptions.

Terminal risk

The terminal risk is the same risk that has killed every memory bull case since 1995: the cycle peaks, new capacity ships, and pricing breaks. If Samsung's HBM4 qualification at Nvidia clears in H2 2026 and Samsung adds 50% to its HBM capacity as guided, the marginal HBM bit ships into a softer pricing environment in 2027. The bear scenario is FY27 revenue holding flat while

gross margins compress 800-1000 bps, EPS halves, and the stock multiple compresses against a lower forward number — a classic memory cycle peak.

Bull case

The three-to-five-year bull case rests on the HBM TAM compounding from $35 billion in 2025 to $100 billion in 2028 and Micron holding 20-25% share — call it $20-25 billion of HBM revenue alone at peak, against $30 billion of HBM-eligible capacity supply and $80-100 billion of total Micron revenue. Operating margins in the high-30s structurally rather than the high-teens through-cycle averages of the last decade. FY28 EPS in the $90-110 range. A 20-25x multiple lands the stock at $1,800-2,750.

The bull case also includes the optionality on HBM4E and HBM5 generations — where the 1-gamma node and Micron's R&D pipeline arguably narrow the gap to Hynix — and on the HBF (High Bandwidth Flash) joint roadmap with SanDisk and Kioxia, where AI inference storage becomes a structurally tight market in the late-2020s. Each of these is a free option at current prices.

The cash-generation arithmetic is what makes this an institutional rather than a retail trade. At consensus FY27 EBITDA of $35-40 billion and capex running $15-17 billion through the Idaho ramp, free cash flow lands in the $18-25 billion range — enough to retire most of the legacy debt, run a meaningful buyback, and initiate a sustainable dividend. The transition from "memory cyclical with reinvestment needs" to "AI-memory franchise with capital-return capacity" is the re-rating mechanism that pulls multiple from the historic 8-12x forward through-cycle range to the 20-25x range that durable secular growers earn. Most sell-side models still anchor on the historical multiple band, which is the consensus error.

Gap / bear case

The market is right about the direction but probably still wrong about the magnitude. Consensus FY26 EPS at $58 still leaves room — sell-side has been chasing the number quarterly and is now structurally behind the LTA disclosures. But the gap the market may also be missing is on the downside: if 2027 is the cycle peak rather than the mid-cycle, and the LTAs reset lower from 2028 onwards, the stock that printed +180% YTD into the print could give back 40-50% in a six-month period. The bear case is not absent — it is delayed by 12-18 months. The asymmetry today still favours long; the asymmetry in mid-2027 may not.

The specific signal to watch for the cycle-peak inflection is HBM ASP movement in the monthly Trendforce and DRAMeXchange data, combined with hyperscaler capex guidance at the Q1 2027 prints. If hyperscaler capex guides flat or down for 2028 against 2027 levels, the LTA structure will hold short-term but the renewal terms in 2028-29 will reset materially lower. The historical memory-cycle peak pattern has been six to nine months of warning signs before the stock breaks; watching the ASP data carefully is the protective discipline.

A more granular signal worth watching is Samsung's HBM3E and HBM4 qualification track at Nvidia. If Samsung clears qualification in H2 2026 and ramps aggressively in 2027, Micron's HBM share at Nvidia could give back 200-400 basis points to Samsung, partially offsetting the volume growth. The mitigation is that the HBM TAM is growing fast enough that Micron's absolute HBM revenue continues to expand even at a lower share; but the share-loss optics would

compress the multiple. Watching Samsung's qualification disclosures at quarterly print cycles is part of the trade discipline.

Optionality

Three free options worth tagging. First, HBF — High Bandwidth Flash — is a NAND-based AI inference tier that Micron is positioned alongside SanDisk and Kioxia to develop; this becomes a real TAM line if AI inference economics shift from training-dominated to inference-dominated, which is the current trajectory. Second, on-package memory for CPUs is an emerging architecture that Micron's LPDDR5X and 3DS DRAM products are uniquely positioned for as the GB-class CPU+GPU hybrid socket gains share. Third, the Idaho fab geography is itself a strategic option for US-DoD and US-cloud sovereign workloads — none of which is currently priced.

The trade

> **Entry:** $760-820 — current zone is acceptable; on a 10-15% pullback to $700 is a clean add. **Size:** 4-5% portfolio weight target, scale on prints. **Stop:** $620 (close below 200-day MA invalidates). **Catalyst date:** Q3 FY26 print June 25, 2026. **Trim/exit:** trim 25% above $1,200; exit at $1,500 or on HBM ASP roll-over in 2027 monthly data. **Conviction: 9/10**.

Layer 3 · SNDK — SanDisk Corporation

One-line thesis

SanDisk is the cheapest NAND pure-play that just printed 78% gross margin on +250% YoY revenue — own it as the AI-storage-tightness call and accept that the parabola will eventually correct.

What SanDisk physically does

SanDisk fabricates 3D NAND flash memory — the storage substrate that sits one tier below DRAM in the memory hierarchy and absorbs the cost-sensitive bulk of AI training datasets, model checkpoints, agentic-AI memory, vector databases, and the enterprise SSDs that increasingly displace HDDs at the warm-tier of hyperscaler storage. Post-separation from Western Digital in early 2025, SanDisk is now the standalone NAND pure-play, jointly operating the Yokkaichi and Kitakami fabs with Kioxia under their decades-long manufacturing JV. Physically, SanDisk's wafers are produced at 100+ layer 3D NAND stacks, packaged into enterprise SSDs, client SSDs, and the new HBF (High Bandwidth Flash) form factor that the Kioxia-SanDisk roadmap is positioning as the AI-inference storage tier.

The AI stack relevance is that storage tightness has finally arrived at NAND. For most of the past four years AI capex was bottlenecked at GPUs, then at HBM, then at advanced-packaging substrates. NAND was the spare-capacity layer — over-supplied, low margin, the loss-leader of the memory complex. That has flipped. AI inference workloads at scale require enormous volumes of warm-tier storage, agentic AI workflows require persistent memory for context and tool-state, and the migration of vector databases onto SSD-class storage has pulled enterprise

NAND demand to levels the industry never planned capex for. SanDisk and Kioxia, both running the same Yokkaichi/Kitakami fabs, are the principal beneficiaries.

The HBF roadmap is the AI optionality. HBF is a NAND-based, HBM-form-factor flash module designed to sit on the same package as the GPU and serve as a 1-2 tier slower but 10-100x cheaper-per-bit substitute for HBM-only architectures in inference workloads. If HBF clears the qualification cycle at Nvidia or AMD for any production inference SKU, it is a step-change re-rating event for the joint SanDisk-Kioxia franchise.

The mechanics of why NAND is suddenly tight in 2026 are worth spelling out because most generalist investors still think of NAND as the perennial-oversupply layer of the memory complex. Three things changed simultaneously. First, the industry under-invested in capex from 2022 through early 2025 — Samsung, Hynix, Kioxia, Micron and SanDisk all cut node-transition capex during the 2022-23 downcycle, which means 2026 supply is structurally below trend. Second, AI inference workloads create a demand pull on enterprise SSD volumes that nobody modelled — every agentic AI deployment, every RAG system, every long-context inference cluster needs persistent storage at multi-petabyte scale. Third, the substitution of HDD warm-tier with SSD at the hyperscaler tier is no longer a 2027-28 story but a 2026 story, because the GB-per-watt economics finally crossed over in the high-density nearline use case. All three pull bias up on the demand curve at the same time supply curve under-shoots. That is the NAND supercycle.

The financial print

Q3 FY26 (March quarter, reported April 30) was a blowout — **revenue $5.95 billion, +250% YoY**, EPS **$23.41 versus $14.62 consensus (+60% surprise)**, non-GAAP gross margin **78.4% — a 5,570-basis-point year-over-year expansion**. These are not normal memory numbers. The 78% gross margin print is what an AI accelerator looks like at peak pricing, not what a NAND fab has historically delivered.

The mechanism is pricing leverage. SanDisk disclosed NAND pricing **+130% QoQ and +200% YoY** — the steepest unit-price increase in NAND's modern history. Capacity is sold out, customer LTAs are being renegotiated higher into 2027, and the cost base is structurally lower than the FY24 trough because the fab JV has been running depreciation-heavy. The stock has closed at **$1,381 on May 17 — up 482% year-to-date** — and prints all-time highs as the analyst community walks consensus higher. Q4 FY26 reports **July 30**.

Customer mix

Customer concentration is hyperscaler-skewed but more diversified than the HBM names. The top five customers — AWS, Microsoft, Google, Meta and one tier-1 OEM — likely account for 50-60% of revenue, with the enterprise SSD line carrying the bulk. Nvidia is an indirect customer through the AI-server OEM channel; the direct relationship is still developing on the HBF roadmap.

The product-mix evolution inside the customer book is what supports the through-cycle margin step-up. Through 2023-24, SanDisk's revenue mix was roughly 60% client/consumer NAND (mobile, PC, USB sticks, SD cards) and 40% enterprise NAND. Through 2026, that mix has flipped to roughly 35% client / 65% enterprise as the hyperscaler SSD demand has compounded. Enterprise NAND carries 1,500-2,500 basis points higher gross margin than client NAND structurally, so the mix-shift alone explains a meaningful slice of the 78% gross margin print without invoking

commodity-pricing peaks. The mix is durable — enterprise SSD demand does not mean-revert the way client NAND does — and that supports the bull-case argument that through-cycle margins are now structurally higher than the historic peer comp.

The customer LTAs are the key signal. SanDisk has explicitly disclosed that **NAND pricing is locked into 2027 at escalators above current spot**, that capacity for the next 12-18 months is contracted, and that LTAs into the back half of the decade are being negotiated. This is a meaningful change from the historic NAND contract structure, which was effectively spot pricing on a quarterly basis. The customer mix is also tightening upward — hyperscalers are now willing to pay premium for guaranteed allocation rather than chase spot, which is itself the indicator that the cycle is structurally different.

Competitive context

The NAND market has consolidated to a five-vendor structure: Samsung, SK Hynix/Solidigm, Kioxia, SanDisk and Micron — and effectively two technology consortiums (Samsung-Hynix and Kioxia-SanDisk via the JV). SanDisk is the western-listed wrapper for the Kioxia-SanDisk fab capacity, which represents roughly 30-35% of global NAND supply. The moat is the JV manufacturing scale, the existing customer qualifications, and the technology roadmap (BiCS 9 and BiCS X high-layer 3D NAND nodes).

The capex-discipline dynamic across the NAND vendor group is what makes the current pricing power sustainable rather than transient. Through 2022-23, Samsung dominated capex expansion and effectively over-supplied the market, which is why the 2022-23 NAND downcycle was so severe. Through 2024-26, Samsung has been more disciplined, Hynix has been focused on HBM rather than NAND, and the Kioxia-SanDisk JV has prioritised cash returns over capacity expansion. The result is industry capex running below the demand growth curve, which is the structural condition for sustained pricing power. Watching the Samsung and Hynix NAND-specific capex commentary at each print cycle is the key forward indicator.

SanDisk's competitive position has improved markedly since the separation from Western Digital. As a standalone NAND pure-play with a tighter cost structure and direct hyperscaler relationships, the operating leverage profile is now clean — no HDD-cycle drag, no portfolio-discount overhang. Against Micron's NAND segment and Samsung's NAND business, SanDisk is the marginal-cost beneficiary because the Yokkaichi/Kitakami JV depreciation is largely behind it, while Samsung is still running large new-fab capex. The competitive context is favourable through the cycle peak and probably for another 18 months after.

The market-share data is informative. SanDisk's standalone NAND share is roughly 14-16% globally; combined with Kioxia inside the JV the manufacturing footprint controls 30-35% of global capacity. Samsung remains the share leader at ~32%, SK Hynix/Solidigm runs ~17-19% combined, Micron ~10-12%. The competitive battleground in the up-cycle is principally the enterprise SSD segment, where the SanDisk-Kioxia JV is gaining share on the back of BiCS-class technology that outperforms Samsung's V-NAND on per-watt at the highest capacities. The enterprise SSD share gain is what is fuelling SanDisk's revenue out-performance versus the sector — the company is taking share inside a tight market, which compounds the pricing-power lever.

Terminal risk

The terminal risk is the parabolic move itself. **+482% in five months** is a sign that the marginal buyer is paying any price for the narrative, and when AI capex visibility blinks — whether for one quarter or one cycle — the unwind is brutal. Historic NAND cycle peaks have given back 50-70% of the upcycle move within 12 months. The risk to the thesis is not that the cycle is wrong; it is that 482% has already priced the cycle. A capex blink at any major hyperscaler in the 2027 print cycle is the trigger.

A secondary technical risk is the Chinese NAND capacity overhang. YMTC's domestic Chinese fab capacity is ramping; while the absolute volume is still small in global context, the trajectory through 2027-28 could meaningfully add bits at the low end of the market, particularly for consumer and client NAND. The displacement risk is at the volume tier rather than the enterprise tier where SanDisk is concentrating, but any pricing pressure on commodity NAND ripples through the wider market and could compress SanDisk's blended ASP.

Bull case

The base bull case is structurally tight NAND through 2027, HBF qualification at one hyperscaler-class customer by end-2026, and FY27 revenue running $22-25 billion versus $18-19 billion FY26 consensus. EPS in the $90-110 range at peak, a 12-15x multiple — a NAND business with HBF optionality and a tight supply curve probably deserves a higher multiple than the historic peer math suggests — and you arrive at $1,400-1,700 as a base bull range.

The stretch bull case adds HBF being a real GPU-package memory tier by 2028, which materially re-rates the entire Kioxia-SanDisk franchise on a sum-of-parts and pulls the multiple to high-teens. The implied upside in that path is $2,500+ over three years, but it is a probabilistic call rather than a base case.

The capital-return scenario is the underappreciated leg of the bull case. SanDisk's FY26 free cash flow is running at $4-5 billion against a market cap of roughly $85-90 billion at $1,381 — a free-cash-flow yield in the high single digits at peak earnings. The board has signalled openness to buybacks once leverage targets are met (the post-separation balance sheet was deliberately structured for capital flexibility), and a sustained buyback at $1,200-1,500 would compound earnings-per-share growth through the cycle peak in a way that pure-volume models do not capture.

Gap / bear case

What the market may be missing is the duration risk on the upside — if NAND ASPs hold or push higher into 2027 because hyperscalers continue to pay for allocation rather than wait for capex to land, the FY27 number is dramatically above consensus and the stock can grind higher despite the 482% YTD move. Conversely, the bear case the market may be under-pricing is operating-margin mean reversion: 78% non-GAAP gross margin is a peak print, not a through-cycle number, and a 1,000-bp compression in FY27 takes the EPS down 25-30% even on flat revenue.

The market is also probably under-pricing the JV-structure risk. Kioxia and SanDisk co-own the fabs; any disagreement on capex pacing, technology roadmap or HBF prioritisation between the two parents creates execution friction. The post-IPO Kioxia has more strategic flexibility now, which is a positive for HBF velocity but a complication for joint-capacity planning.

The capex-cycle timing is also worth highlighting. Through 2026, the JV is running fabs near full utilisation with limited capex expansion; from 2027 onward, the JV may need to commit to a new capacity phase to support continued demand growth. The capex commitment will compress free cash flow temporarily but is required for the bull-case revenue path. Watching the JV capex commentary at the Q4 FY26 print and 2027 capex guidance will be the leading indicator of whether SanDisk is prioritising capacity expansion (revenue-positive long-term, FCF-negative short-term) or maximising cash return (FCF-positive short-term, revenue-cap long-term).

Optionality

Three options. First, HBF — the GPU-package flash form factor — is the single biggest free option in the storage complex. Second, the standalone listing structure gives SanDisk strategic flexibility to spin out the JV stake, do a sale-leaseback on fab capacity, or initiate a buyback that the legacy WDC-attached structure could not. Third, enterprise SSD displacing HDD warm-tier storage — the secular substitution — is accelerating, and at current NAND $/GB the crossover economics are reaching the point where the next leg of HDD-to-SSD migration is on the table for 2027-28.

The trade

> **Entry:** $1,280-1,400 — acceptable on the current melt; add on any 15-20% pullback to $1,150. **Size:** 3-4% portfolio target; treat as a torque sleeve rather than a core. **Stop:** $950 (200-day MA / prior breakout retest). **Catalyst date:** Q4 FY26 print July 30, 2026. **Trim/exit:** trim 25% at $1,800, 50% at $2,100; full exit on NAND ASP roll-over in monthly Trendforce data. **Conviction: 8/10**.

Layer 3 · WDC — Western Digital

One-line thesis

The only listed pure-play on AI hard-disk-drive tightness, just printed 50% gross margin for the first time ever with capacity sold out through 2026 and LTAs into 2029 — own it as the contrarian storage call where the market is finally rebuilding the AI-HDD thesis and the HAMR technology ramp extends the duopoly's addressable warm-tier through the next decade.

What Western Digital physically does

Post the early-2025 separation from SanDisk, Western Digital is now a pure-play hard-disk-drive manufacturer — nearline enterprise HDDs in 30TB+ form factor for hyperscaler and cloud storage tier, plus client HDDs for the legacy PC and surveillance markets. The physical product is a precision-mechanical device — multiple platters spinning at 7,200 RPM, perpendicular magnetic recording heads writing at sub-nanometer precision, sealed helium chambers, electromagnetic actuators. For all the press coverage of NAND substitution, the HDD is still the cheapest $/GB persistent storage medium by a 6-10x factor against SSD at the warm/cold tier, and that economic gap is the moat.

The AI stack relevance is the warm-tier storage explosion. Training a frontier LLM requires multi-petabyte training datasets that have to live somewhere accessible. Agentic AI workflows produce

checkpoint and intermediate state at orders of magnitude greater volume than supervised training did. Vector embeddings for RAG and tool-use produce permanent storage footprints. Hyperscalers ran the numbers in 2024-25 and concluded that even with NAND/HBM tightness, the warm-tier of AI infrastructure is HDD-economic — the gap is just too big to close in this cycle. The result is an AI-HDD supercycle that was not in any sell-side model two years ago and that drove WDC capacity to **sold out through 2026** with **LTAs extending to 2029**.

WDC is one of two listed HDD pure-plays (the other being Seagate); these two duopoly the global nearline HDD supply with roughly 50/50 share. The physical capacity expansion is constrained — building a new HDD fab requires multi-year lead time on head technology, media manufacturing and clean-room build — and the duopoly structure means neither incumbent is rushing to add aggressive supply. That is the supply-side reason for the tightness.

The technology roadmap is the demand multiplier. WDC's HAMR (heat-assisted magnetic recording) technology is the path to per-platter capacity doubling — from 28-30TB drives today to 50TB+ in 2027 and 80TB+ by 2029. HAMR uses a laser-assisted write head to temporarily heat the magnetic media at the write point, allowing smaller magnetic grains and therefore higher areal density. The technology has been in development for over a decade and is now at production qualification at multiple hyperscaler customers. Each per-drive capacity step takes another bite out of the SSD substitution math; at 80TB per drive, the per-GB cost advantage versus enterprise SSD extends to a 12-15x gap, which is structural even if NAND prices fall back to trend.

The financial print

Q3 FY26 (March quarter, reported May 4) printed **revenue $3.337 billion, +45.5% YoY**, EPS **$2.72** beating consensus by 14%, and — the milestone print — **gross margin 50.5%, the first time ever** that WDC has crossed the 50% line. The dividend was raised **+20%** alongside the print, signalling management's confidence in the durability of the cash generation. LTAs extending to **2029** were disclosed, which is unprecedented in HDD history; prior LTA structures topped out at 18-24 months.

The stock closed **$482.02 on May 16** and is up roughly **+170% year-to-date**. The 24/7 Wall St $500 target is essentially consensus now; the more bullish sell-side prints are running $600-700 with the LTA visibility as the anchor. Q4 FY26 reports **July 30**.

The 50.5% gross-margin print is structurally significant because the HDD business has historically operated in the 26-32% gross margin range through cycle averages. The mechanism for the step-up is product-mix toward nearline enterprise (30TB+ drives carry meaningfully higher per-drive gross margin than client/consumer drives), the LTA pricing discipline (no spot-market discount erosion), and the fixed-cost leverage from running fabs near full utilisation. The through-cycle margin floor is now plausibly the 40-45% range rather than the high-20s, which materially rerates the through-cycle earnings power and supports a higher multiple anchor than the historic 6-8x forward earnings.

Customer mix

Hyperscaler concentration is unusually clean for HDD. AWS, Microsoft and Google are the top three customers and together account for the majority of revenue. Meta and Oracle round out the top five. The customer mix is essentially binary — either the buyer is a hyperscaler/large

cloud, or it is a tier-2 OEM with a fraction of the volume; legacy enterprise OEM has shrunk to a rounding error.

The LTAs extending to 2029 are the structural change. Hyperscalers, having watched HBM and now NAND go from spot-priced to allocation-controlled, have moved to lock in HDD pricing and volume on multi-year fixed contracts. WDC has disclosed that LTAs cover the majority of FY26 and a meaningful slice of FY27-28. This is the kind of customer book that produces decade-long earnings durability rather than two-year cyclical recoveries.

The specific structure of these LTAs is informative for understanding the durability. Unlike historical HDD purchase orders — which were 90-day rolling forecasts with reasonably soft commitments — the AI-HDD LTAs include non-cancellable annual volume floors, fixed-price escalators tied to capacity-tier (per-TB pricing rather than per-drive pricing), and option clauses for capacity ramps. The customer is effectively pre-paying for guaranteed allocation in a tight market. From a credit-quality perspective, the LTA structure is closer to an aerospace supplier contract than a semiconductor commodity PO; the through-cycle earnings durability that flows from it has not historically existed in the HDD industry.

Competitive context

The HDD market is a true duopoly: WDC and Seagate, roughly 50/50 share, with Toshiba/Kioxia as a small third player. Both incumbents are price-disciplined, capacity-constrained, and effectively rationing supply to the largest customers at fixed escalators. The competitive context is structurally favourable in a way that no other Layer 3 memory/storage business is — no Korean state-backed competitor, no Chinese fab capacity ramping in stealth, no rising fourth or fifth player. The duopoly is stable, mature, and now operating at peak utilisation with peak pricing. The barriers to entry are extremely high — head-technology IP, media-manufacturing scale, clean-room capex, and the proprietary servo and signal-processing engineering — which structurally prevents the kind of competitive entry that has periodically broken up semi memory pricing power.

WDC's competitive advantage relative to Seagate is principally on the post-separation balance sheet (cleaner cash structure), the technology pipeline on HAMR (heat-assisted magnetic recording — the next-generation areal-density technology that pushes per-platter capacity above 50TB), and the OEM relationship structure that has historically favoured WDC for the largest cloud accounts. Against Seagate the two trade roughly in line; the trade is the HDD duopoly, not the individual name.

The reason a duopoly trade is sustainable rather than competed-away is industry consolidation. Through the 2010s the HDD industry consolidated from eight credible vendors to three to two — Seagate, WDC, and Toshiba (which now exists inside the Kioxia ecosystem as a smaller third player). The consolidation purged the marginal capacity and the price-warring vendors, leaving a structurally disciplined duopoly that does not chase share through pricing. That discipline is the precondition for the current pricing-power episode; without the consolidation that happened over the prior decade, the AI-HDD tightness would have been competed away within two quarters by capacity expansion. As is, both incumbents are deliberately capping capacity expansion at single-digit per cent annual to preserve pricing.

Terminal risk

The terminal risk is two-pronged. First and most cited is SSD substitution — if NAND $/GB falls fast enough to crack the cost-per-bit gap with HDD, the warm tier migrates to SSD and the AI-HDD thesis evaporates. The historical $/GB gap has been 6-10x; that needs to close to 2-3x for substitution to become structural. Current NAND pricing is moving in the opposite direction (up, not down), which is the proximate reason the AI-HDD thesis is holding. Second is hyperscaler capex blink — if any major hyperscaler trims its AI capex plan, HDD demand is the first variable cost to fall, because storage is more elastic than compute. The terminal risk is real but the gating numbers are visible.

The way to track the substitution risk in real time is via the monthly $/GB ratio between enterprise nearline HDD and high-capacity enterprise QLC SSD. As of May 2026 that ratio sits roughly 8-9x in HDD's favour. The trigger for substitution acceleration would be NAND ASPs falling 40-50% from current levels while HDD ASPs hold, which would close the gap to 4-5x. That scenario requires both a meaningful NAND capacity surge AND continued HDD pricing discipline — a combination that is structurally improbable in the next 18-24 months given current NAND tightness. The terminal risk is therefore probably a 2028-29 issue rather than an immediate concern.

Bull case

The three-to-five-year bull case rests on the LTAs holding, HAMR technology pulling the next areal-density step (50TB+ per drive by 2027, 80TB+ by 2029), and the AI-HDD demand curve sustaining through the back half of the decade. FY27 revenue at $14-15 billion versus $12 billion FY26 consensus, EPS in the $14-16 range, a 10-12x multiple — the HDD duopoly probably deserves higher than the historic 6-8x — and the stock works to $600-700.

The stretch bull case adds the dividend trajectory becoming a real income story. A 20% raise in FY26 plus sustained free cash flow north of $4 billion supports a 5-8% dividend yield at current prices, which pulls in a different class of buyer (yield-focused, income-funds, retiree allocations) than the traditional tech-cyclical buyer base. That re-rating mechanism — from cyclical to dividend compounder — is the underappreciated leg.

Gap / bear case

The market is probably under-pricing the duration. WDC has historically been treated as a deep-cyclical and the consensus number assumes mean reversion in FY27-28. The LTAs into 2029 are visible to anyone who reads the print, but the modelling assumption is still that LTAs roll lower as supply catches up. If the structural NAND tightness persists, the SSD substitution gating math does not close, and the LTAs roll at flat or higher pricing into 2030, the FY28-29 numbers are well above consensus.

The bear case the market is right about is the +170% YTD valuation. Stock at $482 against historic high-teens forward multiples is fully priced for the consensus path; the room for multiple expansion from here is finite. The bear case is not that the business breaks; it is that the stock has caught up to the fundamentals and now requires earnings to grind higher.

A specific consensus assumption that may be wrong is the ASP-per-TB trajectory. Most sell-side models bake in 5-7% annual ASP-per-TB declines on the assumption that capacity expansion eventually erodes pricing. The LTAs disclosed at the Q3 print suggest fixed or rising ASP-per-TB through 2027-28, which is materially different from the consensus assumption. If the

LTA-anchored pricing trajectory is the realised path, FY28-29 revenue is roughly 15-20% above current consensus and EPS scales proportionally.

Optionality

Two options worth tagging. First, the HAMR technology ramp is genuinely transformative — if WDC can ship 50TB+ drives at scale in 2027, the $/GB economics shift in favour of HDD against SSD for warm-tier storage, and the duopoly extends its addressable market structurally. Second, the post-separation corporate structure leaves room for capital return acceleration — special dividends, accelerated buybacks, or even a strategic combination — none of which is priced into consensus. A third optionality leg is sovereign-data exposure: the regulated-data warm-tier requirements at government, healthcare and financial-services customers represent a separate demand pool with different growth and pricing dynamics from hyperscaler, and WDC is well-positioned to capture share in that segment.

The trade

> **Entry:** $450-490 — acceptable on the current zone; add aggressively below $420 on any 15% pullback. **Size:** 3% portfolio target. **Stop:** $360 (200-day MA). **Catalyst date:** Q4 FY26 print July 30, 2026. **Trim/exit:** trim 25% at $620; full exit on NAND ASP roll-over OR SSD warm-tier substitution data flipping. **Conviction: 8/10**.

Layer 3 · 285A — Kioxia Holdings

One-line thesis

Kioxia is the Japanese NAND pure-play just out of the IPO base and printing FY operating profit +314% QoQ — the cheapest and most under-owned exposure to AI-storage tightness plus HBF optionality, with the parabolic risk that comes with +1,949% in twelve months, and the Japanese-listed structural diversifier inside a US-NAND-heavy portfolio.

What Kioxia physically does

Kioxia (formerly Toshiba Memory) is the spiritual originator of NAND flash — Fujio Masuoka's invention at Toshiba in 1987 is what created the storage industry the world built around for thirty years. Today Kioxia is the standalone NAND pure-play with co-fab structure alongside SanDisk at the Yokkaichi and Kitakami sites; together the JV represents roughly 30-35% of global NAND supply. Kioxia listed on the Tokyo Exchange in late 2024 after a long private-equity ownership period and has been re-rating against the AI-storage cycle ever since.

Physically, Kioxia produces 3D NAND wafers at the JV fabs, packages into enterprise SSDs and high-density storage modules, and ships into the same hyperscaler/OEM customer base that SanDisk serves. The strategic difference between Kioxia and SanDisk — both jointly manufacturing the same wafers — is geographic listing, capital structure, customer relationships at the OEM tier (Kioxia has historically had stronger Japanese enterprise relationships), and the strategic flexibility of being a Japanese-listed company with a Japanese government interest in semiconductor sovereignty.

The HBF roadmap is shared with SanDisk and is the principal AI-stack optionality. Kioxia and SanDisk co-developed HBF — High Bandwidth Flash — as the GPU-package flash form factor that could displace some HBM-only inference architectures with cheaper-per-bit NAND-class storage. If HBF qualifies at Nvidia or AMD for production inference, the joint franchise re-rates structurally.

The financial print

FY26 (year ended March 2026) printed **revenue ¥2,337.6 billion, +37% YoY**, with Q4 alone at **¥1,002.9 billion — roughly double the prior quarter** — and **full-year operating profit ¥876.2 billion at a record OP margin**. Q4 operating profit alone was **¥599.1 billion, +314% QoQ**. The stock printed **+21.2% on the print day** and has run roughly **+1,949% over the trailing twelve months** — the steepest twelve-month re-rating in Japanese large-cap equity since the early-2000s tech recovery.

The stock closed **¥44,450 on May 15**. The +21% one-day move is the kind of move that says the post-IPO sell-side coverage is still catching up; consensus FY27 estimates are being re-cut weekly. Q1 FY27 reports **August 8, 2026**.

The Q4 operating profit acceleration deserves close reading. **OP +314% QoQ** in a memory business is the kind of operating-leverage print that signals either pricing-power dominance or one-time accounting tailwinds. In Kioxia's case it appears to be the former — NAND ASPs rose, capacity utilisation pushed to 100%, and the JV depreciation profile is currently in a favourable phase. The trailing FY26 operating margin clocks in north of 35%, which is a structural high for the Kioxia business and is broadly comparable to the SanDisk print on the same JV output base. The consistency between SanDisk and Kioxia operating-margin trajectories confirms the JV-output economics rather than vendor-specific accounting choices.

Customer mix

Customer mix mirrors SanDisk's at the JV-output level: hyperscaler enterprise SSD demand at AWS, Microsoft, Google and Meta is the dominant volume line, with high-end OEM enterprise (storage-array vendors) as the second tier. Kioxia retains stronger Japanese-enterprise direct relationships than SanDisk does (Sony, NTT, Japanese cloud operators), which is a small but meaningful diversification slice.

Customer LTAs are similarly extending into 2027 with fixed-pricing escalators. Kioxia has disclosed that NAND ASPs are running well above the FY25 baseline and that capacity is essentially sold out through the year. The visibility is materially cleaner than at any point in Kioxia's recent history.

A useful distinction between Kioxia and SanDisk for portfolio-construction purposes is the geographic-customer skew. SanDisk's customer book leans more heavily toward US hyperscalers (AWS, Microsoft, Google, Meta in the top four); Kioxia's book retains a meaningfully larger slice of Japanese, Korean and broader Asian enterprise customers, plus the strategic Japanese government layer that supports semiconductor sovereignty. The two are not perfectly fungible exposures even though they share fabs; an investor seeking pure US-hyperscaler exposure leans SanDisk, while an investor seeking diversified Asia-Pacific AI-storage exposure leans Kioxia.

Competitive context

Kioxia competes inside the same NAND five-vendor structure as SanDisk: Samsung and SK Hynix/Solidigm on one side, Kioxia-SanDisk JV on the other, with Micron as a smaller fifth player. The competitive context is favourable for the same reason SanDisk's is — the JV manufacturing footprint is post-depreciation cost-advantaged, the customer relationships are sticky, and the technology roadmap on BiCS 9 and BiCS X high-layer NAND is competitive.

The technology roadmap deserves attention because it is the principal differentiator between Kioxia/SanDisk and the Samsung-Hynix consortium in the NAND segment. BiCS-class 3D NAND uses a charge-trap-flash architecture with sequential layer deposition; Samsung's V-NAND uses a floating-gate-style cell with vertical hole etching. The BiCS architecture has scaled to 218-layer and is targeting 300+ layer nodes through 2027-28, which provides bit-density parity-to-leadership against Samsung's V-NAND roadmap. Kioxia has historically been the technology owner of the BiCS process — the original Fujio Masuoka legacy combined with the Yokkaichi engineering team — and the technology direction is set inside Kioxia's R&D rather than at the SanDisk side of the JV.

The differentiator for Kioxia versus SanDisk inside the same JV is principally the listing geography. Japanese-listed Kioxia is the natural home of Japanese institutional capital, the Japanese government's semiconductor strategic-fund money, and the Asia-Pacific allocation buckets that under-own the AI-storage trade because the American HBM names dominate global tech allocations. As more capital rotates into AI-storage as a sub-theme distinct from AI-compute, Kioxia is the natural geographic vehicle, and that mechanical allocation pressure is itself a re-rating engine.

Beyond the JV-output-level moat, Kioxia's standalone competitive positioning rests on three additional legs. First, the Yokkaichi and Kitakami fabs are post-depreciation and operationally efficient — the all-in cost-per-bit is lower than either Hynix or Micron and comparable to Samsung. Second, the customer relationships at the OEM enterprise tier (storage-array vendors like Pure Storage, NetApp, plus Japanese enterprise) provide diversification away from pure hyperscaler concentration. Third, the post-IPO strategic clarity allows Kioxia to invest more aggressively in sales and marketing infrastructure than its previous private-equity-owned structure permitted, which is a positive operational lever that consensus models likely under-weight.

Terminal risk

Terminal risk is identical to SanDisk's plus the parabolic-position risk. **+1,949% trailing twelve months** is not a number that can repeat, and any reasonable mean-reversion math says the next twelve months produce a meaningfully smaller gain or a meaningful give-back. The terminal risk is positioning unwind on any AI-capex blink; the secondary terminal risk is the JV-structure friction with SanDisk if HBF prioritisation or capex disagreements emerge.

A particular technical risk worth flagging is the post-IPO lock-up cadence. Several tranches of legacy-holder shares will become eligible for distribution through 2026-27 under the original IPO lock-up schedule. The market knows about these tranches and prices some of the supply overhang in, but each distribution event creates a discrete period of indigestion. Watching the calendar of unlock dates is part of the trade discipline; coincidence of an unlock with a cycle-peak signal would be the worst-case technical setup.

A tertiary risk is Chinese NAND capacity expansion. YMTC (Yangtze Memory Technologies) has

been ramping domestic Chinese NAND fabs through 2024-26 with state-subsidised capex; the absolute capacity additions are still small in global context but the trajectory is steep enough that by 2028-29 Chinese supply could be meaningful at the volume tier. This is structurally negative for the pricing power of all five major NAND vendors including Kioxia, even if Chinese product is qualitatively behind on technology.

Bull case

The base bull case sees Kioxia capturing the Japanese geographic re-rating bid as institutional allocations move toward Asia-Pacific AI-storage exposure, FY27 revenue running ¥2,800-3,000 billion at sustained 35%+ operating margins, and a multiple expansion to 15-18x forward earnings as the post-IPO sell-side coverage matures. That math takes the stock to ¥55,000-65,000 over twelve to eighteen months.

The stretch bull case adds HBF qualifying as a real GPU-package flash tier by 2027-28 — a step-change re-rating that pulls the multiple to high-teens and the stock to ¥80,000+. The HBF leg is the same option as for SanDisk; Kioxia is arguably the principal technology owner on the joint development.

The post-IPO supply-and-demand dynamic deserves attention. Kioxia's listing in late 2024 brought a meaningfully smaller free-float to market than the eventual long-term flotation will support; legacy private-equity holders (Bain Capital, Korean SK Hynix) retain large stakes that will distribute over time. The float overhang is a near-term technical headwind that is concurrent with the +1,949% TTM move; the eventual conversion to broader institutional ownership is itself a re-rating mechanism over 18-36 months as index inclusion criteria are met and Japanese institutional allocations adjust.

Gap / bear case

What the market may be missing is the durability — Japanese-listed equities in AI sub-themes have historically been under-owned by global investors and the post-IPO re-rating cycles tend to extend for longer than the parabolic move suggests. The bear case the market is right about is the parabolic mean reversion: 19x in twelve months has historically been followed by 40-60% give-backs in Japanese mid-caps. Position sizing has to respect that asymmetry.

The market may also be under-pricing the JV-structure execution risk. SanDisk's larger and faster-moving listed structure may pull HBF prioritisation in a direction Kioxia's slower-moving board does not optimise for; conversely, SanDisk's IPO-investor pressure for short-term capital return could undermine the joint-fab capex programme. These are second-order risks but real.

A subtle bull-case omission worth highlighting is the optionality value of the BiCS technology licensing pipeline. Kioxia has historically licensed elements of its BiCS process technology to non-JV partners (Chinese NAND vendors, Korean memory players) under structured royalty agreements. As the NAND industry consolidates and the technology becomes a strategic asset, the licensing-revenue stream could become a non-trivial high-margin annuity. Consensus models effectively zero this line; if Kioxia commercialises BiCS licensing more aggressively, it adds 200-400 basis points to consolidated operating margin without capex requirement.

The Yen exposure is a complicating dimension. A meaningful slice of Kioxia's costs are JPY-denominated (Yokkaichi and Kitakami fab operations) while a meaningful slice of revenue is USD-denominated (hyperscaler enterprise SSD pricing). A strengthening Yen would compress

reported margins; a weakening Yen would inflate them. The current JPY/USD trajectory has been broadly supportive but is itself a macro variable rather than a fundamental one. Investors taking the Kioxia trade should think about whether they want the FX exposure or want to hedge it via paired US listings.

Optionality

Two big options. First, HBF — same as for SanDisk, the GPU-package flash form factor that could re-rate the joint franchise. Second, Japanese government strategic ownership and METI semiconductor strategy is a potential catalyst — the Japanese state has financial interest in keeping Kioxia as a sovereign semiconductor champion, and any government-backed capital return programme, capex subsidy, or strategic consolidation involving Kioxia would be a structural positive that consensus does not currently price. A third option is the BiCS technology licensing pipeline mentioned above; a fourth is potential consolidation across the NAND industry — a Kioxia-SanDisk full combination or a strategic transaction with a Korean memory peer would each be re-rating events that current consensus does not contemplate.

The trade

> **Entry:** ¥42,000-46,000 — current zone acceptable; aggressive add on any 20% pullback to ¥35,000. **Size:** 2-3% portfolio target; sized smaller than US NAND names due to parabolic risk. **Stop:** ¥30,000 (post-IPO base breakout retest). **Catalyst date:** Q1 FY27 print August 8, 2026. **Trim/exit:** trim 30% at ¥60,000; full exit on NAND ASP roll-over in monthly Trendforce data. **Conviction: 8/10**.

Layer 4 · ALAB — Astera Labs

One-line thesis

Astera Labs is the open-standards UALink and CXL connectivity pure-play just out of a +93% YoY quarter — own it as the agentic-AI back-end network play, the structural alternative to Nvidia's NVLink moat, and the cleanest exposure to the AMD MI400 ramp at the four major hyperscalers.

What Astera Labs physically does

Astera makes the silicon that connects AI accelerators to each other and to memory across the back-end of the data centre fabric — specifically PCIe retimers, CXL memory expanders and now the Scorpio X-Series of switch silicon that anchors the UALink open-standards alternative to Nvidia's NVLink. In a frontier AI training cluster, every GPU sits inside a node that connects to seven other GPUs via NVLink (Nvidia's proprietary fabric); those nodes then connect to other nodes via the back-end network. Astera's silicon is in that back-end interconnect path for every non-Nvidia accelerator family — AMD MI355X/MI400, Intel Gaudi, AWS Trainium, Microsoft Maia, Google TPU — and increasingly inside Nvidia-based clusters too for PCIe-class peripheral connectivity and CXL memory pooling.

The physical product is small but the value capture is large. A typical 100-rack AI cluster might have 500-1,000 Astera retimer chips plus a smaller number of Scorpio switch silicon SKUs, and the design wins are 5-7 year sole-source qualifications because the engineering effort to swap

signal-integrity silicon mid-platform is non-trivial. The AI-stack layer Astera sits in is Layer 4 — interconnect, packaging-adjacent — and it is one of the binding constraints once you scale past a few-hundred-rack training cluster.

The strategic moat is the UALink standard. UALink is the open-standards consortium-backed alternative to NVLink, co-founded by AMD, Intel, Microsoft, Meta, Google, AWS and Astera. The hyperscaler logic is straightforward — Nvidia's NVLink is a proprietary fabric that locks the entire cluster architecture to Nvidia. UALink is the open-standards version of the same thing, and the hyperscalers are committing real engineering resources to making UALink the default back-end for their non-Nvidia accelerator fleets. Astera is the principal silicon supplier inside that ecosystem.

The agentic-AI angle is the under-appreciated demand driver. Training-dominated AI workloads — the LLM pretraining clusters of 2023-25 — were characterised by very high bandwidth requirements between adjacent GPUs (handled by NVLink intra-node) but relatively predictable east-west traffic patterns at the cluster level. Agentic-AI inference workloads are different. They produce highly variable, low-latency, fanout traffic patterns as agents call tools, query vector databases, exchange state with peer agents, and stream context across multiple inference engines. The back-end fabric in an agentic-AI cluster has to be flexible, lossless, and able to handle bursty traffic patterns — which is precisely the architectural niche UALink and Astera's switch silicon are designed for. The shift from "training-dominated" to "inference-and-agentic-dominated" AI capex is the secular demand lever that pulls Astera content per cluster up by 3-5x over the next thirty-six months.

The financial print

Q1 2026 (reported May 5) printed **revenue $308.4 million, +93% YoY**, EPS **$0.61 versus $0.49 consensus**, and management guided to a Scorpio X-Series ramp through the second half of the year that is materially above prior expectations. The growth is organic — no acquisitions, no big one-time licence revenue — and the operating leverage is intact: incremental gross margin in the 70s on the new Scorpio SKUs, operating margin trajectory from the high-20s toward the high-30s through FY27.

The stock closed **$230 on May 17** at a roughly **$39.88 billion market cap**. Q2 2026 reports **August 4**. Analyst price targets clustered in the $250-300 range with the top of the range at $350.

The +93% YoY revenue growth is particularly informative because the prior-year base was already non-trivial — this is not a small-base growth print. Astera's quarterly revenue run-rate has moved from roughly $40-50 million in early 2024 to $300+ million in Q1 2026, a roughly 6-7x expansion in less than two years. The growth is also broad-based rather than single-product; Aries retimers, Taurus Ethernet, Leo CXL, and Scorpio X-Series are all contributing simultaneously, which suggests the underlying TAM expansion is real rather than a product-cycle blip. The consensus FY26 revenue track is now $1.4-1.6 billion against $432 million in FY25, with FY27 consensus near $2.2 billion — and the sell-side has been incrementally raising these numbers each quarter.

Customer mix

Hyperscaler concentration is heavy and is the strength rather than the weakness of the story. AWS, Microsoft, Google and Meta are the four anchor customers; AMD's accelerator platform sale is a downstream multiplier across every AMD MI355X/MI400 socket that ships into those four. The top-four customer concentration is probably 70%+ of revenue. The customer-level engagement runs deeper than typical silicon supplier relationships — Astera's engineers embed with hyperscaler architecture teams during the multi-quarter platform design phase, which creates information asymmetry on the upcoming product roadmaps and effectively locks in the next-generation design win before competitors are in the conversation.

The customer commitment is unusually deep. UALink consortium membership is not a casual marketing line — AMD, Microsoft, Meta and Google have committed multi-year design-win pipelines to UALink, and Astera is the only silicon vendor with first-mover product depth across the retimer + switch + memory-expander stack. The customer mix is structurally favourable because each design-in is a 5-7 year revenue annuity, not a quarterly competitive bid.

The design-in dynamics matter for understanding why concentration is a strength rather than a weakness. Each Astera retimer or switch silicon design-in goes through 12-18 months of joint-engineering with the hyperscaler customer — signal-integrity simulation, package co-design, software stack integration. Once a customer has invested that engineering effort, the cost of switching to a competing vendor (even on a price-attractive bid) is prohibitive because it requires redoing all that work. This is the structural lock-in that produces 30-40% operating margins in a notionally commodity-ish silicon segment; the engineering depth becomes the moat rather than the price.

Competitive context

The competitive set inside the UALink ecosystem is narrow — Marvell competes on some retimer SKUs, Broadcom competes on switch silicon, but neither has the breadth across UALink-specific design wins that Astera has, and neither is a hyperscaler-aligned consortium member with the same level of customer mind-share. Inside the NVLink ecosystem (Nvidia-captive), Astera does not compete directly — that fabric is Nvidia's own silicon.

The moat is the combination of standards body authorship (Astera helped write UALink), the engineering relationships with hyperscaler architects, and the product breadth across the connectivity stack. Margins are evidence — Astera runs operating margins above the chip-design peer average, which is the financial signature of pricing power earned through engineering depth rather than commodity scale.

The product portfolio breakdown matters for understanding the operating-leverage trajectory. The Aries retimer family is the legacy AI-cluster product, currently the largest revenue line, and is in a recurring-design-win cycle as hyperscalers refresh platforms every 18-24 months. The Taurus Ethernet products serve a smaller adjacent market in AI-back-end networking. The Leo CXL memory expanders are the longer-tail high-margin product line that becomes important as CXL deployments scale. And the Scorpio X-Series — the newest line, the one driving the +93% YoY print — is the switch silicon that anchors UALink networks. Scorpio is the highest-ASP product in the portfolio and the principal margin-expansion lever; the Scorpio mix shift from 0% in 2024 to 20%+ in 2026 is what is pulling consolidated operating margin from the high-20s toward the high-30s.

Terminal risk

The terminal risk is Nvidia's response. If Nvidia opens NVLink to non-Nvidia accelerators — which it has flirted with on multiple occasions — the strategic logic for UALink weakens at the margin and Astera's design-win pipeline compresses. The probability of full NVLink opening is low (it would undermine Nvidia's principal moat) but the probability of partial NVLink licensing to specific hyperscalers (cynically, to slow UALink adoption) is non-trivial.

The secondary terminal risk is hyperscaler captive silicon — every major hyperscaler has internal connectivity-silicon programs that could displace Astera's design wins over time. The mitigation is that Astera's engineering velocity and the standards-body alignment make displacement non-trivial; the risk is real but probably a 2028-29 issue rather than a 2026-27 issue.

A third terminal risk is standards-body fragmentation. UALink itself is a consortium effort with multiple competing technical proposals; if any major member (AMD, Intel, Microsoft, AWS) pushes for a competing fabric standard or splits to back its own captive approach, the unified UALink platform could fragment and Astera's first-mover advantage could be compressed across multiple competing standards. This is structurally improbable given the current hyperscaler alignment, but is worth tagging as a tail risk.

Bull case

The three-to-five-year bull case sees UALink adoption reaching 30-40% of non-Nvidia accelerator sockets by 2028, Astera retaining principal silicon supplier position, and FY28 revenue running $2.5-3 billion versus FY26 consensus closer to $1.5 billion. Operating margin in the high-30s. EPS in the $5-7 range. A 35-40x multiple — connectivity-silicon pure-plays with secular standards-body moats deserve a premium — gets the stock to $250-300 on a 24-month view and $400+ on a 36-month view.

The stretch bull case adds CXL memory pooling becoming a real production architecture for AI inference workloads — every CXL-pooled memory module needs Astera's memory-expander silicon — and Scorpio X-Series becoming the dominant back-end switch silicon outside Nvidia's NVL fabric. That path adds 30-50% to the base case and re-rates the multiple.

The bull-case engineering bet is on AMD's MI400 platform shipping at the volume hyperscalers are forecasting. If MI400 captures even 15-20% of net new accelerator sockets at AWS, Microsoft, Google and Meta through 2027-28 — which is the consensus directional call — every MI400 socket is an Astera content opportunity worth $4,000-7,000 in silicon BOM. The math compounds quickly: 500,000 MI400 sockets times $5,000 average content equals $2.5 billion of revenue from a single platform inflection. That is not in the FY26 consensus number; it materialises in FY27-28 if the AMD ramp executes.

Gap / bear case

What the market may be missing is that the +93% YoY print is the first inflection in a multi-year design-win cycle, not the peak. Each UALink-aligned platform that ramps adds an annuity-class revenue line that is not yet visible to consensus. The market is right about the valuation — $40 billion market cap on a ~$1.5 billion revenue run-rate is rich — but the durability and growth slope arguably justify the premium.

The bear case the market may be under-pricing is hyperscaler captive silicon displacement. AWS'

internal connectivity team, Google's Aquila program, Microsoft's Hollow Core fibre programs — these are all real engineering threats that could compress Astera's design-win durability in the 2028+ window.

A second under-priced bear-case is competitive entry by Broadcom and Marvell at the switch-silicon tier. Both companies have the engineering depth and customer relationships to develop UALink-compatible switch products if the market opportunity becomes large enough; their current absence from the segment reflects strategic choice rather than capability gap. If either decides to enter aggressively in 2027-28, the Scorpio X-Series competitive moat compresses meaningfully. The mitigation is that Astera's standards-body authorship and first-mover qualification creates a switching cost, but it is not absolute.

Optionality

Two options. First, CXL — the memory-pooling architecture is still mostly a roadmap rather than a deployed reality, and Astera is the principal silicon vendor when it does deploy. Second, the optical co-packaged silicon ramp — Astera's silicon-photonics roadmap, currently in early product, is a potential adjacency that could materially expand the addressable market over a 36-48 month horizon. A third option is the agentic-AI workload acceleration creating new connectivity-silicon categories (low-latency RDMA fabric, in-network accelerators) that Astera is positioned to enter as the standards mature.

The trade

> **Entry:** \$215-235 — current zone acceptable; aggressive add below \$180. **Size:** 2-3% portfolio target. **Stop:** \$145 (200-day MA / post-IPO breakout retest). **Catalyst date:** Q2 2026 print August 4, 2026; UALink adoption news flow ongoing. **Trim/exit:** trim 25% at \$320, 50% at \$380; full exit on UALink momentum reversal or Nvidia NVLink opening to non-Nvidia ASICs. **Conviction: 8.5/10**.

Layer 4 · LITE — Lumentum Holdings

One-line thesis

Lumentum is the under-shipped hyperscaler optics name with a +90% YoY record quarter, Nasdaq-100 inclusion bid layered on top — own it as the photonics-tightness play with a passive-flow tailwind into a structurally short market, sized for the 1.6T transition through 2027.

What Lumentum physically does

Lumentum manufactures the optical components that move data inside and between AI data centres — laser diodes (VCSELs and EELs), receivers (photodiodes), and the integrated transceiver modules that plug into the back of every AI server's networking port. Inside a frontier training cluster, every node-to-node connection runs over fibre at 400G, 800G or 1.6T optical speeds, and the bidirectional optical link at each end of every fibre is a transceiver. Lumentum sells the silicon photonics and III-V semiconductor components that go inside those transceivers, and increasingly the fully-assembled transceiver modules themselves.

The AI-stack layer is Layer 4 — interconnect, photonic — and it is the second binding constraint after memory in current-cycle AI build-outs. As GPU clusters scale past a few thousand nodes, the back-end fabric becomes the bottleneck, and the back-end fabric is optical. The migration from 400G to 800G to 1.6T transceivers — driven by Nvidia, Broadcom, Marvell and AMD silicon roadmaps — is structurally pulling Lumentum's content per port from $200-300 to $600-1,200, and the unit volume is exploding alongside.

Lumentum's competitive position inside this market is anchored by two product lines: indium-phosphide laser diodes at the highest-speed nodes (where Lumentum has historic technology leadership against AAOI, Source Photonics and Inphi) and the integrated transceiver assembly business that consolidates the III-V components with packaging and test. The recent Cloud Light acquisition added scale in datacom transceivers and is now contributing meaningfully to the revenue line.

The physics of why indium phosphide matters at the highest speeds is worth understanding. Silicon photonics can serve 400G and most 800G applications with modulator-based architectures, but the highest-density 1.6T and beyond transceivers require external modulated lasers (EMLs) built on indium-phosphide substrate rather than silicon. Lumentum is one of three suppliers worldwide with the InP epitaxy capacity and process know-how to produce these lasers at scale — Coherent and a handful of Chinese vendors are the others. The InP supply chain itself is a bottleneck (which is why AXT, the InP substrate vendor, has been a separate AI-photonics trade); Lumentum's vertical integration into InP gives it both supply security and a margin advantage at the highest-speed nodes.

The financial print

Q3 FY26 (reported May 7) printed **record revenue $808.4 million, +90% YoY**, EPS **$2.37** beat consensus, and Q4 guide of **$960 million-$1.01 billion** with EPS **$2.85-$3.05** — well above sell-side. Operating margin printed at **32.2%**, a structural high for Lumentum and confirmation that operating leverage in the AI-datacom mix is now flowing.

The critical disclosure on the call was that Lumentum is **under-shipping demand by approximately 30%** — capacity is the gating constraint, not demand. That is the optical-photonics equivalent of the HBM-sold-out story, and it is what justifies the FY27 consensus walk-up that the sell-side has been doing weekly.

The stock closed **$953.90 on May 17**. Nasdaq-100 inclusion announced for **May 18** (replacing CSGP), which adds a passive-flow tailwind of roughly $4-6 billion of forced buying across the QQQ ecosystem in the days following inclusion. Q4 FY26 reports **August 7**. Median sell-side price target is $1,100 with high $1,400 and low $600.

The Q3 print mechanics deserve a closer look. Revenue acceleration from sub-$500 million quarterly run-rates two years ago to $808 million in Q3 with $960M-$1.01B guided for Q4 implies a roughly doubling of revenue inside an 18-month window. The 32.2% operating margin print is structurally higher than Lumentum's historical 18-22% through-cycle range — a step-change that reflects both the AI-datacom mix shift and the under-shipping pricing power. The combination of revenue doubling and operating margin expanding by 1,000+ basis points produces the kind of EPS compounding that is normally only seen in software-class businesses. The Q4 EPS guide of $2.85-$3.05 implies annualised run-rate EPS approaching $12, materially ahead of consensus FY27 estimates that still anchor to $9-10.

Customer mix

Hyperscaler concentration is heavy and increasing. The top four hyperscalers (AWS, Microsoft, Google, Meta) plus the AI-focused tier (Oracle, CoreWeave, x.AI) account for the bulk of Lumentum's datacom revenue. Apple is a separate large customer through the consumer-3D-sensing line, but that segment is now a fraction of the AI-datacom business. The acquired Cloud Light business deepened relationships at AWS specifically and is now the principal volume engine for the 800G transceiver line; the integration has executed cleanly and is contributing margin accretion ahead of the original deal model.

The customer book is contracted and tightening. Lumentum has disclosed multi-year supply commitments with hyperscaler customers for 800G and 1.6T transceiver production through 2028, which is the kind of visibility that justifies capex acceleration. The customer mix is heavily skewed to the AI-datacom thesis — a positive durability signal in the up-cycle and a concentrated downside if hyperscaler capex blinks.

Competitive context

The optical transceiver market is more fragmented than the memory market but is consolidating. Direct competitors include Coherent (the closest peer, larger but more diversified), AAOI (the high-velocity small-cap challenger with Amazon as anchor customer), Ciena (more diversified network-equipment vendor), Innolight (the dominant Chinese transceiver assembler), and the captive silicon-photonics teams at Broadcom and Marvell. The combined market structure is a mid-single-digit handful of credible suppliers at the highest speeds and a longer tail at the lower-speed legacy nodes. The competitive structure is therefore more accurately a tight-oligopoly at the leading edge and a fragmented commodity market at the trailing edge, which is the precondition for sustained pricing power at the leading edge.

Lumentum's competitive moat is the indium-phosphide laser-diode technology at the highest-speed nodes (1.6T and beyond), the integrated transceiver assembly capability post-Cloud Light, and the hyperscaler relationships at the architect level. Against Coherent, Lumentum runs a more focused datacom mix; against AAOI, Lumentum has scale and is qualified across more customers. The competitive context is favourable in the up-cycle and probably extends through the 1.6T transition.

The pricing dynamic in optical transceivers has flipped over the past 18 months in a way that mirrors the HBM and HDD pricing flips. Through 2023-24 transceiver pricing was negotiated quarterly with the customer holding leverage; through 2025 pricing moved to annual contracts; through 2026 multi-year LTAs are emerging at fixed-pricing escalators. Lumentum has explicitly captured this transition in its commentary. The under-shipping disclosure (shipping 70% of demand) is not just a capacity story; it is a pricing-power story, because demand-above-supply is precisely the precondition for sustained margin expansion. The Q3 FY26 print at 32% operating margin is not a peak — it is a structural step-up that should hold through the 1.6T transition.

Terminal risk

The terminal risk is silicon photonics co-packaging — the move from pluggable transceivers to optically-coupled silicon directly on the switch ASIC. If silicon photonics CPO (co-packaged optics) replaces pluggable transceivers in the 2027-29 window, the entire transceiver value chain compresses and Lumentum's pluggable-module revenue line goes structurally lower. The mitigation

is that Lumentum's component-level III-V semiconductor business (laser diodes, photodiodes) is needed inside CPO modules too, so the displacement is partial rather than total — but the value capture per port is lower.

The secondary terminal risk is the Chinese transceiver assemblers (Innolight, Hisense Broadband) capturing more share at the volume tier, which compresses pricing at the 800G node before the 1.6T transition fully ramps.

A third terminal risk worth flagging is the inventory-cycle pattern in datacom optics. Historically, hyperscaler optical inventory builds for 9-12 months at a time before customers pause to absorb deployed capacity. The current cycle has been running roughly 18 months of continuous ramp without a meaningful pause, which is unusual and probably reflects the AI-cluster scale-up rather than normal capacity-refresh patterns. The risk is that hyperscalers eventually pause to digest deployed inventory, which would produce a one-to-two-quarter air pocket in transceiver orders even if the structural demand curve continues higher. That pattern would be visible in the monthly customer-mix commentary at Q4 FY26 print or the Q1 FY27 print.

Bull case

The three-to-five-year bull case sees Lumentum riding the 1.6T transition through FY28, with revenue running $5-6 billion versus FY26 run-rate closer to $3.5 billion, operating margins holding in the low-30s, and EPS in the $14-16 range. A 25-30x multiple — earned through the AI-datacom-photonics-bottleneck moat plus Nasdaq-100 passive flow — gets the stock to $1,200-1,500 over a 24-month view.

The stretch bull case adds CPO becoming a real architecture that Lumentum is the principal III-V supplier into, plus continued under-supply through 2027, plus the Apple-3D-sensing line stabilising. That path takes the stock to $1,800+ on a 36-month view.

The under-shipping disclosure deserves an additional layer of analysis. When a manufacturer publicly states that it is shipping 70% of demand, the message to customers is that capacity is being rationed and that pricing will reflect scarcity. Lumentum's competitive set — Coherent, AAOI, the Chinese assemblers — has not made equivalent disclosures, which suggests Lumentum is either uniquely positioned at the highest-speed nodes (1.6T) where capacity is structurally tightest, or is more transparent about supply-demand dynamics for strategic-customer-management reasons. Either reading is positive for through-cycle pricing power. The Nasdaq-100 inclusion combined with the supply-tightness disclosure creates a particular kind of technical setup — passive index buying meeting fundamental supply scarcity — that historically produces extended momentum runs.

Gap / bear case

What the market may be missing is the under-shipping disclosure. **Lumentum is shipping 70% of what customers want** — that is a structurally tight market where pricing power persists through the next 24 months. Consensus FY27 revenue at $4-4.5 billion likely understates the trajectory if capacity expansion stays disciplined.

The market is right about the +90% YoY print being a hard number to grow on. The base-effect compression in FY28 is real, and a deceleration to +30-40% YoY in FY28 is mathematically inevitable. The bear case is that the stock has priced perfect execution and any operational stumble (capacity delays, qualification slips, hyperscaler order rephasing) is punished disproportionately.

A specific consensus gap worth flagging is the operating-margin trajectory assumption. Most sell-side models assume operating margin compresses from the Q3 32% print back toward the 25-28% historical range over FY27-28 on competitive pressure and mix normalisation. The bull-case alternative is that the 32% margin holds or expands because the under-supply position persists and the 1.6T mix shift pulls ASPs higher. If operating margin holds at 32% rather than compressing to 28%, that is roughly $1.50-2.00 of additional FY27 EPS that consensus does not capture — a meaningful asymmetry on a stock trading at $953.

Optionality

Three options. First, CPO — silicon photonics co-packaged optics — is an inevitable architecture transition; Lumentum's III-V component business is positioned for it even if the pluggable-module business compresses. Second, the Apple 3D-sensing line stabilising or expanding (LiDAR-class consumer applications, AR/VR headsets) adds a non-AI revenue diversification. Third, the Nasdaq-100 inclusion is a discrete passive-flow event that creates ~$4-6 billion of mechanical buying — the price action over the 30 days post-inclusion is independent of fundamentals. A fourth option is the indium-phosphide supply chain — Lumentum's vertical integration into InP provides margin protection and could become a revenue line if InP capacity is licensed to third-party transceiver assemblers. The combination of these four optionality legs is unusual for a single name and supports the conviction sizing.

The trade

> **Entry:** $920-980 — acceptable; aggressive add below $850. **Size:** 3% portfolio target. **Stop:** $700 (200-day MA). **Catalyst date:** Nasdaq-100 inclusion May 18 (today); Q4 FY26 print August 7. **Trim/exit:** trim 25% at $1,200, 50% at $1,400; full exit on hyperscaler capex blink or CPO displacement evidence. **Conviction: 8.5/10**.

Layer 4 · 522 — ASMPT Limited

One-line thesis

ASMPT is the TC bonder share-taker that just won the SK Hynix HBM4 socket from Hanmi with bookings +72% YoY — own it as the back-end equipment leader at the inflection point in HBM stacking economics, with hybrid bonding and HBM5 as the next-cycle optionality and a Hong Kong listing as the geographically-distinctive vehicle.

What ASMPT physically does

ASMPT (formerly ASM Pacific Technology, Hong Kong-listed) is the global #2 semiconductor back-end equipment vendor after Hanmi for thermo-compression bonding (TC bonding) — the precision pick-and-place machines that stack HBM dies vertically with sub-micron alignment under controlled heat and pressure. Across the company's full portfolio, ASMPT also makes die-attach, wire bonders, mold-equipment for advanced packaging, and SMT (surface-mount-technology) lines for downstream assembly. But the AI-stack-relevant business is the TC bonder unit and increasingly the adjacent hybrid-bonding roadmap.

In an HBM line at SK Hynix, Micron or Samsung, the TC bonder is the single most expensive piece of equipment per stacking step, and the throughput of the bonder is the gating variable for the entire fab output. A 12-Hi HBM4 stack requires 11 bonding operations; a 16-Hi HBM5 stack requires 15. ASMPT's TC bonder family directly competes with Hanmi's TC Bonder 4 for the HBM4 socket and has been winning incremental share at SK Hynix through 2025-26 as Hynix has deliberately diversified away from Hanmi's sole-supply position.

The technical specifics of ASMPT's TC bonder positioning matter for the share-loss/share-gain narrative. ASMPT's HBM-class TC bonder platform supports both MR-MUF (the mass-reflow molded underfill process that SK Hynix favours) and TC-NCF (the non-conductive film process that Samsung and Micron historically used) in a single tool — a capability parity with Hanmi's TC Bonder 4. The qualification process at SK Hynix for HBM4 took roughly fourteen months of engineering engagement, which is itself a moat against further competitive entry: any new vendor seeking to compete at SK Hynix's HBM4 line in 2027-28 is structurally late. ASMPT has converted that qualification into recurring tool orders and is now positioned for the HBM4E generation that ramps in late-2026 to mid-2027.

The strategic significance for ASMPT is that the AI-memory ramp is the principal demand driver for its back-end equipment business through this cycle. Hanmi's share loss is ASMPT's share gain — Trendforce reported in December 2025 that ASMPT now operates roughly half of the active HBM4 TCBs at SK Hynix, up from effectively zero in 2024. That share migration is what is showing in the bookings number.

The financial print

Q1 2026 (reported April 29) printed **revenue +32% YoY**, **bookings +72% YoY**, and the headline structural metric: SK Hynix HBM4 TC bonder order confirmed as taken from Hanmi. Q2 guidance is **$540-600 million in revenue, +37% YoY**, above consensus, with the bookings velocity suggesting Q3 and Q4 should sequentially accelerate. Operating margin trajectory is improving from mid-teens toward the 20s as the higher-ASP TC bonder mix flows through.

The stock closed **HKD 176.20 on May 17**. Q2 2026 reports **July 28, 2026**.

The bookings-to-revenue lag is the key metric to track. ASMPT's bookings +72% YoY at Q1 will convert to revenue over the subsequent two to four quarters depending on tool type and customer delivery schedule. That suggests Q3 and Q4 2026 revenue can run materially above current consensus, with full-year revenue plausibly $2.2-2.4 billion against consensus closer to $2.1 billion. The book-to-bill ratio at Q1 was running well above 1.0, which by itself is the predictive signal for sequential revenue acceleration. The Q2 print on July 28 should confirm whether the bookings acceleration is sustaining or whether Q1 was a one-time pull-forward.

Customer mix

Customer mix has flipped favourably. SK Hynix is now the principal HBM customer (the share-take from Hanmi), Micron is a smaller but ramping account, and Samsung remains a longer-tail option pending HBM4 qualification. The non-HBM business — SMT lines for general electronics assembly, wire bonders for legacy packaging — provides diversification and represents probably 40-50% of the total revenue line. Per Trendforce's December 2025 reporting, ASMPT was operating roughly half of the active HBM4 TC bonders at SK Hynix's M14/M16 fabs by year-end, up

from effectively zero in early 2024 — the share-shift velocity has been faster than any equipment vendor share-shift in recent memory.

The customer book in HBM is structurally improving. Once an HBM line is built out with ASMPT TCBs, the recurring revenue (consumables, service, upgrades) plus the next-generation tool replacement cycle creates an annuity. The customer mix is moving toward higher-quality (HBM, advanced packaging) and away from cyclical commodity assembly.

The geographic distribution of the customer base is part of the strategic story. ASMPT is Hong Kong-listed with Singapore operational headquarters and engineering footprint in multiple Asian semiconductor hubs. This geopolitically-distributed profile is itself an advantage in a world where US-China semiconductor decoupling continues — ASMPT's tools are not subject to US export controls in the same way Hanmi's Korean-origin tools or BESI's European-origin tools might be, and the Hong Kong listing gives the equity a regulatory profile that is distinct from Korean or European peers. As capital allocators increasingly think about supply-chain resilience, that geographic distribution becomes a small but meaningful positive feature of the ASMPT exposure.

Competitive context

The TC bonder competitive set is narrow: Hanmi (#1, ~71% share globally per TechInsights 2025), ASMPT (#2, share rising), Hanwha Semitech (#3, Korean local share-taker, small base), and the longer tail of Japanese assembly-equipment vendors. ASMPT's strategic advantage is the Hong Kong corporate structure (geopolitically more neutral than Korean), the broader back-end portfolio (which is a cost-of-ownership argument for fabs that want one-vendor consolidation), and the hybrid-bonding roadmap that runs through the same corporate parent as the TC bonder business. The Samsung-captive SEMES TCB programme is an internal subsidiary rather than an open-market competitor, so the addressable competitive market is effectively narrowed to the three commercial vendors plus the Japanese long tail.

Against BESI (the hybrid-bonding incumbent at TSMC for logic SoIC), ASMPT is the credible second in the hybrid-bonding race with a different customer footprint. Against Hanmi, ASMPT is the share-taker at the marginal HBM4 socket. The competitive context is structurally favourable in the medium term — ASMPT is moving up the share curve, not defending share.

The recent SK Hynix HBM4 wafer-bonder patent context is informative. SK Hynix filed and was granted patents in early 2026 covering specific wafer-bonding process steps that ASMPT's tool family is positioned to deliver. The implication is that SK Hynix has effectively co-developed elements of the ASMPT HBM4 production flow — a deeper engineering relationship than the typical customer-vendor arrangement and a moat that competitors cannot replicate quickly. The Hanmi share-loss at SK Hynix is therefore not principally a pricing or commercial-relationship issue; it is a process-integration depth issue that ASMPT has earned through the qualification cycle.

Terminal risk

The terminal risk is the same as for Hanmi but in reverse: hybrid bonding eventually displaces TC bonding for HBM5/HBM6 generations in the 2027-29 window. ASMPT has a hybrid-bonding roadmap through the same corporate parent that owns the TCB business, so the displacement risk is partial rather than total — but the value capture per HBM line in hybrid bonding may not match the value capture in TCB. The mitigation is that ASMPT's hybrid-bonding qualification

track at SK Hynix and Samsung is concurrent with BESI's, which positions the company for the transition rather than against it.

The secondary terminal risk is HBM cycle peak. ASMPT's bookings are running +72% YoY off a constrained base; the mathematics of base effects mean FY28-29 will deliver structurally lower growth even if absolute volumes hold. The terminal risk is positioning rather than fundamental.

A third structural risk is the non-HBM portion of ASMPT's business — the wire bonder, die-attach and SMT lines that account for roughly 40-50% of revenue. These lines are cyclical and exposed to broader semiconductor capex cycles, automotive electronics demand, and consumer-electronics assembly volumes. If global semiconductor capex outside HBM softens — which is a plausible 2027 scenario given the inventory positions at the end-customer level — the non-HBM revenue line could compress 10-15% in a single year, partially offsetting the HBM tool-order strength. The consolidated revenue trajectory depends on the HBM acceleration outpacing the non-HBM cyclical compression.

Bull case

The three-to-five-year bull case sees ASMPT capturing 40-50% of new HBM4 TC bonder sockets globally through 2027, FY27 revenue running HKD 22-25 billion versus FY26 closer to HKD 17 billion, operating margin expanding to the mid-20s, and EPS at HKD 6-8. A 25-30x multiple — TC bonder duopolist with hybrid-bonding optionality — gets the stock to HKD 200-250 over twelve months.

The stretch bull case adds hybrid bonding qualifying at SK Hynix for HBM5 in 2027 and ASMPT taking principal supplier position alongside BESI, plus continued share gains at Samsung HBM4 production. That path takes the stock to HKD 280-320 on a 24-month view.

A particular dynamic to watch is the Samsung HBM4 qualification path. Samsung has historically been TCB-supplied principally through its in-house SEMES subsidiary (~13% global TCB share) plus a marginal slice from Hanmi. If Samsung's HBM4 generation qualification at Nvidia clears in late 2026 and Samsung's HBM4 capacity expansion runs aggressively in 2027, Samsung will need external TCB capacity beyond what SEMES can supply — and ASMPT is the most likely beneficiary because of its broader process technology and its already-qualified position at the SK Hynix tier. That Samsung optionality is worth 10-15% of incremental revenue in the bull-case path; consensus has not yet modelled it.

Gap / bear case

What the market may be missing is the structural durability of the share gain. Once an HBM line is qualified on ASMPT TCBs, the cost of switching back to Hanmi (re-qualification, line downtime, process validation) is prohibitive — these design wins are sticky 5-7 year revenue annuities, not contestable quarterly orders. Consensus FY27-28 revenue probably understates this durability.

The bear case the market may be right about is the cycle peak — ASMPT's bookings inflection is concurrent with Hanmi's, BESI's, Lumentum's and the entire AI capex complex, and any blink in 2027 hits the whole sector simultaneously. Sizing has to respect that correlated downside.

A specific gap in consensus modelling is the operating-margin trajectory. ASMPT's historical margins have run in the 10-15% range through cycle averages; the current mid-teens margin reflects mix-shift toward HBM-class higher-ASP equipment. If the HBM mix continues to grow

as a share of revenue through 2027-28, the consolidated operating margin can structurally lift to the 20-25% range — a 500-1000 bp improvement that would materially re-rate the multiple. Most sell-side models do not yet bake in that mix-shift trajectory at full magnitude.

Optionality

Two options. First, hybrid bonding — ASMPT has the roadmap and the customer engagement, and an HBM5 qualification win would be a step-change re-rating. Second, Singapore corporate-restructuring optionality — the ASMPT parent structure has historically created sum-of-parts value that occasionally surfaces through dividend programs, spin-offs or strategic transactions; none of which is priced into consensus. A third option is the AI-3D-packaging logic-stacking adjacency — as TSMC, Samsung and Intel scale 3D logic packaging for AI accelerators, the TC bonder qualification at the logic-packaging tier becomes a separate revenue line that ASMPT could enter alongside its HBM-focused business; the cross-over from HBM to logic-packaging is a 24-36 month strategic opportunity that is not currently in consensus.

The trade

> **Entry:** HKD 165-185 — current zone acceptable; aggressive add on any 15% pullback to HKD 150. **Size:** 3% portfolio target (already owned 10K sh per v3 watchlist). **Stop:** HKD 125 (200-day MA / prior breakout retest). **Catalyst date:** Q2 print July 28, 2026. **Trim/exit:** trim 25% at HKD 230, 50% at HKD 270; full exit on HBM cycle peak in monthly Trendforce data. **Conviction: 8/10**.

Layer 4 · BESI — BE Semiconductor Industries

One-line thesis

BE Semiconductor is the hybrid-bonding incumbent at TSMC with bookings +104.5% YoY and Lam/AMAT takeover approaches on the table — own it as the next-cycle HBM5+ stacking technology winner with embedded M&A optionality and a TSMC-anchored SoIC revenue base that compounds independent of HBM cycle timing.

What BESI physically does

BE Semiconductor Industries makes back-end assembly equipment, and specifically the hybrid-bonding tools that are positioning to displace thermo-compression bonding for the HBM5/HBM6 generation and that already dominate the leading-edge logic stacking market at TSMC. Hybrid bonding is the direct copper-to-copper die bonding process that eliminates the µbump entirely — copper pads on adjacent dies are bonded directly with no solder, no underfill, no thermal-cycle bottleneck. The pitch shrinks from 25-40 µm (TC bonding) to 1-5 µm (hybrid bonding), enabling 20+ Hi stacks and dramatically lower power per bit.

The AI-stack relevance is the next-generation HBM. Current HBM3E and HBM4 use TC bonding (Hanmi and ASMPT). HBM5/HBM6 — targeting commercial production around 2027-28 — is expected to use hybrid bonding for the higher die counts and the tighter power envelope. BESI is the incumbent technology owner via the TSMC SoIC partnership (3D logic packaging, currently in production for Apple M-series and forthcoming AI ASICs) and is the principal candidate for

HBM5 qualification at SK Hynix and Samsung. The Q1 26 bookings inflection — +104.5% YoY at €269.7 million — is the signal that the HBM5 capacity build is starting.

Physically, the hybrid bonder is a precision-alignment tool that does sub-micron die placement followed by a low-temperature anneal step for the Cu-Cu interdiffusion bond. It is a different physical product from a TC bonder (cleaner room class, different chemistry, different metrology) and the moat is the qualification track at TSMC and the integration with the upstream wafer-thinning and TSV processing equipment. BESI's installed base at TSMC for logic SoIC gives it the qualification incumbency that competitors do not have.

The TSMC SoIC programme is the under-appreciated revenue anchor that sits inside BESI's story but does not get much coverage outside specialist semiconductor research. SoIC — System on Integrated Chips — is TSMC's 3D logic packaging technology that stacks compute dies vertically with hybrid bonding, used for Apple M-series processors (front-and-back die stacking) and forthcoming AMD MI400 and Nvidia 3D-packaged AI accelerators. Every SoIC stack requires multiple hybrid-bonding operations on BESI tools, and the SoIC volume is expanding 25-30% CAGR through 2028 independent of HBM5 timing. This logic-packaging revenue line is more visible and shorter-cycle than the HBM5 line, and it is what gives BESI a quality of earnings even before HBM5 ramps.

The financial print

Q1 2026 (reported April 23) printed in-line revenue, beat EPS, and the structural number: **bookings €269.7 million, +104.5% YoY**. Q2 guide is **+30-40% QoQ** — a sequential acceleration that confirms the hybrid-bonding capex cycle is ramping. Operating margin runs in the high-20s structurally and is trending up as the hybrid-bonding mix grows.

The stock closed **€261.80 on May 17**. The Lam Research and Applied Materials takeover approaches reported May 12 are the market-positioning event of the quarter — neither has filed a formal bid but both have reportedly approached BESI's management about strategic combinations. Q2 2026 reports **July 23, 2026**.

The bookings doubling is the cleanest signal of the hybrid-bonding capex cycle starting in earnest. Through 2023-25 BESI's bookings ran in the €130-150 million per quarter range, reflecting normal cyclical back-end equipment demand. The Q1 2026 print at €269.7 million represents a step-function move that aligns with the timing of HBM5 qualification programmes at SK Hynix and Samsung. If the Q2 +30-40% QoQ guide is delivered, bookings reach €350-380 million for a single quarter — annualised that approaches €1.4-1.5 billion of bookings, which materially exceeds current consensus FY26 revenue guidance and implies sell-side estimates are walking up through every print cycle.

Customer mix

Customer mix is more diversified than Hanmi's or ASMPT's. TSMC is the principal hybrid-bonding customer through the SoIC programme — Apple M-series, forthcoming Nvidia 3D-packaged AI silicon, AMD MI400 3D-stack — and represents a meaningful slice of revenue. The HBM customers (SK Hynix, Samsung, Micron) are emerging customers as hybrid bonding qualifies for HBM5. Beyond AI, BESI has a broad back-end equipment portfolio across automotive, industrial and consumer semiconductor assembly that provides diversification. The broad legacy back-end equipment book provides earnings ballast through cyclical periods and supports operating-

margin stability even when the leading-edge advanced-packaging revenue lines have volume air pockets. The geographic distribution of customers — Taiwan, Korea, Japan, US, Europe — creates a globally-balanced revenue profile that compares favourably to the Korea-concentrated TC bonder competitive set.

The hybrid-bonding customer book is structurally improving. TSMC SoIC orders are recurring; HBM5 qualification at SK Hynix and Samsung adds incremental revenue lines that are currently zero. The customer mix migration is from cyclical back-end equipment toward strategic-bottleneck advanced packaging.

The TSMC-anchored customer book is uniquely valuable for two reasons. First, TSMC's process roadmap is the closest thing to a guaranteed long-term volume curve in the semiconductor industry — TSMC has visibility on customer demand and capacity planning that propagates into BESI's order book on a multi-year basis. Second, the SoIC qualification at TSMC is fungible across multiple downstream customers (Apple, Nvidia, AMD), which means BESI's revenue at TSMC scales with the broad AI accelerator and consumer-silicon demand curve rather than any single customer's product cycle. The TSMC anchor is the structural quality-of-earnings differentiator versus pure HBM-equipment plays like Hanmi.

Competitive context

The hybrid-bonding competitive set is narrow: BESI (the incumbent at TSMC), ASMPT (the credible second with concurrent roadmap), EV Group (the Austrian wafer-bonding specialist, smaller scale), and the longer tail of research-grade tool makers. Hanmi's hybrid-bonding programme is structurally late. Hanwha Semitech is shipping hybrid-bonding sample equipment to SK Hynix but is small-base.

BESI's moat is the TSMC SoIC qualification incumbency (which competitors cannot easily replicate), the engineering depth on the Cu-Cu bonding process (15 years of R&D), and the European corporate-tax structure that provides operating-margin advantage. Against ASMPT, BESI has the higher-quality customer book and the more advanced technology position; against EV Group, BESI has scale.

The strategic acquirer logic is also worth tagging because it is the M&A optionality leg. For Lam Research, BESI fills a gap in the back-end advanced packaging portfolio — Lam is a deposition and etch specialist with limited assembly-equipment exposure, and the hybrid-bonding adjacency is technologically aligned with Lam's existing TSV-etch business. For Applied Materials, BESI complements the company's existing back-end packaging programme (AMAT has a hybrid-bonding R&D effort but is structurally behind BESI in the TSMC qualification track) and would accelerate AMAT's positioning for HBM5. The acquisition logic is sound on both sides; the question is whether either Lam or AMAT will commit the $20-25 billion equity-value transaction that BESI's current valuation would imply.

Terminal risk

The terminal risk is the timing of the hybrid-bonding transition. If HBM5 is delayed beyond 2027-28 — which is possible if TC bonding can be extended for 20-Hi stacks via new die-thinning techniques — BESI's hybrid-bonding revenue ramp is pushed out by 12-24 months. The mitigation is that the TSMC SoIC business is already ramping for logic, so BESI is not entirely dependent on HBM5 timing; but the upside scenario requires HBM5 to happen on schedule.

The secondary terminal risk is the M&A speculation reversing. If neither Lam Research nor Applied Materials proceeds with a formal bid, the takeover-premium component of the current price compresses and the stock gives back 15-25% on the unwind. The mitigation is that the operating story is independent of M&A — the bookings inflection is the structural anchor.

A third terminal risk is European competitive entry — specifically the risk that ASM International (the lithography-adjacent Dutch equipment vendor) or another European semiconductor-equipment specialist enters the hybrid-bonding tool market through a partnership or acquisition of one of the smaller hybrid-bonding R&D players (EV Group, SUSS MicroTec). The probability is low because the customer-qualification cycle at TSMC and SK Hynix is multi-year and BESI's lead is structural, but the geographic concentration of hybrid-bonding R&D in Europe makes a competitive entry slightly easier than for the Asian-headquartered TC bonder market.

Bull case

The three-to-five-year bull case sees BESI capturing principal hybrid-bonding share at HBM5 qualification at SK Hynix and Samsung, TSMC SoIC volume expanding through the AI-3D-packaging ramp, FY28 revenue running €1.2-1.5 billion versus FY26 closer to €750 million, operating margin expanding to the low-30s. EPS in the €12-14 range. A 25-30x multiple — hybrid-bonding monopolist plus M&A optionality — takes the stock to €350-400 over 24 months.

The stretch bull case adds a formal Lam Research or AMAT acquisition at a 30-40% premium — €340-365 take-out — which is a discrete event in the next 12-18 months. That path is binary but real.

The hybrid-bonding TAM math is what makes this a structurally larger opportunity than the TCB cycle. A single 12-Hi HBM line at SK Hynix requires roughly $200-300 million of TC bonder capex; an equivalent hybrid-bonding line for HBM5 16-Hi or HBM6 20-Hi requires $350-500 million of equipment capex because the tool ASPs are higher and the throughput is lower per tool. The transition from TCB to hybrid bonding therefore expands the addressable equipment TAM by roughly 70-100% per HBM line. If hybrid bonding becomes the standard for the 2028-2030 HBM generations, the TAM for back-end stacking equipment doubles structurally — and BESI is the principal share-holder going into the transition.

Gap / bear case

What the market may be missing is the durability of the TSMC SoIC ramp independent of HBM. The logic-3D-packaging business — Apple M-series, AMD MI400, Nvidia next-gen — is itself a structural growth driver that consensus probably under-weights because the HBM narrative dominates the headlines. The hidden upside is the SoIC business growing 25-30% CAGR through 2028 even before HBM5 contributes.

The bear case the market may be right about is the timing risk. Hybrid bonding for HBM has been "next year's technology" for three years now, and SK Hynix has explicitly said HBM4 16-Hi will use TC bonding (MR-MUF) rather than hybrid bonding. If the technology generation slips one more cycle, BESI's revenue ramp is delayed and the multiple compresses against a flatter near-term earnings curve.

A specific timing watchlist item: the SK Hynix HBM5 process technology disclosure cadence. SK Hynix has typically disclosed forward HBM roadmap at SEMICON Korea (February) and at internal capital markets days. Watching for explicit HBM5 hybrid-bonding qualification timing

at the 2027 disclosure cycles will be the gating signal for BESI's earnings inflection. A delay or technology choice announcement that favours TC bonding over hybrid bonding for HBM5 would be a 20-30% downside trigger for the BESI stock; a clean hybrid-bonding qualification commitment would be a comparable upside trigger.

Optionality

Three options. First, the M&A approach itself — Lam Research and Applied Materials are credible strategic acquirers and a formal bid at 30-40% premium is plausible inside the next 18 months. Second, the SoIC logic-packaging business — broader than HBM, less visible to consensus, structurally growing. Third, the European corporate-tax structure provides operating-margin advantage that flows directly to FCF and supports capital-return optionality. A fourth option worth tagging is the chiplet ecosystem — as AMD, Intel, and Nvidia all migrate to chiplet-based architectures for next-generation accelerators, the inter-die hybrid-bonding step becomes a recurring revenue line; BESI is the only equipment vendor positioned across both HBM-stacking and logic-chiplet hybrid bonding, which is a structurally unique competitive position.

The trade

> **Entry:** €245-275 — current zone acceptable; aggressive add below €220. **Size:** 3-4% portfolio target (already owned per current portfolio). **Stop:** €185 (200-day MA / prior breakout). **Catalyst date:** Q2 print July 23, 2026; M&A news flow continuous. **Trim/exit:** trim 25% at €340, 50% at €380 (formal M&A bid zone); full exit on hybrid-bonding HBM5 delay or M&A approach unwind. **Conviction: 9/10**.

Layer 7 · GEV — GE Vernova

One-line thesis

GE Vernova is the power-infrastructure pure-play with $163 billion backlog, +71% organic order growth and FY26 guide raised — own it as the AI-power-bottleneck call that consensus is still under-modelling versus the gas-turbine slot reservation pipeline.

What GE Vernova physically does

GE Vernova is the spun-out power-equipment arm of the old GE — gas turbines, steam turbines, wind turbines, grid-electrification equipment (transformers, switchgear, HVDC converters), and the digital/services layer that supports the installed base. Inside the AI-stack story, the relevant businesses are the gas turbine business (which is supplying the new combined-cycle gas plants being built specifically for hyperscaler data centre power) and the Electrification segment (which is supplying transformers, switchgear and grid equipment for the data centre interconnection build-out).

The AI-stack layer is Layer 7 — power infrastructure — and it is the deepest binding constraint in the entire AI capex stack right now. You can have all the GPUs, HBM, transceivers and substrates you want; if you cannot get a 100 MW grid connection and the gas turbine slot to back it up, the data centre does not run. The US grid interconnection queue is now backed up to 2029-2031 in most ISO regions. Gas turbine lead times have stretched from 18 months in 2022 to 4-5 years in

2026. GEV is the principal global supplier of utility-scale gas turbines and is one of the few names where the supply-side constraint is literally physical capacity rather than financial willingness.

The Q1 26 disclosure that **Electrification booked $2.4 billion in data centre orders in a single quarter — more than the full year 2025** captures the inflection. The data centre power demand curve has gone vertical, and GEV's order book is the financial signature of that curve.

The financial print

Q1 2026 (reported April 22) printed **revenue $9.34 billion**, EPS consensus beat, **orders +71% organic YoY**, **backlog $163 billion (+$13 billion QoQ)**, and management raised FY26 revenue guide to **$44.5-45.5 billion** and FCF guide to **$6.5-7.5 billion**. The Electrification segment's $2.4 billion data-centre order print in Q1 alone was the headline number that re-rated sell-side models.

The backlog math is structural. **$163 billion versus $116 billion at the spin** (less than two years ago) is a 41% backlog expansion in 18 months, on a business where backlog converts to revenue over 2-5 years depending on segment. Gas turbine slot reservations are now being booked through 2029-2030 and management has guided to **110 GW of slot reservations targeted by year-end 2026**. The stock closed **$1,049.23 on May 17**, up roughly **+145% one-year total return**. Q2 2026 reports **July 22**.

The orders +71% organic YoY metric deserves separate attention. Organic order growth at that rate, on a $40+ billion revenue base, is unprecedented in modern utility-equipment history. The closest analogue is probably the early-2000s combined-cycle gas-turbine cycle that lifted GE Power orders by 50-60% YoY for two consecutive years. That cycle ran for roughly five years before saturating; the current AI-driven order cycle has plausible duration through 2030-32 based on grid-interconnection queue depth and gas-turbine slot-reservation forward bookings. The historical comparison suggests the order book has another two to three years of structural expansion ahead before mean reversion sets in.

Customer mix

Customer mix is unique among the Tier A names — it is principally utility companies (Duke, NextEra, Southern, Entergy, Dominion, plus the Canadian, European and Middle Eastern utility customers), independent power producers, and the hyperscalers' own captive-power vehicles. The customer book is also unique in that the contracting is long-dated (5-15 year purchase agreements with deposit structures) and the cancellation costs are prohibitive, which makes the backlog quality very high.

The Middle Eastern utility customer base is increasingly material to the GEV story. Saudi Arabia's National Center for AI Capability (NCAC) is committing tens of billions of dollars to data centre build-out across the kingdom, paired with new gas-turbine generation to support the load. UAE's G42 and the broader Mubadala AI infrastructure programmes are doing the same at slightly smaller scale. GEV is a principal supplier to both regions through long-standing utility relationships with SEC (Saudi Electricity) and EWEC (UAE). The Middle Eastern AI-power capex is an under-appreciated incremental driver because most US-focused sell-side models do not properly capture the regional contribution to GEV's order book.

The hyperscaler-driven demand is funneled through three channels: direct captive-power purchases (Amazon's nuclear+gas+renewables ladder, Microsoft's Three Mile Island restart deal, Google's geothermal+SMR programmes), utility-led data-centre interconnection (where the utility procures GEV equipment on behalf of the hyperscaler load), and IPP-driven (where merchant generators like Vistra, Constellation, Calpine procure GEV equipment to serve hyperscaler PPAs). All three channels are accelerating, and the second-and-third channels are arguably even more durable than the first because the utility and IPP customer base is regulated and capital-disciplined in a way that hyperscaler capex is not.

The customer-quality angle is unusual for an AI-themed name. Unlike GPU/HBM/transceiver customers — which are concentrated in a handful of hyperscalers and a smaller tier of AI-cloud merchants — GEV's customer base is broadly diversified across hundreds of utility, IPP and industrial accounts globally. The hyperscaler-driven incremental demand is layered on top of a structural utility-replacement base that does not depend on AI capex continuing. Even in a hypothetical scenario where AI data centre demand collapses entirely, the gas-turbine fleet replacement cycle alone supports 15-18 GW/year of new build globally; the AI-driven incremental is what takes the curve to 25-30+ GW/year. The downside cushion is structurally larger than for any other Tier A name.

Competitive context

The competitive set in utility-scale gas turbines is a global oligopoly: GEV, Siemens Energy, Mitsubishi Power, and to a lesser extent Ansaldo Energia. The four together supply >95% of new utility-scale gas-turbine capacity globally. GEV has roughly 35-40% global market share and is the technology leader in the HA-class (high-efficiency advanced) gas-turbine segment that dominates new combined-cycle builds. Siemens Energy is the credible #2 and has been ramping its order book concurrently.

The competitive moat is the installed base (GEV's installed turbines globally generate ~30% of world electricity, providing decades of service revenue), the technology depth on combustion efficiency, and the manufacturing scale at the Greenville, SC and Belfort, France factories that competitors cannot easily match. The competitive context is structurally favourable through the AI-power-bottleneck cycle and probably extends through the next decade given the grid build-out requirements.

The Electrification segment is the second-engine story that gets less attention than the gas-turbine line but is structurally just as important. Transformers, switchgear, HVDC converters and grid-interconnection equipment are the second tier of the AI-power-bottleneck — even after you have the generation capacity, you need the equipment to get the power from the substation into the data centre. The market for utility-grade transformers and switchgear has tightened to the point where lead times for major substation equipment are now 36-48 months, up from 12-18 months pre-pandemic. GEV is the principal global supplier of this equipment alongside Siemens, ABB, Schneider Electric and Hitachi Energy. The $2.4 billion of data-centre orders booked in Q1 2026 alone — more than the entire 2025 — captures the Electrification inflection.

Terminal risk

The terminal risk is the AI capex blink translated into the slowest-moving part of the supply chain. If hyperscaler capex blinks in 2027-28, the immediate effect is GPU and memory orders rephased

— the gas turbine orders, with multi-year lead times and committed deposits, are stickier. But over a 3-5 year window, sustained AI capex moderation would cool the gas-turbine order book and the +71% organic growth would mean-revert to the 8-12% structural utility CAGR.

The secondary terminal risk is the execution risk on slot oversubscription. **110 GW of slot reservations by year-end 2026** is a massive scaling-up for a manufacturing footprint that delivered 15-20 GW/year historically. Quality issues, capex bottlenecks, or supply-chain failures during the ramp would erode margin and damage customer relationships. Management has flagged this risk; it is real but historically GE Power has navigated similar ramps reasonably well.

A third terminal risk is regulatory and policy disruption. Gas-turbine combined-cycle plants are increasingly contested at the state and federal regulatory level — particularly in California, the Pacific Northwest, and the EU — where emissions rules and decarbonisation mandates could constrain new-build approvals. If regulatory friction increases meaningfully in 2027-28, the gas-turbine order book could face slot cancellations even if AI demand is intact. The mitigation is that the bulk of GEV's gas-turbine order book is in the US South-East, Texas, and the Middle East — geographies with materially more permissive regulatory profiles.

Bull case

The three-to-five-year bull case sees GEV converting the $163 billion backlog into $50+ billion of annual revenue by FY28, gas-turbine deliveries running 25-30 GW/year by 2028, Electrification revenue running $18-20 billion (versus FY26 closer to $13 billion), and operating margin expanding to the mid-teens consolidated. FCF in the $10-12 billion range. EPS in the $30-35 range. A 30x multiple — power-infrastructure monopolist with secular AI bottleneck moat — gets the stock to $900-1,050 on a 24-month view and $1,300-1,500 on a 36-month view.

The stretch bull case adds the small-modular-reactor (SMR) joint venture programmes maturing — GEV has positioned through BWXT and other strategic partnerships on the next-generation nuclear technology — plus the European grid build-out (REPowerEU, German Energiewende) layering on incremental orders independent of US AI demand. That path takes the stock to $1,800+ over five years.

The services revenue compounding is the often-overlooked third leg of the bull case. GEV's installed-base service business generates 35-40% operating margins on recurring multi-year contracts and grows at 6-8% CAGR independent of new-build orders. As the new-build cycle peaks and rolls off, the services line continues to compound, providing earnings durability and a high-quality-of-earnings anchor that pure capital-equipment names lack. By 2030, services revenue could be running $12-15 billion against $4-5 billion of operating income from services alone — a meaningful slice of consolidated earnings that does not depend on AI capex.

Gap / bear case

What the market may be missing is the durability of the backlog conversion. Gas-turbine deliveries through 2030 are essentially price-locked through the existing slot reservations and the deposit structures make cancellation prohibitive. Consensus FY27-28 revenue at $48-50 billion probably understates the conversion velocity if the manufacturing ramp executes.

The bear case the market may be under-pricing is the +145% one-year valuation expansion. The stock has moved from a re-rating story to a fully-priced compound story, and at ~30x EBITDA on consensus 2026 the entry-point asymmetry is less favourable than it was 12 months ago. Price

has softened from $1,275 to $1,049 over the past eight sessions on broader power-sector profit-taking — the operating story is durable but the technical positioning is heavy, and unlike CEG and VST there is no STRONG_EXIT signal flagged in the May 11 daily brief for GEV, so this remains a buy-the-pullback setup rather than a tactical-downgrade setup.

A specific bear-case scenario worth modelling is a wind-turbine segment write-down. GEV's wind business has been a periodic drag on consolidated earnings, with offshore wind projects in particular running over budget and behind schedule. While management has improved execution materially since the spin, a residual project write-down at any point through 2026-27 would hit earnings even as the gas-turbine and Electrification segments continue to ramp. The wind-business cleanup is the housekeeping item that could surprise to the downside in any given quarter.

Optionality

Four options. First, SMR — small modular reactors are the long-tail nuclear option that GEV has positioned through partnerships; if SMR clears regulatory hurdles in 2027-29, GEV is positioned. Second, European grid build-out — REPowerEU and German Energiewende create a second-engine demand source independent of US AI. Third, services revenue compounding — the installed base produces high-margin recurring service revenue that compounds at 6-8% CAGR through cycle and provides downside cushion. Fourth, capital return acceleration — at FCF of $6.5-7.5 billion and a deleveraging balance sheet, GEV has room to accelerate buybacks or initiate a special dividend; none currently priced. A fifth optionality leg is hydrogen-enabled gas turbines: GEV's HA-class turbines are increasingly designed for hydrogen co-firing, which positions the installed base for decarbonisation retrofits across the next decade — a structural revenue line that would extend the post-AI cycle without depending on new-build demand.

The trade

> **Entry:** $1,000-1,080 — current zone acceptable but watch the technical base; aggressive add below $920. **Size:** 4-5% portfolio target. **Stop:** $820 (prior consolidation low / 200-day MA). **Catalyst date:** Q2 print July 22, 2026; slot reservation milestones quarterly. **Trim/exit:** trim 25% at $1,350, 50% at $1,500; full exit on hyperscaler capex blink or 110-GW slot target shortfall. **Conviction: 9/10**.

Part 1 portfolio synthesis

Nine names. Three layers. One thesis: the AI silicon supply chain remains the structurally tightest physical-asset bottleneck in any global equity market, and the price-action through May 18 2026 has compressed entry-level R/R but has not changed the underlying narrative.

Layer 1 (silicon design). NVDA, AVGO, MRVL sit at the apex of the stack. The combined target weight is 8-11% of NLV across the three names, with NVDA carrying the largest single allocation (4-5%) given the platform-monopoly thesis, AVGO second (3-4%) on the diversified custom-silicon + networking franchise plus the software cash-flow ballast, and MRVL third (1.5-2.5%) on the higher single-customer-concentration risk and the +135% momentum that argues for tactical

entry discipline. The triple owns the two non-NVIDIA layers of credible AI accelerator silicon plus the broader networking and electro-optics economics.

Layer 2 (foundry and equipment). TSM, KLAC, LRCX, CAMT, ONTO sit at the manufacturing tier. Combined target weight 11-15% of NLV. TSM is the largest single allocation (4-5%) — the only sub-5nm logic foundry that can absorb AI silicon volume. KLAC + LRCX is the equipment-pair-trade at 5-6% combined (2.5-3.5% KLAC, 2.5-3% LRCX) covering inspection / metrology and etch / deposition respectively. CAMT + ONTO is the bump-inspection duopoly pair at 3-4% combined (1.5-2.5% each), with ONTO carrying the larger valuation cushion to consensus PT and CAMT carrying the cleaner HBM-pure-play exposure.

Layer 3 (memory). 000660 (SK Hynix) is sized at 4-5% of NLV as the HBM monopoly position. Paired with MU (covered in Part 2 of this Tier A series) the combined HBM exposure runs 7-9% of NLV. The Korean equity carries a structural geopolitical / governance discount versus the US-listed analogue, which is the reason the pair is sized for diversified HBM exposure rather than single-name concentration.

Aggregate Tier A Part 1 weight: 23-31% of NLV across the nine names. This is a meaningfully concentrated AI silicon supply-chain expression. The principal cycle-risk variables to monitor in priority order: (1) NVDA Q1 FY27 print May 20 (single biggest market event), (2) hyperscaler capex-to-OCF ratios into Q3 26, (3) SK Hynix HBM ASP trajectory through 2027 LTAs, (4) Samsung HBM4 qualification status at NVIDIA, (5) TSMC monthly revenue trajectory and the Taiwan-strait geopolitical context, (6) China export-restriction policy direction through 2026.

The single biggest known-unknown across all nine names is the agentic-AI transition. If hyperscaler inference workloads scale 10-100x per user interaction as multi-step agentic-AI frameworks displace one-shot LLM serving, the demand curve for every name in this Tier A list shifts higher by 30-50% on a multi-year basis. If the transition stalls — if hallucination, reliability, or economic-cost ceilings prevent agentic-AI from achieving production-ready scale — the entire AI capex cycle is in the third-inning rather than the second-inning of growth. The base case assumes the transition is underway and is one of the principal drivers of the consensus FY27-28 EPS revisions across this cohort.

Print-to-print roadmap, May 18 - August 30, 2026.

Date	Name	Event
May 20	NVDA	Q1 FY27 print after-hours
May 27	MRVL	Q1 FY27 print after-hours
June 3	AVGO	Q2 FY26 print after-hours
June 25	MU	Q3 FY26 print (Part 2)
July 17	TSM	Q2 print
July 22	LRCX	Q4 FY26 print
July 24	000660	Q2 print
July 30	KLAC	Q4 FY26 print
August 6	ONTO	Q2 print
August 11	CAMT	Q2 print
Late August	NVDA	Q2 FY27 print

Two months of binary events across the Part 1 portfolio. The cleanest single signal — the one to

position against if you can only act on one — is the NVDA May 20 print, both because the size of NVDA's exposure dominates the cycle and because the read-through to every other Layer-1 and Layer-2 name in this cohort is unambiguous.

Part 4 — Tier B · Add on Pullback / Hold

Sorted by AI stack layer ascending; within each layer by conviction rank. 49 names. Credible AI exposure with valuation or moat caveats — sized smaller, added on weakness.

Layer 0 · ANET — Arista Networks

Thesis

The cleanest pure-play on back-end AI Ethernet — Meta, Microsoft, Oracle migrating from InfiniBand to 800G/1.6T Ethernet fabrics. Tier B not Tier A because a -16% post-print sell-off ($170 to $142) told us the crowd is now selling rallies on hyperscaler-timing fears, even as the print itself was a clean double-beat.

What it does + financial print

Arista builds the high-radix data-centre switches that knit together AI clusters — Tomahawk-5 and Jericho3 silicon, EOS software, Etherlink for back-end RoCE. The competitive race is binary: either AI back-ends standardise on Ethernet or NVDA's Spectrum-X eats the fabric. Right now both vendors win, but Arista is the share-leader inside the Magnificent-Seven hyperscaler footprint.

Q1 2026 printed $2.71B revenue versus $2.66B consensus — a +35% YoY beat — and $0.87 EPS versus $0.79 (+10%). The guide was technically in-line but management leaned cautious on hyperscaler timing, which is what drove the -16% reaction. Sell-side average PT sits at $182 (FX Leaders, 25-35% upside). The fundamentals are fine; the tape is suspicious.

Bull case

A Meta/MSFT 800G order acceleration in H2 — even one disclosed mega-PO — flushes the digestion fear and rerates the multiple back toward the $170-180 zone. Add Oracle, xAI and the sovereign deals (KSA, UAE) ramping their own Ethernet back-ends and you have the same revenue trajectory ANET had through 2025 but on a wider customer base.

Gap / bear case

Two bears co-exist: AVGO's Tomahawk-6 silicon means more hyperscalers can build their own switches and skip Arista's premium box; NVDA's Spectrum-X bundles networking with GPUs and is the path-of-least-resistance for Tier-2 buyers. Until ANET prints a quarter where front-end + back-end Ethernet wins are both visible, the multiple stays capped.

Trigger to upgrade / downgrade

Upgrade to Tier A on a clean $130 reset that holds (the $130 zone is the prior gap-fill and 200-day moving average) — or on the next print if back-end AI Ethernet revenue is explicitly disclosed at >$1B annualised. Downgrade to C only if a hyperscaler publicly defects to Spectrum-X for a tier-1 cluster.

The trade

- ***Entry zone***: *$128-132 on pullback (gap fill); chase only above $158*
- ***Stop***: *$118 close (2×ATR below entry, breaks 200-day)*
- ***Position size***: *1.5% of NLV starter, 3% full*
- ***Catalyst date***: *Q2 print August 5 2026; hyperscaler Computex-style commentary mid-summer*
- ***Conviction***: *7/10*

Layer 0 · CSCO — Cisco Systems

Thesis

The unexpected mega-cap AI-networking rerate — Cisco raised its FY26 AI infrastructure orders target from $5B to $9B mid-year, a doubling of the framework metric for the whole thesis. Tier B because the stock is already +21% in 3 months and +14% in the last 30 days, so the easy money is gone — but the moment is real.

What it does + financial print

Cisco is the legacy enterprise networking franchise being repositioned around three AI-era vectors: Silicon One ASICs winning hyperscaler design slots, Splunk's observability stack feeding AI ops, and Nexus 9000/HyperFabric capturing the long tail of enterprise AI clusters that don't buy Arista. The bear case for a decade was "Cisco is dying" — the bull case now is "Cisco has just bought back into the AI conversation."

Q3 FY26 printed $15.8B revenue (record, +12% YoY) versus $15.0B consensus — and $1.06 EPS versus $1.00. FY26 revenue guided to $62.8-63.0B. The headline number was the AI orders target raised from $5B to $9B for FY26. Morgan Stanley raised its PT post-print. The stock gapped +13% AH.

Bull case

If Cisco crosses $10B AI orders for FY26 — and the trajectory through Q3 says they will — the narrative flips from "old-tech rebound" to "second AI networking franchise" and the multiple expands from low-20s P/E toward mid-20s. Splunk-AI synergy disclosures and a tier-1 hyperscaler Silicon One design win would unlock Tier A.

Gap / bear case

The stock has already run from $58 to $99 — the rerate is partially priced. Mid-cycle AI capex anxieties (the same ones bruising ANET) apply here too. A miss on the Q4 print, or an AI orders

number that comes in below the raised $9B framework, would be punished asymmetrically given the new sell-side enthusiasm.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) a pullback to $88-92 that holds the 50-day, OR (b) the Q4 FY26 print on August 13 confirming the $9B AI orders run-rate with FY27 colour suggesting $12-15B. Downgrade only on an AI orders disappointment.

The trade

- ***Entry zone**: $88-92 on consolidation; tolerate up to $96 for a starter*
- ***Stop**: $82 close (below 50-day + prior breakout shelf)*
- ***Position size**: 2% NLV*
- ***Catalyst date**: Q4 FY26 print August 13 2026*
- ***Conviction**: 7/10*

Layer 0 · DELL — Dell Technologies

Thesis

Dell is the dominant AI-server ODM-with-brand — xAI Colossus, CoreWeave fleet build-outs, sovereign deployments. Tier B because the next print is May 28 (consensus +112% EPS YoY) and the stock is already +121% over 12 months with a UBS downgrade in the bag — too much beta to a single-day print to add fresh capital here.

What it does + financial print

Dell builds the boxes. PowerEdge XE9680 / XE9712 / XE9740 servers fitted with B200/B300/GB200 NVL72 racks ship to xAI, CoreWeave, Oracle, and a long roster of Tier-2 GPU clouds. AI server backlog disclosed last quarter ran into the high-double-digit billions. The bear case has always been margins (these are commodity boxes); the bull case is that customer concentration in xAI/CoreWeave gives Dell pricing power as long as those two are the swing buyers.

Q4 FY26 (reported February 27) was a beat-and-raise on AI server backlog. May 28 prints Q1 FY27 with consensus EPS at +112% YoY. JPM raised PT to $280 (from $205) on May 15. UBS downgraded to Neutral on May 11 — both reactions tell you positioning is now two-sided.

Bull case

A clean $30-35B AI server backlog disclosure on May 28 — combined with margin commentary that says ODM competition is not eroding gross margins as feared — moves the stock to $280+ and earns Tier A. The sovereign / xAI pull-through is real; Dell is the default vendor for non-hyperscaler buyers who want the GB200 NVL72 yesterday.

Gap / bear case

ODM competition (Supermicro, Hon Hai, Quanta) compresses Dell's margin franchise as the AI server cycle matures. The +121% TTM run leaves no margin for an in-line print. UBS-style downgrades will multiply if Q1 prints anything less than +30% AI server revenue sequential growth.

Trigger to upgrade / downgrade

Upgrade to Tier A on a beat-and-raise May 28 print where AI server margins are flat-to-up sequentially. Downgrade to C if AI server margins compress >200 bp and backlog growth decelerates.

The trade

- ***Entry zone****: $215-225 on pre-print pullback OR $260+ chase on confirmed beat-raise*
- ***Stop****: $198 close (50-day support)*
- ***Position size****: 2% NLV starter*
- ***Catalyst date****: Q1 FY27 print May 28 2026*
- ***Conviction****: 7/10*

Layer 0 · IREN — IREN Limited

Thesis

The Bitcoin-miner-pivoting-to-AI-cloud story that just signed a 5-year NVDA infrastructure partnership and is targeting $3.7B ARR by end-CY26. Tier B because the Q1 EPS missed materially (-$0.33 vs -$0.24 est), Bitcoin revenue is decaying, and execution risk on the back-end-weighted ARR ramp is real — but the partnership signature is structural and validates the thesis.

What it does + financial print

IREN owns and operates renewable-powered data-centre capacity at Childress (TX), Sweetwater (TX), and West Texas — originally Bitcoin mining campuses being converted to GPU-as-a-service for AI inference and training. The NVDA partnership locks supply preference and reference-architecture credibility. Microsoft Blackwell deployments are now in flight. The model: convert depreciating Bitcoin economics into multi-year ARR contracts at much higher margins.

Q3 FY26 (May 7) printed $145M revenue (-22% QoQ on Bitcoin price decline) with AI cloud revenue ~$34M, doubling QoQ. EPS missed at -$0.33 vs -$0.24 consensus. The $3.7B ARR target by end-CY26 was reaffirmed but is back-end-weighted. Stock is +33.8% in the last month on the NVDA news and +89% TTM.

Bull case

ARR crosses $1B exit run-rate in Q4 FY26 — that's the visibility checkpoint that confirms the back-end-weighted plan is achievable. Two more hyperscaler contracts at >$500M each TCV. Bitcoin stabilises at >$80k so the legacy mining segment doesn't bleed cash. All three things together get you to Tier A.

Gap / bear case

The EPS miss is not a rounding error — it's nearly 40% worse than consensus and reflects D&A drag from rapid capex on GPU clusters. If ARR ramp slips by a quarter, the stock loses its "growth at any price" bid and reverts to a Bitcoin-miner multiple. Single-customer reliance on NVDA reference-design is a strength and a vulnerability.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) ARR exit run-rate disclosed >$1.5B in Q4 FY26, OR (b) a second hyperscaler contract signed publicly. Downgrade to C if the ARR target gets pushed to CY27.

The trade

- ***Entry zone****: $42-46 on pullback (50-day moving average zone)*
- ***Stop****: $36 close (below NVDA-announcement gap)*
- ***Position size****: 1% NLV (small-cap concentration risk)*
- ***Catalyst date****: Q4 FY26 print August 27 2026; interim NVDA/MSFT pull-through updates*
- ***Conviction****: 6/10*

Layer 0 · NBIS — Nebius Group

Thesis

The best print in the entire scaling-cloud cohort — $399M revenue (+684% YoY), 45% EBITDA margin, $27B Meta contract + $2B NVDA equity stake. Tier B not Tier A because the stock is +50% in a single month and capex is being raised to $20-25B, which materially raises balance-sheet risk even on a Meta-anchored book.

What it does + financial print

Nebius is the European answer to CoreWeave — formerly Yandex's cloud unit, now a standalone NASDAQ-listed AI cloud with anchor sites in Finland, Israel, France, and now a 1.2 GW Pennsylvania mega-site announced May 12. NVDA's reference architecture, Meta's $27B 5-year contract, and $2B in fresh NVDA equity — the model is GPU-as-a-service with secular take-or-pay structure.

Q1 2026 printed $399M revenue (+684% YoY) versus ~$370M consensus. Net income swung to $621M positive (large one-time on equity remeasurement). ARR was $1.92B at end-March, exit-CY26 target raised to $7-9B (3-4x current). Capex raised to $20-25B for FY26. The print is the cleanest scaling-cloud number we've seen.

Bull case

ARR exits CY26 at >$8B (high end of guide), Meta contract triggers additional volume, and the 1.2 GW Pennsylvania site comes online on schedule in 2027. EBITDA margins hold at 45%+ as scale grows. At that trajectory the stock is a 2026 Tier A holding by the August print.

Gap / bear case

+50% in a month is a parabolic move on top of an already-extended chart. $20-25B of capex on a balance sheet still developing access to investment-grade debt is a real risk — any speed-bump in ARR ramp combined with rising rates and the equity has to fund the gap dilutively. Meta concentration: one customer is ~70%+ of contracted ARR.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) a pullback to $170-185 that holds 50-day, AND (b) Q2 print confirming ARR run-rate ≥$3B with no capex surprise to the upside. Downgrade only on Meta contract renegotiation.

The trade

- ***Entry zone****: $170-185 on the first real pullback; do not chase $230+*
- ***Stop****: $158 close (below recent breakout shelf)*
- ***Position size****: 1.5% NLV starter; 3% on confirmation*
- ***Catalyst date****: Q2 print August 7 2026*
- ***Conviction****: 7/10*

Layer 0 · WULF — TeraWulf

Thesis

The smaller-cap AI HPC pivot story — 60% of Q1 revenue is now from AI HPC versus zero a year ago, with $3.1B cash on the balance sheet to fund the build-out. Tier B because the Q1 EPS missed by 5x consensus, the financial mechanics are still ugly (D&A drag on capex), and execution risk on the ramp is severe.

What it does + financial print

TeraWulf converts hydroelectric and nuclear-adjacent Bitcoin mining sites in upstate New York into AI HPC colocation capacity. The Lake Mariner facility is the flagship. The model is similar to IREN — depreciating crypto economics being converted into multi-year colo / GPUaaS ARR — but with a domestic-renewable angle that resonates with hyperscaler ESG sourcing.

Q1 2026 printed $34.0M revenue versus $32.6M consensus (HPC lease ~$21M, AI HPC now 60% of mix). EPS missed at -$1.01 versus -$0.19 — a major miss reflecting capex run-rate plus D&A acceleration. $3.1B liquidity (cash + restricted) keeps the build-out funded.

Bull case

The AI HPC pivot continues to ramp — by Q4 FY26, HPC is 80%+ of revenue and the company prints a positive EBITDA quarter. A second hyperscaler or Tier-2 GPU cloud signs a multi-year colo contract anchored at Lake Mariner. The $3.1B cash funds capacity expansion without further dilution.

Gap / bear case

The EPS miss is a warning that the model is not yet self-funding — D&A on rapidly-deployed GPU clusters is eating the lease economics. Small-cap, single-site concentration means any operational hiccup (power curtailment, contract delay) is binary. The stock is +27% on the month and would unwind fast on bad print.

Trigger to upgrade / downgrade

Upgrade to Tier A on a positive EBITDA quarter combined with a second anchor customer disclosed. Downgrade on a second consecutive >3x EPS miss or capex revision higher without ARR proof.

The trade

- ***Entry zone****: $17-19 on pullback (gap-fill from April breakout)*
- ***Stop****: $14.50 close (200-day)*
- ***Position size****: 0.75% NLV (smaller-cap risk)*
- ***Catalyst date****: Q2 print August 12 2026*
- ***Conviction****: 6/10*

Layer 1 · AMD — Advanced Micro Devices

Thesis

Strong beat-and-raise — DC segment +57% YoY, MI400 launch ahead, custom hyperscaler wins. Tier B not Tier A because the stock is already +98% YTD and absorbing two analyst downgrades (HSBC, Daiwa) plus insider selling — the easy-money phase of the AI accelerator rerate has been claimed.

What it does + financial print

AMD is the second GPU vendor — Instinct MI300X / MI325X today, MI400 (CDNA 4) launching, MI500 in design. The data-centre CPU franchise (EPYC Turin, Venice) is the cash-cow tied to inference workloads. The structural bull case: as inference demand grows, customers diversify away from NVDA's CUDA monopoly toward AMD's open ROCm stack, and AMD captures 10-20% AI accelerator share.

Q1 2026 printed $10.25B revenue (+38% YoY) versus $9.95B consensus, and $1.37 EPS versus $1.25. The data-centre segment hit $5.78B (+57% YoY). Q2 guide raised to ~$11.2B (+46% YoY). The stock initially gapped +16% pre-market then faded -7% as insiders sold into strength and HSBC/Daiwa downgraded on valuation.

Bull case

MI400 ramps cleanly through H2 26, hyperscaler custom design wins materialise (rumours of Anthropic/Meta MI-series volume orders), and the ROCm software ecosystem gets enough developer mass that the CUDA-lock thesis weakens. A clean print where DC segment crosses $8B in a single quarter would be the catalyst.

Gap / bear case

+98% YTD with insider selling is the textbook setup for a 15-25% correction. NVDA Blackwell Ultra and Rubin keep raising the bar — every cycle AMD has to ship a credible competitor and price below NVDA, which means margin pressure. The MI300X ramp last cycle taught everyone that customer concentration (just MSFT and Meta) is fragile.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) a clear MI400 launch with three named hyperscaler design wins, OR (b) a pullback to $185-195 that holds 200-day. Downgrade on a delayed MI400 ship date or DC segment sequential growth dropping below 10%.

The trade

- ***Entry zone**: $190-205 on pullback; tolerate up to $220*
- ***Stop**: $175 close (200-day)*
- ***Position size**: 2% NLV starter, 3.5% full*
- ***Catalyst date**: Q2 print August 4 2026; MI400 launch event (TBD H2)*
- ***Conviction**: 7/10*

Layer 1 · ARM — Arm Holdings

Thesis

Record FY26 — DC royalty doubled YoY, first Arm-branded chip in 35 years announced, edge-AI / physical-AI royalty stream emerging. Tier B because the multiple sits at 60x+ forward earnings and the stock is +15% in the last month and +40% in 3 months — the structural bull case is intact but the valuation cushion is non-existent.

What it does + financial print

Arm licenses CPU IP — Cortex / Neoverse cores, plus the architecture itself. Every smartphone in the world runs on Arm; the new growth vector is Arm in data centres (Neoverse V2/V3 in Graviton, Cobalt, Axion) and Arm in AI edge devices (Cortex-M, Ethos NPU). The own-chip announcement — first in-house silicon in 35 years, targeted at AI workloads — is a structural shift in the business model from pure IP-licensing toward chip-design royalties.

Q4 FY26 (May 6) printed $1.49B revenue (record), FY $4.92B, Q4 royalty $671M. EPS $0.60 vs $0.54 consensus (+11%). Data-centre royalty doubled YoY. Guide in-line. TIKR set $599 PT post-print.

Bull case

The in-house chip announcement is the catalyst that justifies a higher multiple — if Arm captures custom-silicon royalties at $5-10 per chip on top of its core licensing, FY27 EPS expands materially. Data-centre royalty doubling is the second leg. Edge-AI / physical-AI device proliferation is the third.

Gap / bear case

60x+ forward P/E means any miss or in-line guide is punished. RISC-V open-ISA competition is real and growing — Tenstorrent, SiFive, Western Digital all moving away from Arm in specific embedded segments. The own-chip strategy will antagonise some existing licensees (Apple/QCOM-adjacent).

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) the first in-house chip shipping with named customer commitments, OR (b) a pullback to $170-180. Downgrade if a tier-1 licensee defects to RISC-V for a flagship product.

The trade

- ***Entry zone****: $170-185 on the first real pullback*
- ***Stop****: $160 close (50-day support)*
- ***Position size****: 1.5% NLV*
- ***Catalyst date****: Q1 FY27 print August 5 2026; own-chip launch event (TBD)*
- ***Conviction****: 6/10*

Layer 1 · INTC — Intel Corporation

Thesis

The earnings turnaround is real — Q1 was a massive beat ($0.29 EPS vs $0.01 consensus), DC +22% YoY, AI is 60% of revenue. Tier B not Tier A because the stock is +360% TTM and at an all-time high $95.73 — the turnaround narrative is largely priced and execution risk on 18A foundry is still meaningful.

What it does + financial print

Intel is the integrated CPU + foundry story. Xeon (Granite Rapids / Sierra Forest) in data centres, Core Ultra in PCs, Gaudi in AI accelerators, and the 18A node ramping in Arizona as the foundry play. Under Lip-Bu Tan the narrative has shifted from “Intel can’t execute” to “the 18A node is on schedule, AI silicon is shipping, and the foundry has customers” — speculation about a US government / TSMC JV adds optionality.

Q1 2026 (April 23) printed $13.58B revenue versus $12.42B consensus — a large beat — and $0.29 EPS versus $0.01 consensus (massive). Q2 guide raised to $13.8-14.8B versus $13.07B consensus, and $0.20 EPS versus $0.09. Stock hit ATH $95.73 on May 8.

Bull case

18A volume ramps Q4 26 with a named external customer disclosure (the unnamed third-party design Intel referenced becomes real). Gaudi3 / Gaudi4 captures a credible 5-10% AI accelerator share. A US-government-backed foundry JV is announced. Any one of these unlocks Tier A.

Gap / bear case

+360% TTM means the easy turnaround money is gone. The 18A node has to deliver real customer revenue, not just pilot lots — and Intel's history of foundry execution disappointments is long. AMD continues to take Xeon share in cloud. AI accelerator share remains in the low single digits.

Trigger to upgrade / downgrade

Upgrade to Tier A on a named 18A external customer disclosure with volume commitment, OR a pullback to $75-80 that holds. Downgrade on a 18A node delay or a Gaudi roadmap setback.

The trade

- ***Entry zone***: *$75-82 on pullback*
- ***Stop***: *$68 close (50-day support)*
- ***Position size***: *1.5% NLV*
- ***Catalyst date***: *Q2 print July 23 2026; 18A external customer announcement (TBD)*
- ***Conviction***: *6/10*

Layer 1 · QCOM — QUALCOMM Inc.

Thesis

Beat + record auto ($1.33B, +38% YoY), data-center CPU plans, tariff de-escalation. Tier B because the stock has consolidated after hitting $247 ATH and now sits at $200 — the rerate is partially through, and the Apple modem business sunsets in 2027 is a structural headwind that hasn't fully played out.

What it does + financial print

QCOM is the smartphone modem-and-SoC franchise being repositioned around three AI-era vectors: Snapdragon for AI PCs (Microsoft Copilot+ partnership), Snapdragon Cockpit / Ride for AI in automotive ($1.33B record quarter), and the new data-centre CPU effort (Nuvia-based, server-class Arm cores). The bear case has always been "Apple goes in-house on modems" — the bull case is QCOM has three new growth engines outside handsets.

Q2 FY26 (April 29) printed $10.60B revenue versus ~$10.5B consensus, and $2.65 EPS versus $2.56 (+3.5%). Auto revenue was a record $1.33B (+38% YoY). Guide in-line. The stock hit ATH $247.90 on May 11 then settled back to $200 at writing.

Bull case

The data-centre CPU announcement (rumoured for late 2026) lands with a named hyperscaler design partner. Auto continues to compound at +30%+ YoY. AI PC volumes ramp in Copilot+ as Microsoft pushes Snapdragon. Any one of these — especially the DC CPU — could trigger a Tier A move.

Gap / bear case

Apple modem sunsets in late 2026/early 2027 — that's ~$3-4B annual revenue at high margins disappearing. The DC CPU effort is unproven and competes with AMD EPYC + NVDA Grace at the high end. ATH at $247 means the market has digested the bull case already; the consolidation phase tells you positioning is full.

Trigger to upgrade / downgrade

Upgrade to Tier A on a named DC CPU customer disclosure, OR a pullback to $175-185 that holds. Downgrade on Apple modem cliff materialising faster than expected or DC CPU launch slippage.

The trade

- ***Entry zone****: $175-185 on pullback*
- ***Stop****: $162 close (200-day)*
- ***Position size****: 1.5% NLV*
- ***Catalyst date****: Q3 FY26 print July 30 2026; DC CPU announcement (TBD H2)*
- ***Conviction****: 6/10*

Layer 2 · AMAT — Applied Materials

Thesis

Record Q2, guide raised — but Tier B not Tier A because AMAT is less HBM-pure than LRCX or KLAC, and the China export overhang caps the multiple. The thesis works but the path to outperformance vs the WFE basket is less obvious than the pure HBM names.

What it does + financial print

AMAT is the broadest WFE vendor — deposition (Endura, Producer), etch (Sym3 Magnum), implant, CMP, ion-beam — supplying every major fab globally. AI tailwind hits AMAT through HBM packaging (through-silicon-via deposition, hybrid bonding tools), advanced logic (DRAM trench, gate-all-around finFET), and the new EUV-deposition Sciencia tool family.

Q2 FY26 (May 14) printed $7.91B revenue (record, +11.4% YoY) versus $7.69B consensus — a beat — and $2.86 EPS versus $2.68 (+6.7%). Q3 guide raised to EPS $3.36 ± $0.20 (above consensus $3.10). The fundamentals are clean.

Bull case

A specific HBM-tool design win disclosure — say, AMAT capturing the deposition pull on HBM4E ramp at SK Hynix or Samsung — re-rates AMAT closer to LRCX's multiple. China demand resilience continues (mature-node tools not yet under export control). Sciencia EUV-deposition lands a tier-1 customer.

Gap / bear case

The China export licence regime tightens under the Trump administration, hitting both DUV-adjacent tools and HBM-relevant deposition gear. Consensus already models 15% revenue growth — anything less will reset the multiple. AMAT's general-purpose exposure is a weakness in a market that rewards HBM/AI-purity.

Trigger to upgrade / downgrade

Upgrade to Tier A on an explicit HBM-tool win disclosure with named customer, OR a pullback to $390-410. Downgrade on China export rules expanding to AMAT's HBM-adjacent SKUs.

The trade

- ***Entry zone****: $390-410 on pullback*
- ***Stop****: $370 close (200-day)*
- ***Position size****: 1.5% NLV*
- ***Catalyst date****: Q3 print August 14 2026*
- ***Conviction****: 6/10*

Layer 2 · ASML — ASML Holding

Thesis

The lithography monopoly intact — FY26 guide raised to EUR 36-40B, High-NA EUV ramping, Intel 18A and TSMC A14 pulls accelerating. Tier B because the China DUV export overhang (roughly 25% of pre-controls revenue) is a persistent headwind that caps multiple expansion until either it stabilises or proves immaterial.

What it does + financial print

ASML is the sole supplier of EUV lithography tools on the planet — every leading-edge logic and DRAM fab below 7nm runs ASML tools. Low-NA EUV is the workhorse; High-NA EUV (EXE:5000, EXE:5200) is the next-node tool starting volume ramp now. The DUV tools (immersion + dry) are the mid-node franchise still highly profitable but increasingly under China export scrutiny.

Q1 2026 (April 15) printed EUR 8.8B revenue versus EUR 8.5B consensus — a beat. EUR 2.8B net income vs 2.5B consensus. FY26 guide raised to EUR 36-40B (from 34-39B). Q2 guide EUR 8.4-9.0B. Stock initially sold off -6% on China export-control tightening news the next day.

Bull case

High-NA ramps cleanly at Intel 18A, TSMC A14, and Samsung 2nm/1.4nm — each High-NA tool is ~EUR 380M, so even a single-digit number of incremental tool wins matters. China DUV pressure stabilises (either via licence flexibility or non-China demand offsetting). FY27 guide stays >20% growth.

Gap / bear case

China DUV revenue is ~25% of historical mix and under sustained pressure. Even a clean monopoly trades at a discount if a meaningful share of TAM is geopolitically encumbered. Trump-administration export controls could escalate further — there are scenarios where DUV-to-China gets fully restricted.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) High-NA Q3 or Q4 print showing >3 tools shipped to leading-edge customers, OR (b) a pullback to $1,300-1,400. Downgrade on China DUV revenue dropping below 10% of total without offset.

The trade

- ***Entry zone****: $1,300-1,400 on pullback*
- ***Stop****: $1,200 close (200-day)*
- ***Position size****: 2% NLV*
- ***Catalyst date****: Q2 print July 16 2026*
- ***Conviction****: 7/10*

Layer 2 · TSEM — Tower Semiconductor

Thesis

EPS beat, record Q2 guide ($455M, +22% YoY), silicon photonics narrative for AI optical interconnects, defense chip wins. Tier B not Tier A because the Q1 revenue actually slightly missed ($414M vs $419M consensus) even as EPS beat, and the gap between sell-side median PT ($142) and spot ($254) creates near-term pressure as analysts catch up.

What it does + financial print

Tower is a specialty foundry — RF SOI (smartphone), SiGe (5G/mmWave), power-management, silicon-photonics, image sensors. The AI thesis runs through silicon photonics — Tower fabs the photonic-integrated-circuit (PIC) wafers that feed downstream optical-transceiver assembly at AOI, Coherent, Lumentum, etc. The defense angle (US fab in New Mexico) is a sovereign-supply chain hedge.

Q1 2026 (May 13) printed $414M revenue versus $419M consensus — a slight miss but +15% YoY. EPS $0.65 vs $0.57 (+14% beat). Q2 guide raised to $455M (record, +22% YoY). Stock gapped +15.4% pre-market on the print.

Bull case

Silicon photonics PIC volumes ramp at scale as 800G/1.6T optical transceivers move to PIC-based designs — that's the architecture Tower has been investing toward for 3 years. Defense chip wins compound (the 300mm New Mexico ramp is part of this story). A clean Q3 print with continued sequential growth flips this to Tier A.

Gap / bear case

The Q1 revenue miss tells you Tower's volumes are not yet exponential — they're growing nicely but specialty foundry is inherently lumpy. Median PT $142 versus spot $254 means analysts have not endorsed the rerate — that's either a buying signal or a warning that fundamentals don't yet justify the price.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) a clean Q2/Q3 print with revenue and guide both up, AND (b) at least one PIC-related design-win disclosure with a named photonics OEM. Downgrade if Q2 prints below guide or PIC ramp gets pushed.

The trade

- ***Entry zone**: $220-235 on pullback (gap fill)*
- ***Stop**: $205 close (50-day)*
- ***Position size**: 1.5% NLV*
- ***Catalyst date**: Q2 print August 12 2026*
- ***Conviction**: 7/10*

Layer 3 · HY9H — SK Hynix Sponsored GDR

Thesis

The European wrapper for 000660 — same HBM monopoly, same record Q1, same 2027/2028 supply tightness — listed in Frankfurt for EUR investors. Tier B not Tier A purely to avoid double-counting Tier A exposure that's already in the book via 000660; this is a valid B-tier proxy for accounts that can't hold KRW directly.

What it does + financial print

SK Hynix is the dominant HBM3E supplier to NVIDIA (>50% share through Hopper, locked through Blackwell, primary on Rubin pending HBM4 qualification), the global #2 DRAM supplier, and a top-3 NAND vendor. The HBM monopoly is the entire bull case — HBM supply is sold out through 2027 and the TAM is expanding from $35B in 2025 to $100B+ by 2028.

Q1 2026 (April 23) printed KRW 52.6T revenue (+198% YoY, +60% QoQ) — a beat. OP KRW 37.6T at 72% operating margin. Net income KRW 40.3T at 77% NM. Guidance: HBM demand far exceeds capacity next 3 years; HBM4E samples 2H26. Stock at all-time-high KRW ~1.83-1.95M; HY9H GDR at EUR 594.

Bull case

HBM4 NVDA Rubin design-win confirmation, HBM4E sample-to-volume transition through 2027, and supply tightness extending to 2028 (per Samsung's tracker referenced in the data). EUR-based investors get the same fundamentals as 000660 holders with simpler settlement.

Gap / bear case

The whole memory complex is at cycle-high valuations — +35% in 30 days, +77% in 90. Any signal that Samsung's HBM4 qualifies at NVDA (potentially rebalancing share) would compress SK Hynix's premium. KRW FX risk is real for the underlying.

Trigger to upgrade / downgrade

Upgrade to Tier A on a pullback to EUR 520-560 (the breakout shelf), reclassifying as the primary HBM exposure for EUR books. Downgrade if Samsung wins majority HBM4 share at NVDA Rubin.

The trade

- ***Entry zone****: EUR 520-560 on pullback*
- ***Stop****: EUR 480 close (50-day)*
- ***Position size****: 1.5% NLV (or 0% if already holding 000660)*
- ***Catalyst date****: Q2 print July 24 2026*
- ***Conviction****: 7/10*

Layer 3 · 005930 — Samsung Electronics

Thesis

Record Q1 — memory profit +48x YoY, OP exceeded full-year 2025 in a single quarter, FY26 record. Tier B not Tier A because the HBM4 NVIDIA qualification is the gating issue — Samsung still hasn't cleared NVDA HBM3E qual, the share gap to SK Hynix persists, and the -8.6% single-day drop on May 15 signals the positioning unwind has begun.

What it does + financial print

Samsung is the conglomerate — memory (DRAM, NAND, HBM), foundry (3nm, 2nm in ramp), mobile (Galaxy), display, consumer electronics. The AI thesis runs almost entirely through memory and foundry. Memory is exceptional. Foundry has been losing money. NVDA HBM3E qualification has been pending for ~18 months and is the single binary catalyst for this stock.

Q1 2026 (April 30) printed KRW 133.9T revenue (+69% YoY) — record. OP KRW 57.2T — record, 8x+ YoY. Memory profit +48x YoY. HBM revenue to 3x in 2026 vs 2025. HBM4E samples Q2. Foundry still drag. Stock -8.6% on May 15 to KRW 270,500 after touching ATH KRW 299,500 on May 14.

Bull case

NVDA finally qualifies Samsung HBM3E (and/or HBM4) for volume — that one event closes the share gap with SK Hynix, expands Samsung's revenue, and triggers a multi-quarter re-rate. Foundry 2nm stabilises with one named external customer (Tesla rumour, Google rumour). The conglomerate trades at a sum-of-parts discount that closes.

Gap / bear case

The qualification has been "imminent" for 18 months and hasn't happened. Every quarter it doesn't happen is another quarter of SK Hynix locking in NVDA Rubin allocations. Foundry losses continue. Mobile is structurally slowing. Labour-strike rumours add execution noise.

Trigger to upgrade / downgrade

Upgrade to Tier A on NVDA HBM3E / HBM4 qualification announcement with volume commitment. Downgrade on a public NVDA decision to single-source SK Hynix on HBM4.

The trade

- ***Entry zone***: *KRW 250,000-265,000 on pullback*
- ***Stop***: *KRW 235,000 close (50-day)*
- ***Position size***: *2% NLV*
- ***Catalyst date***: *NVDA HBM3E qual announcement (anytime); Q2 print July 30 2026*
- ***Conviction***: *6/10*

Layer 3 · SMSN — Samsung Electronics GDR

Thesis

The London-listed GDR proxy for 005930 — same binary NVDA HBM qualification catalyst, same memory tailwind, same foundry drag, USD-denominated. Tier B because it inherits 005930's catalyst structure exactly — if you don't own the underlying, this is the dollar-settled way to express the view.

What it does + financial print

SMSN is the GDR — five-to-one underlying conversion ratio, USD-priced, trades on LSE / SIX. The underlying business is exactly Samsung Electronics: memory (HBM3E, HBM4 pending qual), foundry (3nm/2nm), and mobile/consumer.

Q1 2026 was in line; memory beat, foundry dragged. HBM3E pricing +20% confirmed for 2026 shipments. NVDA HBM3E qual still pending. GDR price $4,538 on May 18 — down ~7% in the prior 24h on profit-taking after the May 14 melt-up.

Bull case

Same as 005930 — NVDA HBM qualification flips the whole memory franchise from "trailing SK Hynix" to "co-leader" and the GDR rerates with it. USD investors get a clean way to hold the catalyst without KRW exposure.

Gap / bear case

Same as 005930 — qualification has been pending too long, SK Hynix is widening the lead, and foundry losses keep eating earnings. Plus the GDR specific: liquidity is thinner than the local listing, and tracking error vs underlying can drift in stress.

Trigger to upgrade / downgrade

Upgrade to Tier A simultaneously with 005930 on NVDA HBM qualification. Downgrade on a confirmed NVDA single-source decision favouring SK Hynix on HBM4.

The trade

- ***Entry zone**: $4,100-4,300 on pullback*
- ***Stop**: $3,850 close (50-day)*
- ***Position size**: 1.5% NLV*
- ***Catalyst date**: NVDA HBM3E qual announcement (anytime); Q2 print July 31 2026*
- ***Conviction**: 6/10*

Layer 4 · 007660 — ISU Petasys

Thesis

Korean AI server HDI PCB pure-play — direct beneficiary of the AI server / supercomputer backplane mix shift. 2025 revenue +30% YoY, trailing 12-month gain +360% as the structural tightness in AI HDI PCB capacity gets priced. Tier B because the price source disagreement (TradingView KRW 106,400 vs Investing.com KRW 141,400) prevents Tier A confidence until reconciled with a clean local source.

What it does + financial print

ISU Petasys makes high-density interconnect (HDI) PCBs for AI servers, networking switches, and supercomputer backplanes. The AI server build-out is structurally tightening HDI capacity because the layer count and board complexity of GB200 NVL72 and equivalent racks is materially higher than prior-generation enterprise servers — a single Blackwell tray uses several times the PCB content of a Hopper tray.

Q1 2026 (May 8) printed revenue beat (2025 full-year +30% YoY to KRW 1.09T). EPS beat. Guidance constructive on AI server HDI PCB demand. Stock at KRW 106,400 on May 16 per primary source; needs_manual_check for the variance vs Investing.com.

Bull case

AI server HDI demand continues to outrun supply through 2027 as hyperscalers can't get enough backplane capacity. ISU captures a tier-1 hyperscaler design win for a next-gen networking switch. Q2 print confirms the +30% YoY revenue trajectory. Price-source reconciliation pushes the stock to the higher of the two prints.

Gap / bear case

+360% TTM is a parabolic move that invites profit-taking on any disappointment. Small-cap Korean OSAT exposure means liquidity is thin and FX-translated returns can lag the local print. The price source variance suggests one of the venues has stale data — needs to be sorted out before sizing.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) price source reconciliation showing clean data above KRW 100,000, AND (b) a named hyperscaler / Korean defense backplane win. Downgrade on revenue deceleration below +20% YoY.

The trade

- ***Entry zone**: KRW 92,000-100,000 on pullback (subject to price-source reconcile)*
- ***Stop**: KRW 84,000 close*
- ***Position size**: 1% NLV (already owned 5K sh; treat as add-on)*
- ***Catalyst date**: Q2 print August 10 2026*
- ***Conviction**: 7/10*

Layer 4 · 009150 — Samsung Electro-Mechanics

Thesis

Three-leg AI thesis — ABF substrate share gains as Ibiden / Unimicron run at capacity, MLCC content per AI server rising, and camera-module / semi packaging diversification. Tier B because v3 had this at WATCH 7/10 and the price action validates the watch but hasn't yet earned the upgrade — needs a single quarter where ABF revenue is explicitly disclosed at materially higher mix.

What it does + financial print

Samsung Electro-Mechanics (SEMCO) is a Samsung-affiliated component supplier — three segments: MLCC (multi-layer ceramic capacitors, ~30-40% of mix), camera modules (for Galaxy phones), and semiconductor packaging (ABF substrates, FCBGA). The AI angle runs through ABF substrate share-gain (the duopoly with Ibiden / Unimicron is so tight that any qualified third supplier wins business) and MLCC content increasing per AI server.

Q1 2026 (April 29) printed in-line revenue, in-line EPS. Guidance constructive on ABF/MLCC AI server mix rising. Stock at KRW 1,024,000 on May 14, near ATH.

Bull case

ABF substrate volumes ramp at scale — every Blackwell/Rubin GPU package needs a substrate, and SEMCO is the third source. A clean quarter where SEMCO discloses ABF revenue at >10% of total mix, or names a hyperscaler design partner, would unlock Tier A. MLCC pricing tightening with AI server build adds the second leg.

Gap / bear case

Samsung captive customer dependency — the smartphone OEM business and Galaxy camera-module mix is mature and not growing. MLCC pricing is cyclical and the up-cycle is in its later innings. The ABF story is real but not yet quantified in the financial disclosures.

Trigger to upgrade / downgrade

Upgrade to Tier A on an explicit ABF substrate revenue disclosure with named hyperscaler / AI customer. Downgrade on smartphone-mix collapse hitting >20% of revenue.

The trade

- ***Entry zone****: KRW 920,000-980,000 on pullback*
- ***Stop****: KRW 870,000 close (200-day)*
- ***Position size****: 1% NLV*
- ***Catalyst date****: Q2 print July 29 2026*
- ***Conviction****: 6/10*

Layer 4 · 067310 — Hana Micron

Thesis

Direct SK Hynix HBM packaging supplier — small-cap KOSDAQ name that has 4x rerated from KRW 10,030 to KRW 42,800 over 12 months as the SK Hynix HBM volume ramp pulled through its capacity. Tier B because the single-customer concentration risk is severe (this is essentially a leveraged SK Hynix HBM call) and KOSDAQ liquidity is thin.

What it does + financial print

Hana Micron is a Korean semiconductor packaging supplier — FCBGA, WLCSP, HBM packaging support. The direct SK Hynix relationship is the entire bull case: as SK Hynix's HBM volume scales through HBM3E / HBM4 / HBM4E, the packaging supply chain has to scale with it, and Hana Micron is one of the named beneficiaries.

Q1 2026 (May 6) printed in-line revenue (2024 full-year +29% YoY). EPS beat. Guidance constructive — HBM packaging tied to SK Hynix volume. Stock near 52-week high KRW 44,850 with the 52-week low at KRW 10,030 — a 4x rerate on the year.

Bull case

SK Hynix HBM volumes continue to outrun supply through 2027 — every incremental HBM stack pulls more Hana Micron packaging revenue. Capacity expansion announcements (FCBGA + WLCSP) come through with named utilisation. A clean Q2 print confirms the +30% YoY trajectory.

Gap / bear case

This is a leveraged single-customer name. If SK Hynix's HBM share at NVDA gets eroded by Samsung qualifying, Hana Micron is the second-derivative loser. KOSDAQ small-cap liquidity means any drawdown is sharper than the underlying fundamentals justify. The 4x rerate is largely priced.

Trigger to upgrade / downgrade

Upgrade to Tier A on an explicit HBM packaging revenue disclosure at >50% of mix AND a SK Hynix HBM4E capacity-expansion sign-off naming Hana Micron. Downgrade on SK Hynix HBM share loss to Samsung.

The trade

- ***Entry zone****: KRW 36,000-40,000 on pullback*
- ***Stop****: KRW 32,000 close (50-day)*
- ***Position size****: 0.75% NLV (small-cap concentration)*
- ***Catalyst date****: Q2 print August 10 2026*
- ***Conviction****: 6/10*

Layer 4 · 3037 — Unimicron Technology

Thesis

Taiwanese ABF substrate duopolist with Ibiden — the bottleneck that gates every Blackwell, every Rubin, every AWS Trainium 2 package. Q1 26 printed a profit-margin inflection that validates the structural-tightness thesis (NI +451% YoY, op margin 14% vs 3% YoY). Tier B not Tier A only because the stock is +263% trailing year and a -2.3% intraday on May 18 hints that the marginal buyer is finally tiring — the cleanest expression of the substrate bottleneck, but the entry has to be patient.

What it does + financial print

Unimicron sits at Layer 4 — the advanced packaging / substrate layer that sits underneath every leading-edge accelerator package. ABF (Ajinomoto Build-up Film) substrates are the printed-circuit-board-on-steroids that carry the GPU die and the HBM stacks on a CoWoS interposer; Ibiden and Unimicron are the only two suppliers who can hit the layer count (20+ build-up layers), the line/space (sub-15μm), and the yield required for AI accelerator packages. Samsung Electro-Mechanics is the credible third source but is still ramping. Capacity is sold out through 2026 and the lead times have stretched from six months to nine-plus.

Q1 FY26 printed NT$37.4B revenue (+25% YoY), EPS NT$3.00 vs NT$0.60 YoY (a massive beat — consensus had ~NT$1.80), net income +451% YoY. Operating margin expanded to 14% from 3% — that's the operating-leverage signature of a capacity-constrained business taking price. Guidance described capacity as "sold through" — management code for "we are pricing what we can ship." Stock at NT$802 on May 18, ~5% off the recent high.

Bull case

ABF substrate tightness persists through 2027 because the only meaningful capacity adds (Ibiden's Ogaki North, Unimicron's Taoyuan expansions, SEMCO's Sejong line) all come online into a demand curve that grows faster — every Blackwell→Rubin→Vera-Rubin Ultra generation increases substrate area per package. Unimicron at sub-25x forward earnings is materially

cheap if you believe the cycle has another 18-24 months. The optionality on top: CoPoS / glass substrate platform sets up the next-decade tail with TSMC as the named partner.

Gap / bear case

+263% TTM is the gap — this is a stock that has already priced the entire near-term substrate-tightness narrative. The -2.3% intra-day on May 18 isn't a thesis-break but it is a positioning warning: when a Taiwanese substrate name with this much momentum starts trading heavy intra-day, the next test is whether NT$750 holds. If it doesn't, the rerating from "structural tightness" back to "cyclical winner" can take 20-30% off the multiple before anything in the fundamentals changes. Secondary risk: if ABF substrate yields improve faster than expected at the third-source (SEMCO), the duopoly premium compresses.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) a clean break and hold above NT$830 with confirmation of Q2 revenue continuing >+20% YoY, AND (b) explicit AI/HPC substrate revenue disclosure at >50% of mix. Downgrade on loss of NT$720 close (50-day proxy) — that level marks the line between healthy consolidation and trend break.

The trade

- ***Entry zone****: NT$740-780 on pullback*
- ***Stop****: NT$700 close (50-day MA)*
- ***Position size****: 1.5% NLV (already owned 10K sh; treat as add-on at zone)*
- ***Catalyst date****: Q2 print July 29 2026*
- ***Conviction****: 7/10*

Layer 4 · 4062 — Ibiden

Thesis

Japanese ABF substrate co-leader — the other half of the substrate duopoly that gates every AI accelerator package. The narrative is intact (capacity sold through, profit-surge commentary, electronics segment strong), but Q4 FY26 EPS missed consensus (-13.5%) and the stock is up +881% trailing year. Tier B not Tier A because at +881% TTM with an earnings miss in the last print, the margin of safety is gone — you wait for a real reset before sizing up.

What it does + financial print

Ibiden is the Japanese half of the ABF substrate duopoly — the supplier that has historically held the highest-layer-count and highest-yield substrates at Ibiden's Ogaki facility. Ibiden's substrates carry the package for Intel server CPUs (legacy strength) and increasingly for NVIDIA Blackwell / Rubin GPUs and AWS Trainium 2 ASICs. Ibiden also makes ceramic honeycomb structures for automotive catalytic converters (legacy, cyclical, low-multiple) — that mix dilution is part of why the multiple has historically been compressed and why the rerating once the AI mix dominated was so violent.

Q4 FY26 (May 8) revenue mixed, EPS JPY 31.98 vs JPY 36.97 consensus (-13.47%). The miss did not break the narrative — guidance was "constructive — strong profit projection growth on electronics." Stock hit an ATH at JPY 16,375 and then advanced to JPY 17,060. Market cap JPY 4.33T. +881% trailing year.

Bull case

ABF substrate tightness through 2027; Ibiden runs at capacity and prices accordingly. The Ogaki North capacity expansion comes online into a demand curve that is still steepening, not flattening. The auto-catalytic-converter cyclicality fades as a mix issue because AI / DC electronics grows faster than the legacy book shrinks. FY27 OP guide raises to +25-30% would push the multiple to mid-30s forward — still expensive in absolute terms but justified if the structural-tightness story remains intact.

Gap / bear case

The EPS miss in Q4 FY26 is the warning. When a stock at +881% TTM misses its first print, the burden of proof shifts — every subsequent quarter has to print a clean beat to keep the multiple. JPY FX is a second factor: if BoJ normalisation accelerates and JPY runs to USD 140, the export earnings are translation-headwind without any operational change. Third risk: if SEMCO qualifies meaningfully as a third source, Ibiden's premium pricing power on the highest-layer-count substrates compresses.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) the next quarter printing a clean EPS beat with electronics segment >+30% YoY, AND (b) a pullback to JPY 14,000-15,000 zone that restores margin of safety. Downgrade on loss of JPY 13,500 close (which would mark a ~20% drawdown from current and a real positioning unwind).

The trade

- ***Entry zone**: JPY 14,500-15,500 on pullback*
- ***Stop**: JPY 13,500 close*
- ***Position size**: 1.5% NLV (already owned 5K sh; treat as add-on)*
- ***Catalyst date**: Q1 FY27 print August 5 2026*
- ***Conviction**: 7/10*

Layer 4 · 6315 — TOWA Corporation

Thesis

Japanese semi-equipment niche — TOWA holds a quiet near-monopoly in compression-molding equipment for HBM stack encapsulation. Every HBM4 stack that goes through TC bonding or hybrid bonding then has to be molded, and TOWA's KMS / IPMS systems are the de-facto standard. Tier B because the bottleneck is real and verifiable, but price-source variance (JPY 2,579

vs 3,050 across venues) means we are operating with a needs_manual_check flag and we cannot underwrite Tier A conviction until that clears.

What it does + financial print

TOWA makes the semiconductor manufacturing equipment that lives downstream of die-attach: **compression molding** for HBM stacks and advanced packages, **singulation** equipment that separates molded packages into individual units, and the precision molds that go into both. The AI angle: every HBM stack — whether the 8-Hi we have today, the 12-Hi shipping now, or the 16-Hi that comes with HBM4 — has to be encapsulated in a precisely controlled compression-mold step before it goes into the CoWoS process. TOWA's share inside that step is the kind of monopoly nobody talks about because the company is small (JPY 100B-ish market cap) and the disclosure is opaque.

Q4 FY26 (May 8) results need manual check — the data quality on this name is the single largest barrier to Tier A. Guidance described as "constructive." Stock at JPY 3,050 on May 15. Analyst PT JPY 3,200 (+4.92% upside on consensus). Intra-month range JPY 2,579 to JPY 3,050 — wide and source-dependent.

Bull case

HBM compression-molding demand scales linearly with HBM volume, and HBM volume is the cleanest secular growth story in the semi cycle. As 16-Hi HBM4 ramps in late 2026 / 2027, the precision required for molding intensifies (taller stacks, tighter tolerances), which is exactly where TOWA's installed base of skilled-labour + proprietary mold designs gives it a moat. A clean disclosure of FY27 guide breaking out HBM equipment as a separate revenue line would be the catalyst that closes the price-source gap and re-prices the multiple to a proper semi-cap-eq comp set.

Gap / bear case

Data quality. We cannot get a clean read on Q4 FY26 EPS, the price varies across venues by ~18%, and the disclosure is in Japanese without an aggressive English-language IR effort. For a Tier A position size we need verifiable numbers — Tier B is the honest classification. Secondary risk: TOWA's TAM is small in absolute terms; if HBM4 volume comes in below expectations (Samsung qualification slip, NVDA demand pause), there is no diversified-revenue cushion.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) clean Q1 FY27 print with HBM equipment revenue explicitly disclosed and growing >+30% YoY, AND (b) price-source reconciliation across at least two venues at consistent levels. Downgrade on FY27 guide cut or HBM4 ramp-delay headlines.

The trade

- ***Entry zone**: JPY 2,700-2,900 (subject to price reconciliation)*
- ***Stop**: JPY 2,500 close*
- ***Position size**: 0.5% NLV (small-cap + data-quality discount)*
- ***Catalyst date**: Q1 FY27 print August 8 2026*

- ***Conviction**: 5/10*

Layer 4 · AAOI — Applied Optoelectronics

Thesis

US-listed photonics integrator with a $324M 800G/1.6T backlog locked in and an Amazon volume-ramp catalyst that is already visibly inflecting (stock +18.49% in a single session on May 13). The financial print is still ugly — EPS negative, Q1 revenue missed consensus — but the order book is the leading indicator and the market is rewarding it. Tier B because the qualification catalysts (Oracle, hyperscaler #2) are binary and the company is still losing money on a GAAP basis — too volatile for Tier A sizing.

What it does + financial print

AAOI is a US-listed vertically integrated photonics supplier — they make the laser diodes, the optical sub-assemblies, and the integrated transceiver modules used in datacom, CATV and FTTH applications. Inside the AI stack, AAOI plays at Layer 4 — the optical-interconnect layer that connects every GPU rack to every switch, every switch to every spine. The 800G transceiver volume ramp at hyperscalers (Amazon named, Microsoft / Oracle in qualification) is the entire bull case. AAOI is small enough that a single $324M backlog disclosure moves the multiple visibly.

Q1 26 (May 7): revenue $151.1M vs $154.81M consensus (small miss), EPS -$0.07 vs -$0.06 (slight beat on a negative number), Q2 guide $180-198M (midpoint approximately in-line with $192.57M consensus). The headline was the $324M 800G/1.6T backlog disclosure that came with the print. Stock then ran +18.49% on May 13 to $223.10, ATH $233.67. Rosenblatt raised PT to $220 from $140 on May 11. Stock at $190.36 on May 17 — has given back some of the rip.

Bull case

The Amazon 800G volume ramp converts the $324M backlog into recognized revenue through 2H 2026 and 2027, taking AAOI to GAAP profitability for the first time in years. A second hyperscaler qualification (Oracle most likely on near-term timing, Microsoft on longer) doubles the backlog into 2027. Vertical integration (in-house laser diodes) becomes a margin-defender as the broader transceiver market commoditises around 800G. The multiple rerates from “speculative photonics” to “scaled photonics platform” — and at $190 the implied valuation has plenty of room if revenue inflects to a $1B+ annualised run-rate.

Gap / bear case

EPS is still negative. The Q1 revenue miss in a print that should have been a celebration is exactly the kind of execution wobble that gets a story-stock punished. Hyperscaler customer concentration is double-edged: if Amazon’s 800G volume ramps slower than the backlog implies, AAOI carries the working capital hangover. Competition from Coherent (much larger), Lumentum (more diversified), Innolight (Chinese, lower cost) compresses gross margin in the long run. The +18% one-day move on May 13 is exactly the kind of melt-up that invites a -25% mean-revert on any disappointment.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) Q2 print confirming backlog conversion (revenue +30%+ YoY with positive GAAP EPS), AND (b) explicit Oracle or second-hyperscaler qualification announcement. Downgrade on backlog cancellation headline or Q2 revenue miss.

The trade

- ***Entry zone***: *$150-170 on pullback (the May 11 Rosenblatt PT $140-220 brackets the volatile zone)*
- ***Stop***: *$130 close*
- ***Position size***: *0.75% NLV (small-cap + binary catalyst discount)*
- ***Catalyst date***: *Q2 print August 6 2026; Amazon 800G volume disclosure*
- ***Conviction***: *6/10*

Layer 4 · AXTI — AXT Inc

Thesis

Indium-phosphide substrate supplier — the upstream wafer that every InP laser diode (every 100G+ datacom transceiver, every 1.6T module of the next cycle) is fabricated on. Q1 26 printed a +80% EPS surprise on +38.7% YoY revenue and the gross-margin inflection (29.9% vs -6.1% YoY) is the kind of operating leverage that confirms a structural tightness. But the stock has rocketed from a 52-week low of $1.38 to an ATH of $134 — a 100x move — and entering at $124 with that history is hazardous. Tier B because the bottleneck is real but the entry is not.

What it does + financial print

AXT is a Beijing-listed-and-headquartered (US ADR) compound-semiconductor wafer supplier — they grow indium-phosphide (InP), gallium-arsenide (GaAs) and germanium boules and slice them into the substrates that III-V foundries pattern. InP is the relevant AI bottleneck: it is the substrate of choice for the high-speed laser diodes that drive 800G, 1.6T and beyond datacom optics. The InP supply chain is genuinely tight — there are only three credible suppliers globally (AXT, Sumitomo Electric, JX Nippon), Chinese InP capacity has been export-controlled, and the AI datacom demand inflection has stripped the safety-stock buffer out of the chain.

Q1 26 (April 30) revenue $26.9M (+38.7% YoY), EPS -$0.01 vs -$0.05 (+80% surprise). Non-GAAP gross margin 29.9% vs -6.1% YoY — that's a 36-percentage-point swing in twelve months. Guide constructive: InP capacity doubling 2026. Wedbush raised PT to $93 from $80. Stock at $123.78 on May 16 — above the analyst PT after the May 13 push to an ATH $134.

Bull case

InP demand growth runs ahead of capacity adds through 2027 because the 1.6T and 3.2T datacom optical roadmap consumes 2-3x the InP-substrate area per port versus 400G/800G. AXT doubles InP capacity in 2026 and runs sold-out; gross margin steady-state moves to 35-40% on the AI mix; EPS turns durably positive in 2H 2026 with the $1+ annualised run-rate that anchors

a $200+ price target. China export-control workarounds (Tianjin facility, foreign-government qualifications) prove to be the small risks the bulls said they were.

Gap / bear case

This is a 100x stock — $1.38 to $134 in under two years. The historical comp for what happens to a 100x compound-semi name when the cycle hesitates is something like Cree in 1999-2001 (down 80% from peak). The valuation at $124 prices a multi-year ramp at near-perfect execution. China geopolitical risk is real and material — AXT's Beijing operations sit in the export-control crosshairs every time US-China optical-component licensing is renegotiated. Liquidity is thin; the move down can be as violent as the move up.

Trigger to upgrade / downgrade

Upgrade to Tier A on a pullback to $80-90 with the bottleneck thesis intact and Q2/Q3 confirming the gross-margin inflection. Downgrade on loss of $90 close + China export-control negative headline.

The trade

- ***Entry zone**: $80-95 on pullback (would be Bucket A territory)*
- ***Stop**: $75 close (cuts 38% from current — sized for the volatility)*
- ***Position size**: 0.5% NLV at current; 1% if it pulls into zone*
- ***Catalyst date**: Q2 print July 30 2026; InP capacity ramp disclosure*
- ***Conviction**: 6/10*

Layer 4 · CIEN — Ciena Corporation

Thesis

US-listed coherent-optical equipment leader — Ciena's WaveLogic 6 (1.6T per-wavelength) is the platform that hyperscalers deploy for DCI (datacenter interconnect) and metro-AI fabric. Q1 FY26 (March 4) was a clean beat-beat print but the forward guide was softer than the AI-narrative crowd wanted, and the stock has run to $551 on May 17 — a $70 to ~$600 trailing-year range. Tier B because the WaveLogic 6 ramp is real but the soft guide blunts the near-term upside; the June 4 Q2 print is the catalyst.

What it does + financial print

Ciena sits at Layer 4 — the optical / networking equipment layer that connects data centers to each other, that builds the AI fabric backbone between hyperscaler campuses, that runs the metro and long-haul transport infrastructure. The product line spans optical line systems (the chassis that carries the wavelengths), coherent transceivers (the electronics that modulate and demodulate), and the WaveLogic DSP family that has held a technology lead at 800G coherent for the last two cycles. WaveLogic 6 at 1.6T per wavelength is the current generation — first hyperscaler deployments in 2026, scale ramp through 2027.

Q1 FY26 (March 4): revenue beat, EPS beat. Guidance: mixed — described as “softer than hoped” by analysts who had positioned for an AI inflection in the next-quarter print. Stock $551.80 on May 17. 52-week range $70.80-$598.84 — Ciena is one of the violent rerates of the cycle.

Bull case

WaveLogic 6 hyperscaler ramp accelerates through 2H 2026 as the 800G→1.6T transition pulls forward; Ciena’s optical line-system installed base means every WaveLogic 6 transceiver tends to drive a corresponding chassis attach. The metro-AI fabric build-out (interconnecting GPU clusters across multiple data centers) is structurally additive to Ciena’s TAM because hyperscalers are building out networks that look more like telecom long-haul than enterprise data center. The multiple has room to run if WaveLogic 6 revenue gets explicitly broken out — Ciena trades on a much lower forward earnings multiple than the AI optics peers because the legacy telecom mix anchors the comp set.

Gap / bear case

The Q1 guide already signaled that the AI inflection is not yet hitting the P&L — Ciena ships into hyperscaler timelines that are notoriously lumpy, and the +700% trailing 52-week range tells you the market is already paying for an inflection that hasn’t quite landed. If the June 4 Q2 print shows another quarter of guide-management without a clean upward revision, the stock can give back 20% before any thesis-break. Cisco’s Silicon One + Arista’s networking-stack push compresses Ciena’s competitive position at the high-end switching edge.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) June 4 Q2 print delivering a clean beat-and-raise with explicit WaveLogic 6 revenue disclosure, AND (b) a pullback to $450-480 that restores margin of safety. Downgrade on a second-consecutive soft guide or a hyperscaler win-loss to Cisco/Arista at the optical layer.

The trade

- ***Entry zone***: *$460-490 on pullback (or buy half-size into the June 4 print)*
- ***Stop***: *$420 close*
- ***Position size***: *1% NLV at zone; 1.5% on Tier A upgrade*
- ***Catalyst date***: *Q2 FY26 print June 4 2026*
- ***Conviction***: *6/10*

Layer 4 · COHR — Coherent Corp

Thesis

Optical components and laser systems supplier — Coherent (post the II-VI / Coherent merger) is the diversified optical platform that captures the AI datacom ramp through 800G/1.6T transceivers, indium-phosphide lasers, and the SiC/photonic-IC layer. Q3 FY26 (May 6) was a revenue beat with a modest EPS miss — the kind of “good but not perfect” print that gets a +400% one-year stock to consolidate, not crash. Tier B because the rerating has been violent

(+400% TTM) and the EPS miss tells you the operating leverage is not yet bulletproof; needs one more clean print before earning Tier A.

What it does + financial print

Coherent sits across multiple Layer 4 bottlenecks: **datacom transceivers** (the optical modules that connect every GPU to every switch), **indium-phosphide laser diodes** (the source for high-speed datacom), **silicon carbide power semis** (the EV / power-electronics business — separately monetisable), **industrial lasers and optical instruments** (the legacy II-VI book, cyclical but high-margin). The AI mix has grown materially over the last six quarters as the datacom transceiver revenue scaled into 800G — Coherent now derives an estimated 30-40% of revenue from AI / datacom optical, with the rest split between communications, industrial and semiconductor capital equipment.

Q3 FY26 (May 6): revenue $1.81B vs $1.78B consensus (clean beat), EPS $1.41 vs $1.43 (modest miss). Guidance constructive on AI optics. Jefferies PT raised to $375. Stock ATH $413, +400% trailing year. Currently $385.40 on May 17 — has held above the post-print levels but not breaking out.

Bull case

The 800G → 1.6T transceiver transition pulls Coherent revenue and margin together: 1.6T modules carry ~2x the ASP of 800G with materially better gross margin once volumes scale. The SiC power-electronics business is a free option — if Coherent monetises it through divestiture or spin (semi-frequent management commentary suggests this is on the table), the optical-pure-play remaining business rerates to a higher multiple. Customer concentration is more diversified than the smaller photonics peers (AAOI, Lumentum) — Coherent ships across the named hyperscalers and into telecom carriers.

Gap / bear case

+400% TTM with an EPS miss is the early-warning signal that operational execution is not yet at the level the multiple requires. Coherent's gross margin is lower than the pure-play AI photonics names (Lumentum, AAOI on a normalised basis) because the legacy II-VI industrial mix dilutes the headline number — and that dilution may continue if the industrial-laser cycle stays weak. SiC pricing has compressed materially over the last two quarters as Chinese capacity floods the market; the SiC slice of Coherent's revenue is a drag, not a tailwind.

Trigger to upgrade / downgrade

Upgrade to Tier A on (a) Q4 FY26 print delivering both rev AND EPS beat with an explicit 1.6T volume disclosure, AND (b) clarity on the SiC business (divest, spin, or run-rate). Downgrade on a second consecutive EPS miss or SiC pricing-led gross margin compression below 35%.

The trade

- ***Entry zone**: $340-365 on pullback*
- ***Stop**: $315 close*
- ***Position size**: 1.5% NLV at zone*

- ***Catalyst date****: Q4 FY26 print August 12 2026*
- ***Conviction****: 7/10*

Layer 4 · CRDO — Credo Technology

Thesis

Active electrical cable (AEC) pure-play — Credo's SerDes-and-cable platform replaces optical transceivers inside the rack at short reach with a power-efficient copper alternative, and the hyperscalers (Microsoft and Amazon named) have made AECs a standard part of the GPU-rack build. Q3 FY26 printed revenue +201% YoY — that's the kind of growth that earns a Tier A but Credo holds at Tier B for one reason: customer concentration is severe (MSFT + AMZN ~70%+ of revenue) and the June 1 Q4 FY26 print is the test of whether the AEC adoption broadens beyond the first two adopters.

What it does + financial print

Credo sits at Layer 4 — specifically the inside-the-rack interconnect layer that connects every GPU to every NIC to every switch within a single rack. Active electrical cables are the technology bet: copper cables with embedded SerDes / DSP chips that re-time and equalise the signal, achieving 100G/200G/400G per lane at meaningfully lower power than optical transceivers for sub-7-meter reaches. The economics: AECs cost less than optics, consume less power than optics, and are field-replaceable in a way that copper-passive cables (DAC) are not. The hyperscaler adoption pattern is the validation — Microsoft was the first, Amazon followed, and the consensus expectation is that Meta and Google qualify next.

Q3 FY26 (March 4): revenue $407.01M (+201.5% YoY), EPS $1.07 (+13.75% surprise on consensus). Q4 FY26 guide strong — AEC + optics roadmap intact. Median analyst PT $200 (high $260, low $170). Stock $183.32 on May 17 — has consolidated between the Q3 print and the June 1 Q4 print.

Bull case

The June 1 Q4 print confirms (a) MSFT and AMZN volume continuing to scale, AND (b) at least one of Meta / Google / Oracle entering qualification or initial deployment. Revenue scales from a $1.6B annualised run-rate (Q3 implied) to $3B+ in FY27 as the customer count broadens. The AEC TAM expands from the $4B-by-2028 consensus to a higher number as 200G-per-lane and 400G-per-lane SerDes unlock new use cases. Credo's optics roadmap (introduced as a hedge against AEC compression) starts contributing revenue, diversifying the platform away from pure-AEC.

Gap / bear case

Customer concentration is the gap. If either MSFT or AMZN pauses AEC orders for any reason (architecture change, internal cable program, alternative supplier qualification), Credo's revenue can drop 30-50% within a single quarter. Copper-to-optics substitution is a real long-term risk: as port speeds climb to 200G-per-lane and beyond, the reach over which AECs make economic sense shrinks, and the optics-instead-of-copper crossover point moves closer to the rack. Marvell

and Broadcom both ship competing SerDes / DSP solutions at scale — Credo's technical lead is narrow.

Trigger to upgrade / downgrade

Upgrade to Tier A on the June 1 Q4 FY26 print confirming (a) revenue growth >+100% YoY again, AND (b) at least one new named hyperscaler customer in qualification. Downgrade on Q4 print showing customer-mix concentration deepening (single customer >50% of revenue) or a guide-cut from a captive hyperscaler insourcing program.

The trade

- ***Entry zone****: $155-175 on pullback (or half-size into June 1 print)*
- ***Stop****: $140 close*
- ***Position size****: 1% NLV at zone; 1.5% on Tier A upgrade*
- ***Catalyst date****: Q4 FY26 print June 1 2026*
- ***Conviction****: 7/10*

Layer 5 · 2327 — Yageo Corporation

Thesis

Yageo is the world's largest MLCC and chip-resistor supplier outside Murata — passive components are the unglamorous ballast of every AI accelerator board, and Yageo prints the bulk of them. Q1 confirmed the AI MLCC tightening narrative with a clean +22.7% revenue beat and a 25.2% op margin. It's Tier B not Tier A because the stock just printed a -12.7% single-day pullback off NT$481, which screams positioning unwind, not thesis break.

What it does + financial print

Yageo sits at Layer 5 of the AI stack — the passive-component layer that lives underneath every GPU board, every server PSU, every HBM module. AI servers consume roughly 3-4x the MLCC content of a conventional server, and the high-temperature, high-capacitance grades that AI demands are exactly where Yageo, Murata and TDK have pricing power. Yageo's global MLCC share sits in the high teens, second only to Murata, with a uniquely diversified base across Taiwan, Japan (via the Pulse Electronics and KEMET acquisitions) and China.

Q1 FY26 revenue came in at NT$38.2B, +22.7% YoY. Net income of NT$8.0B was +44.7% YoY — operating leverage is alive. Gross margin printed 38.1%, operating margin 25.2%, both above pre-cycle peaks. EPS NT$3.90. Management's commentary on AI passive-component demand was the most bullish in five quarters — they explicitly called out tightening across the high-cap grades.

Bull case

What promotes this to Tier A is a clean break and hold above NT$480 with confirmation that the -12.7% pullback was a positioning event, not a fundamental warning. The MLCC up-cycle has historically lasted 6-8 quarters once tightness sets in; we are arguably four quarters into this

one. If AI server unit growth holds at current trajectory and Murata's FY26 +34.8% OP guide proves directionally correct for the industry, Yageo at sub-15x forward earnings is materially cheap relative to its own history.

Gap / bear case

The -12.7% one-day move is the gap. Taiwanese small-mid caps have a habit of front-running cycle peaks, and the chart now needs to show a higher low before fresh capital is committed. Secondary concern: the MLCC cycle has historically been the most violent of the passive components — when it turns, it goes fast.

Trigger to upgrade / downgrade

Upgrade trigger: hold of NT$420-440 zone on the next test, then break NT$481 on volume. Downgrade trigger: loss of NT$380 (50-day MA proxy) closes the door on the current up-cycle.

The trade

*- **Entry zone**: NT$405-425 - **Stop**: NT$378 (2×ATR below current; below pullback low) - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 print mid-July; HBM4 OEM passive-component flow-through - **Conviction**: 7/10*

Layer 5 · 3443 — Global Unichip Corp (GUC)

Thesis

GUC is TSMC's captive ASIC design house — the unsung middleman between hyperscaler chip teams and TSMC's leading-edge fabs. Q1 was a blowout: EPS NT$12.28 beat consensus by 44%, revenue +63% YoY, AI/HPC now 83% of mix. It earns Tier B not Tier A because the +44% EPS surprise has already produced a vertical chart, and ASIC tape-out timing introduces lumpy quarter-to-quarter cyclicality that the current valuation doesn't fully respect.

What it does + financial print

GUC sits at Layer 5 as the design-services pivot between hyperscaler ASIC programs (Google TPU successors, AWS Trainium variants, custom inference silicon) and TSMC's 3nm/2nm production lines. The thesis is simple: every hyperscaler now wants its own silicon, and most lack the deep design IP and TSMC relationships to do it alone — GUC fills that gap, taking design fees up front and royalty-equivalent pulls on production volume.

Q1 FY26 revenue printed NT$11.4B, +63% YoY. EPS NT$12.28 versus consensus that had clustered around NT$8.50 — a 44% beat. FY26 guide raised to NT$52.3B revenue and NT$42.06 EPS. AI/HPC is now 83% of revenue mix versus 64% a year ago, telling you the mobile and consumer ASIC business has been actively de-emphasized.

Bull case

What promotes this to Tier A is a second consecutive quarter of consensus-trouncing beats plus visibility into the 2027 ASIC tape-out pipeline. The structural setup — TSMC's captive design arm, exposed to every hyperscaler custom-silicon program — is genuinely best-of-breed inside the Asian ASIC supply chain.

Gap / bear case

ASIC design-services revenue is famously lumpy because tape-out fees recognise on milestones. The +44% beat was helped by milestone timing, and the next two quarters may show normalisation that hits the multiple even if the secular story is unchanged. NT$2,730 already prices in a multi-year ramp.

Trigger to upgrade / downgrade

Upgrade trigger: Q2 print shows revenue stays above NT$11B (no milestone-timing reversion) plus management nods to additional N2 design wins. Downgrade trigger: any flagged customer push-out on AWS Trainium 3 or Google TPU successor lines.

The trade

*- **Entry zone**: NT$2,500-2,650 - **Stop**: NT$2,350 (technical, prior breakout level) - **Position size**: 1% of NLV - **Catalyst date**: Q2 print early August - **Conviction**: 7/10*

Layer 5 · 3661 — Alchip Technologies

Thesis

Alchip is the pure-play independent AI ASIC design house — the closest thing the public market has to a listed bet on "everyone wants their own chip." Q1 EPS NT$17.23 beat by 16.4%; 3nm production ramp is locked for Q2; AI/HPC is 83% of revenue. It's Tier B because AWS Trainium concentration is the elephant in the room and the N2 ramp is already priced in at NT$5,310.

What it does + financial print

Alchip's stack position is identical to GUC's at Layer 5 — ASIC design services bridging hyperscaler-owned IP and TSMC manufacturing. The differentiator is independence. GUC is a TSMC subsidiary; Alchip can theoretically migrate to Samsung Foundry or Intel 18A if the customer demands it. In practice, almost all of Alchip's leading-edge work runs through TSMC, but the optionality matters at the margin.

Q1 EPS came in at NT$17.23 versus a NT$14.80 consensus — a 16.4% surprise. Management confirmed 3nm AI accelerator production ramps in Q2 and flagged that N2 projects will tape out by end of CY26. The AWS Trainium relationship remains the single largest customer concentration in the book — somewhere north of 40% of revenue runs through that single program.

Bull case

What gets this to Tier A is customer diversification away from AWS Trainium plus a clean N2 tape-out without slippage. Alchip has been signing additional hyperscaler design wins quietly through the year; any disclosed second mega-customer (Meta, Microsoft, ByteDance) would re-rate the multiple cleanly.

Gap / bear case

AWS Trainium concentration is real, and the May 2026 commentary out of AWS on the slower-than-expected Trainium 2 ramp is exactly the kind of headline that hurts Alchip first. NT$5,310 prices in a smooth N2 transition; any tape-out slip would be felt.

Trigger to upgrade / downgrade

Upgrade trigger: disclosed second hyperscaler at >15% of revenue or clean N2 first-silicon news. Downgrade trigger: any AWS Trainium 3 timeline pushout from re:Invent or analyst-day disclosures.

The trade

*- **Entry zone**: NT$4,900-5,150 - **Stop**: NT$4,500 (2×ATR; prior swing low) - **Position size**: 1% of NLV - **Catalyst date**: AWS re:Invent late November; Q2 print early August - **Conviction**: 6/10*

Layer 5 · 6762 — TDK Corporation

Thesis

TDK is the Japanese passive-components conglomerate with three legs of AI exposure: MLCC for AI server boards, lithium polymer batteries for data-center UPS and HDD suspension assemblies for the AI-driven nearline-storage build-out. FY26 was a record — NI +34% YoY. Tier B because the passive-components guide is only +0-3%, which holds back the top-line acceleration narrative even as profitability prints record numbers.

What it does + financial print

TDK is a Layer 5 passive-components house with a legacy HDD-component franchise that has unexpectedly turned into an AI thesis. The HDD suspension-assembly business — once viewed as a structural decline story as SSDs displaced spinning disks — has been bailed out by the AI storage hierarchy: hyperscalers run mountains of cold/warm data on nearline HDDs because SSD is too expensive at exabyte scale, and HDD suspension assemblies are core enabling components. TDK has ~85% global share.

FY26 print: net income +34% YoY, EPS JPY 98.78, record full-year profit. Stock hit a 3-year high May 1 with a +6.3% pre-market move. 1-month return +26.8%, 3-month +37.9%. Battery and HDD suspension were the standout segments. Passive-components guide of only +0-3% for FY27 is the disappointment — management is taking a conservative stance on the front half despite the MLCC tightness their peers (Yageo, Murata) are calling out.

Bull case

What promotes this to Tier A is an upward revision of the passive-components guide at the H1 review in October. If TDK joins Murata at the +30% OP guide level, the stock has fresh legs. The HDD suspension business is also under-appreciated — once the Street internalises that nearline HDD demand is structurally AI-linked, the multiple expands.

Gap / bear case

The +0-3% passives guide is conservative for a reason — TDK customers run higher inventory than Murata's, and management is wary of a Q3-Q4 destock. Stock has moved +27% in 1M; chasing here is mechanical risk.

Trigger to upgrade / downgrade

Upgrade trigger: H1 review (October) raises FY27 passive-components guide above +10%. Downgrade trigger: any inventory-correction commentary out of US/Korean OEMs.

The trade

*- **Entry zone**: JPY 2,700-2,850 - **Stop**: JPY 2,550 (2×ATR; below 50-day MA) - **Position size**: 1% of NLV - **Catalyst date**: H1 FY27 review October; quarterly print early August - **Conviction**: 6/10*

Layer 5 · 6981 — Murata Manufacturing

Thesis

Murata is the global MLCC king — 40% global share, structural moat, the single most important passive-component supplier in the AI server stack. FY26 OP guide came in +34.8% on AI data-center demand — the most bullish industry signal in the cycle. It's Tier B because the live JPY price wasn't surfaced in our scan (needs_manual_check flag) and we don't add Tier A without a clean price print on file.

What it does + financial print

Layer 5 passive components — MLCCs (multi-layer ceramic capacitors), inductors, RF modules, EMI filters. The Murata MLCC franchise is the cleanest moat in passive components: their ultra-small, high-capacitance grades run in every iPhone, every AI accelerator board, every hyperscaler PSU. Substitution risk is low because the dielectric ceramic technology is genuinely proprietary.

FY26 (ending March 2026) revenue ~$12.1B trailing twelve months. Market cap $65.7B at last surfaced print. FY27 operating profit guide came in +34.8% on the back of generative-AI data-center capex tailwinds — that guide rate is the highest Murata has issued in over a decade. Morningstar fair-value estimate sits +22% above last surfaced price.

Bull case

What gets this to Tier A is a fresh JPY price print confirming the current trading band plus a Q1 FY27 (calendar Q2 2026) update confirming the +34.8% OP guide is on track. Murata is structurally the highest-quality AI-passive-components name in the world.

Gap / bear case

Modules and Devices is the laggard segment — recovery there is necessary for the full +34.8% OP guide to materialise. FX-dependency: JPY weakness has flattered the print; if BoJ moves on rates and JPY strengthens, the export-translation tailwind reverses.

Trigger to upgrade / downgrade

Upgrade trigger: clean Q1 FY27 print confirming guide trajectory; fresh price verification. Downgrade trigger: Modules and Devices segment misses by >5%.

The trade

*- **Entry zone**: TBD pending price verification — broker check on JPY level versus 50-day MA before sizing - **Stop**: 2×ATR below entry once verified - **Position size**: 1% of NLV - **Catalyst date**: Q1 FY27 print late July / early August - **Conviction**: 7/10 (would be 8/10 with verified price)*

Layer 5 · APH — Amphenol Corporation

Thesis

Amphenol prints the high-speed interconnects, copper cabling and backplane assemblies that physically wire AI clusters together. Q1 was record — $7.62B revenue, +58% YoY, with AI accounting for virtually all of the IT datacom organic growth. Tier B for two reasons: BofA dropped it from the US 1 list and the live USD price wasn't on file at scan time (needs_manual_check). Otherwise a clean Tier A candidate.

What it does + financial print

Layer 5 connector-and-cable infrastructure — the layer below silicon, above the rack frame. AI clusters require massive amounts of high-speed copper interconnect (NVIDIA NVLink cables, InfiniBand DAC cables, server-to-switch trunks). Amphenol holds the dominant share with Molex (private, inside Koch) as the only meaningful peer. The CommScope acquisition is now closing and adds further AI-datacom capability.

Q1 FY26: revenue $7.62B, +58% YoY (+33% organic). EPS +68%. Book-to-bill 1.24 — a healthy forward signal. Q2 guide $8.1-8.2B revenue and $1.14-1.16 EPS, both above Street. AI is described in the call as "virtually all" of the IT datacom organic growth, with Comms Solutions ~60% of revenue. JPM, GS, Jefferies and Citi all hiked PTs into the print.

Bull case

What gets this to Tier A is a clean price reconciliation plus a successful CommScope integration — the bolt-on adds backbone-cable capacity right when AI-DC interconnect cable demand is structurally tight. Add BofA reinstating the US 1 list rating and the consensus is fully aligned.

Gap / bear case

Comms Solutions concentration at 60% of revenue creates AI-capex-peak exposure. BofA's drop from US 1 reflects exactly this — they think the AI-interconnect cycle is closer to peak than the bulls assume.

Trigger to upgrade / downgrade

Upgrade trigger: BofA reinstatement or any second sell-side US 1 / Conviction List add. Downgrade trigger: CommScope integration disclosure of larger-than-flagged dilution.

The trade

*- **Entry zone**: Pending price verification — look for entry on a 5% pullback from current - **Stop**: 2×ATR below entry once verified - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 print late July; CommScope integration milestones through H2 - **Conviction**: 7/10*

Layer 5 · CDNS — Cadence Design Systems

Thesis

Cadence is one half of the EDA duopoly (Synopsys the other) — every AI accelerator on the planet is designed using Cadence or Synopsys tools, full stop. Q1 +19% YoY, $8B record backlog, AgentStack agentic-AI launch with TSMC partnership confirmed. It's been downgraded from Tier A to Tier B on Hexagon acquisition dilution ($0.28 EPS drag) plus the live China-export-control overhang.

What it does + financial print

Layer 5 design-software pillar. EDA (electronic design automation) tools are the CAD systems that chip designers use — Cadence and Synopsys together hold ~70% global share, with Siemens EDA the only meaningful third. The agentic-EDA transition is the next narrative leg: Cadence launched AgentStack, ViraStack and InnoStack at CadenceLIVE in May, integrating Google Gemini into the ChipStack AI Super Agent. TSMC has partnered with Cadence on next-gen AI chip design flows.

Q1 FY26 revenue $1.474B, +19% YoY. Non-GAAP EPS $1.96. Backlog hit $8B — a record. FY26 guide raised to $6.125-6.225B revenue (+17% YoY) and $7.85-7.95 EPS. The Hexagon Manufacturing Intelligence acquisition closed in the quarter and creates ~$0.28 of EPS dilution; this is the swing factor that pushed CDNS out of Tier A.

Bull case

What promotes this back to Tier A is one full quarter of post-Hexagon-integration earnings (Q3 FY26 print in October) showing the dilution headwind is finite and bounded. The agentic-EDA narrative — AI tools designing AI chips — is structurally the highest-multiple narrative in the EDA space.

Gap / bear case

China export controls are the unhealed wound. Cadence and Synopsys both faced new licensing requirements through the year, and any material restriction of mainland China revenue (~15% of CDNS) would re-base the growth model. Hexagon dilution may also be larger than guided once full purchase-price accounting flows through.

Trigger to upgrade / downgrade

Upgrade trigger: clean Q2/Q3 print showing Hexagon dilution bounded plus no new China licensing restrictions. Downgrade trigger: any new BIS rule restricting EDA exports to China.

The trade

*- **Entry zone**: $325-340 - **Stop**: $305 (2×ATR; prior consolidation base) - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 FY26 print late July - **Conviction**: 7/10*

Layer 5 · GLW — Corning Incorporated

Thesis

Corning re-rated as an AI infrastructure name on May 6 when NVIDIA disclosed a $500M pre-funded warrant plus 15M shares at a $180 strike plus a 10x optical-connectivity production commitment. Q1 EPS $0.43 versus $0.18 — a clean beat. It's Tier B because YTD is +123%, the price wasn't surfaced in our scan, and we're not adding Tier A while the consensus is racing to reprice.

What it does + financial print

Layer 5 optical-glass and connectivity. Corning is the global optical-fibre and high-purity glass leader; the AI thesis here is single-mode optical fibre and connectorised assemblies for hyperscaler campus interconnect. The NVIDIA deal is structural: NVIDIA pre-funded a $500M warrant and committed to 10x current optical-connectivity volumes from Corning, with new manufacturing capacity going into North Carolina and Texas plants.

Q1 FY26 revenue $4.14B, +20% YoY. Net income +136% YoY. EPS $0.43 versus $0.18 prior. The Springboard plan now targets a $20B run rate by year-end, $30B by 2028, $40B by 2030. BofA upgraded. The stock moved +12% on the NVDA deal day. YTD +123%.

Bull case

What promotes this to Tier A is a clean pullback that takes some of the +123% YTD froth out plus a verified live price. The structural setup — NVDA committed off-take, named plant builds,

multi-year revenue ramp — is genuinely Tier A material on fundamentals.

Gap / bear case

The 123% YTD print is the gap. Display Technologies is cyclical and Hemlock semiconductor materials is exposed to solar capex. Execution on NC/TX plant ramps is binary — any slip would be felt.

Trigger to upgrade / downgrade

Upgrade trigger: clean 10-15% pullback plus a Q2 print confirming the Springboard $20B run-rate trajectory. Downgrade trigger: Display Technologies guide cut or any NVDA off-take volume revision.

The trade

*- **Entry zone**: Pending price verification — target a 10-15% pullback from May highs - **Stop**: 2×ATR below entry once verified - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 print late July; NC/TX plant ramp milestones H2 - **Conviction**: 7/10*

Layer 5 · SNPS — Synopsys

Thesis

Synopsys is the other half of the EDA duopoly — Q1 revenue $2.409B beat consensus, +66% YoY including the Ansys contribution. Strong AI-EDA demand, AgentEngineers agentic roadmap unveiled. Tier B because the price was not surfaced clearly in our scan (needs_manual_check) and the May 27 Q2 print needs to confirm Ansys integration is on track. Same China-export overhang as Cadence.

What it does + financial print

Layer 5 EDA — see CDNS write-up for stack context. Synopsys is paired with Cadence in the design-software duopoly. The Ansys acquisition closed last year and the integration is the swing factor: Ansys brings multiphysics simulation (thermal, mechanical, fluid) to the SNPS portfolio, which matters for AI-server cooling and packaging design, but the integration carries dilution risk that is still being absorbed.

Q1 FY26 revenue $2.409B (high end of guide range, beat consensus of $2.39B), +66% YoY including Ansys. Non-GAAP EPS $3.77. Strong AI EDA demand called out — HBM, advanced packaging, power-efficient design. AgentEngineers agentic roadmap announced in May. Stock +15.61% trailing month.

Bull case

What promotes this back to Tier A is the May 27 Q2 print confirming Ansys integration is performing in line — same playbook as CDNS. Synopsys' agentic-EDA narrative (AI tools designing AI chips) is structurally the highest-multiple thread in EDA.

Gap / bear case

Ansys integration risk is the gap. Multiphysics simulation has different sales cycles and customer profiles than EDA; integration friction is possible. China export-control overhang same as CDNS.

Trigger to upgrade / downgrade

Upgrade trigger: May 27 Q2 print confirms Ansys integration on track plus no new China-export licence issues. Downgrade trigger: any disclosed Ansys integration cost overrun or China revenue cut.

The trade

*- **Entry zone**: Pending price verification — bias entry into May 27 pullback - **Stop**: 2×ATR below entry once verified - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 FY26 print May 27 — single most important EDA catalyst this month - **Conviction**: 7/10*

Layer 6 · 4063 — Shin-Etsu Chemical

Thesis

Shin-Etsu is the world's largest silicon-wafer supplier — every leading-edge logic and HBM die starts life on a Shin-Etsu wafer. Q3 FY26 EPS beat by 5.1% but net income was down 8.6% YoY on margin pressure. Tier B because wafer demand is normalising, JPY strength is eroding the export-translation tailwind, and Shin-Etsu is not best-in-layer for the AI thesis specifically — that's GlobalWafers' relative-value story right now.

What it does + financial print

Layer 6 — bulk silicon wafer manufacturing. Shin-Etsu (along with SUMCO and GlobalWafers) sits at the foundational layer of the semiconductor stack: 300mm polished and epitaxial wafers feed every fab. AI demand has materially tightened the leading-edge wafer market — HBM stacks consume more wafer per bit than DDR5, and advanced-logic dies on N3/N2 are larger than the N5 generation.

Q3 FY26: EPS JPY 67.51, beat by 5.1%. Revenue JPY 649.5B, -2.1% YoY. Net income down 8.6% YoY. Wafer demand is normalising — the post-COVID inventory destock is mostly through, but pre-AI peak pricing has not fully returned. Live price JPY 7,476 (May 14).

Bull case

What promotes this to Tier A is a re-acceleration of wafer ASPs on HBM4 ramp plus a JPY weakening cycle. Shin-Etsu has structural moat at the foundational layer of the stack; in a clean cycle re-rate, it works.

Gap / bear case

Net income -8.6% YoY is the print that holds this back. JPY strength erodes export margins. PV and silicone segments remain a drag. For an AI-specific wallet, GlobalWafers (greenfield ramp

leverage) and SUMCO (HBM4 timing) offer cleaner exposure.

Trigger to upgrade / downgrade

Upgrade trigger: Q4 FY26 print shows revenue inflection back to positive YoY plus wafer-ASP guide raised. Downgrade trigger: another quarter of NI decline >5% YoY.

The trade

*- **Entry zone**: JPY 7,200-7,400 - **Stop**: JPY 6,900 (2×ATR; prior swing low) - **Position size**: 1% of NLV - **Catalyst date**: Q4 FY26 print late July / early August - **Conviction**: 6/10*

Layer 6 · 6857 — Advantest Corporation

Thesis

Advantest is the HBM tester monopoly — when an HBM stack comes off a SK Hynix or Samsung packaging line, it gets tested on an Advantest V93000. FY26 guide came in +26% revenue with the SoC tester segment +32%. Tier B is purely tactical: the stock dropped -7.88% on a conservative flat-margin guide despite the revenue print. Fundamentals are Tier A and would re-promote on stabilisation.

What it does + financial print

Layer 6 ATE — semiconductor automated test equipment. The HBM testing flow is genuinely sole-sourced to Advantest in practice (Teradyne plays in SoC and is the second source on memory but lags on HBM). HBM4 brings 16-Hi stacks with new known-good-die testing requirements; Advantest's roadmap is the only one with confirmed coverage.

Q4 FY26: EPS JPY 174.55 versus JPY 138.67 consensus — a clean beat. Revenue JPY 328.07B versus JPY 281.48B consensus. FY26 guide (the year ending March 2027): revenue +26%, SoC tester market +32% on the year. But the operating-margin guide came in flat to slightly down despite the revenue acceleration, and the stock printed -7.88% on the result. Price JPY 26,360. Next earnings July 29.

Bull case

What promotes back to Tier A is a clean stabilisation of the price around JPY 25-27K plus a Q1 FY27 print confirming the +26% revenue trajectory is on. The HBM testing thesis is single-best inside the AI test stack.

Gap / bear case

The conservative flat-margin guide is the gap. Customer concentration on the TSMC/NVIDIA/HBM trio creates beta to any single-name AI capex revision. The -7.88% drop signals positioning fragility.

Trigger to upgrade / downgrade

Upgrade trigger: hold of JPY 25,500 zone plus July 29 Q1 print confirming margin trajectory. Downgrade trigger: loss of JPY 24,000 (technical breakdown level).

The trade

*- **Entry zone**: JPY 25,500-26,500 - **Stop**: JPY 24,000 (technical; below post-print swing) - **Position size**: 2% of NLV - **Catalyst date**: Q1 FY27 print July 29 - **Conviction**: 8/10*

Layer 6 · FORM — FormFactor

Thesis

FormFactor prints the probe cards that test wafer-level HBM and leading-edge logic before dicing — record DRAM probe-card revenue $82.9M in Q1, +69.5% YoY on HBM. Q2 guide ahead at $240M ±$5M. Tier B because the live USD price wasn't surfaced (needs_manual_check) and we don't add Tier A without a clean price reconciliation. Fundamentals are clean Tier A.

What it does + financial print

Layer 6 wafer-level test consumables. Probe cards are the physical card-and-needle assemblies that touch down on individual dies during wafer-level test — different from Advantest's testers, which are the test chassis. FormFactor leads in DRAM/HBM probe cards; SK Hynix, Micron and Samsung HBM lines all run on FormFactor. The HBM4 inflection drives a re-platform of probe-card design; FormFactor has been quietly winning second-customer ramps on its Smart Matrix tech.

Q1 FY26 revenue $226.1M, +32% YoY. EPS $0.26 (beat). DRAM segment revenue $82.9M, +69.5% YoY — a record. Q2 guide $240M ±$5M, gross margin 49.5%, EPS $0.61. HBM probe-card business +50% H1 25 to H1 26.

Bull case

What promotes to Tier A is a verified live price plus second consecutive quarter of record DRAM probe-card revenue. FormFactor is the cleanest pure-play probe-card asset in the test layer, period.

Gap / bear case

Customer concentration is the gap — DRAM/HBM probe-card revenue concentrates in three customers (SK Hynix, Micron, Samsung). Foundry & Logic probe market is also lumpy with leading-edge node timing.

Trigger to upgrade / downgrade

Upgrade trigger: price verified plus Q2 print confirming $240M guide hit. Downgrade trigger: any pushout on HBM4 qualification timing from SK Hynix or Samsung.

The trade

- ***Entry zone**: Pending price verification* - ***Stop**: 2×ATR below entry once verified* - ***Position size**: 1.5% of NLV* - ***Catalyst date**: Q2 print late July; HBM4 qualification milestones through Q3* - ***Conviction**: 7/10*

Layer 6 · TER — Teradyne

Thesis

Teradyne is the SoC and memory tester #2 globally — Q1 revenue +87% YoY at $1.282B, a record print. Tier B because the Q2 guide came in soft and the stock dropped -19% post-earnings; the technical needs to rebuild before fresh capital. Fundamentals — AI accelerator + HBM tester demand — are Tier A material.

What it does + financial print

Layer 6 ATE — Teradyne plays in SoC testers (versus Advantest the leader in HBM testers; the two split memory and trade share in compute). The AI thesis here is SoC test capacity for AI accelerators — every NVIDIA Blackwell, every AMD MI400, every custom ASIC has to test through someone's tester, and Teradyne owns meaningful share at the SoC end.

Q1 FY26 revenue $1.282B, +87% YoY (beat by 5.45%). Net income $398.9M. EPS $2.55, beat. Q2 guide $1.15-1.25B revenue, $1.86-2.15 EPS — sequentially soft, which is why the stock dropped 19% on the print. AI accelerator and HBM tester demand are the stated growth drivers in the SoC segment.

Bull case

What promotes to Tier A is a clean stabilisation plus a Q2 print that beats the soft guide. Teradyne is structurally exposed to the AI accelerator volume ramp — the +87% YoY print isn't a one-off, it's the new run rate.

Gap / bear case

The Q2 guide softness is the gap, and the -19% drop confirms market disappointment. Tester demand is cyclical quarter-to-quarter even within a structural up-cycle; sequencing matters.

Trigger to upgrade / downgrade

Upgrade trigger: reclaim and hold $370 plus Q2 print July 28 beats the soft guide. Downgrade trigger: loss of $330 (post-drop consolidation base).

The trade

- ***Entry zone**: $340-365* - ***Stop**: $325 (2×ATR; below post-print low)* - ***Position size**: 1.5% of NLV*
- ***Catalyst date**: Q2 print July 28* - ***Conviction**: 7/10*

Layer 7 · CEG — Constellation Energy

Thesis

Constellation is the largest US nuclear operator and the cleanest hyperscaler PPA bench-mark. Q1 FY26 GAAP EPS $4.49 (huge Calpine-acquisition boost), revenue +64% YoY, FY26 adj EPS $11-12 reaffirmed plus 20% CAGR to 2029. Fundamentals are Tier A. It's tactically Tier B because the May 11 daily brief flagged a STRONG_EXIT technical signal and the price has continued to soften from $275.26 to $267.20 — wait for a base.

What it does + financial print

Layer 7 baseload nuclear generation. CEG operates the largest US nuclear fleet (~22 GW) and was the structural winner of the IRA Production Tax Credit (PTC) framework, which fixed nuclear economics for the next decade. The Calpine acquisition added 27 GW of gas + geothermal, broadening the product set. Hyperscaler PPAs — Microsoft 20-year Three Mile Island restart, others under wraps — anchor the AI-DC power thesis.

Q1 FY26: GAAP EPS $4.49 versus $0.38 prior year (huge step-up from Calpine consolidation plus PTC monetisation). Adjusted EPS $2.74. Revenue +64% YoY. FY26 adjusted EPS $11-12 reaffirmed; management reaffirmed a 20% adjusted-EPS CAGR through 2029. Live price $267.20 versus a 52-week range of $243-413 — currently in the lower third.

Bull case

What gets this back to Tier A is a clean technical reversal — close back above the 50-day MA on volume, then a higher-high higher-low base formation. Fundamentals are unambiguous: nuclear scarcity premium plus hyperscaler PPAs plus IRA PTC monetisation is a multi-year compounding setup.

Gap / bear case

The STRONG_EXIT signal is the gap — May 11 brief flagged it and the subsequent price action has confirmed downside pressure. Calpine integration execution is the secondary risk. Trading near the lower end of the 52-week range is a position-sizing constraint.

Trigger to upgrade / downgrade

Upgrade trigger: technical reversal — reclaim of 50-day MA on volume and base-formation pattern (higher low confirmed). Downgrade trigger: loss of $243 (52-week low).

The trade

*- **Entry zone**: $260-275 only on technical reversal confirmation; do NOT add in the falling-knife window - **Stop**: $240 (below 52-week low) - **Position size**: 2% of NLV (scaled — full size only on Tier A re-promote) - **Catalyst date**: Q2 print early August; Calpine integration milestones - **Conviction**: 7/10 fundamentals, 5/10 tactical*

Layer 7 · VST — Vistra Corp.

Thesis

Vistra is the second IPP merchant-power + nuclear AI-DC story — Q1 revenue $5.64B beat consensus by 8%, FY26 EPS estimate $11.22 on AI data centre demand, Cogentrix 5,500 MW gas portfolio acquisition closing in 2H 2026. Like CEG, fundamentals are Tier A but the May 11 daily brief flagged a STRONG_EXIT technical signal not yet reversed — tactical B until the chart stabilises.

What it does + financial print

Layer 7 IPP merchant power. Vistra is the largest competitive power generator in Texas with significant nuclear (Comanche Peak), coal-to-gas, and gas-peaker assets. The AI thesis: long-term PPAs with Amazon and Meta on a substantial fraction of EBITDA, plus the Cogentrix acquisition adds 5,500 MW of gas portfolio in load-pocket markets (PJM, NEPOOL) where AI data centres are clustered.

Q1 FY26: revenue $5.64B versus $5.24B consensus (beat by 8%). Adjusted EBITDA $1.49B. FY26 EPS estimate $11.22 on AI demand. Cogentrix 5,500 MW gas portfolio acquisition closes 2H 2026. Long-term PPAs with Amazon and Meta. 60%+ of EBITDA secured by contracts — high contracted-coverage gives downside protection. Live price $139.10 (May 17). 52-week range broadly intact but momentum has cooled.

Bull case

What gets this back to Tier A is the technical reversal — same playbook as CEG. Reclaim of the 50-day MA on volume, base-formation pattern, then add. Fundamentals are unambiguous and the Cogentrix close is a structural step-up.

Gap / bear case

The STRONG_EXIT signal from May 11 hasn't been reversed by price action. Cogentrix integration execution is the operational risk. Natural gas pricing is the cyclical-input variable.

Trigger to upgrade / downgrade

Upgrade trigger: technical reversal — reclaim of 50-day MA on volume plus higher-low base. Downgrade trigger: loss of $125 (technical breakdown level).

The trade

*- **Entry zone**: $135-145 only on technical reversal; do NOT add in falling-knife window - **Stop**: $125 (technical breakdown) - **Position size**: 2% of NLV (scaled — full size on Tier A re-promote) - **Catalyst date**: Q2 print late July; Cogentrix closing milestones 2H - **Conviction**: 7/10 fundamentals, 5/10 tactical*

End Tier B Part 2 — Layers 5-7, 22 names.

Layer 7 · BE — Bloom Energy

Thesis

Bloom is the fuel-cell distributed-power play for AI data centres — Oracle's 2.45 GW Project Jupiter NM site is sole-sourced to Bloom's solid-oxide fuel cells. Q1 revenue +130% YoY, product revenue +208% YoY, swung to GAAP profit. Tier B because the stock is +1,511% on the AI-power re-rate and the valuation premium is extreme; A on the next pullback.

What it does + financial print

Layer 7 power generation. Bloom Energy makes solid-oxide fuel cells (SOFC) that run on natural gas, biogas or hydrogen. AI data centres are the breakthrough vertical: hyperscalers are out of grid headroom in core markets (Virginia, Texas, Arizona) and behind-the-meter power generation is the bridge solution. Bloom is the only at-scale SOFC vendor with a deployable product today.

Q1 FY26: revenue $751M, +130% YoY. Product revenue +208% YoY. GAAP net income $70.7M — swung to profitability. Operating cash flow $73.6M (positive). FY26 guide raised to $3.4-3.8B from $3.1-3.3B. Oracle expanded the partnership to 2.8 GW total fuel-cell capacity, with Bloom sole-sourcing the 2.45 GW Project Jupiter AI campus in New Mexico. Live price $274.71 (May 17).

Bull case

What promotes to Tier A is a clean pullback to $230-250 zone plus a second hyperscaler signing for behind-the-meter SOFC at GW-scale. The Oracle deal is structural — it validates SOFC as a deployable AI-DC power solution and creates a reference design.

Gap / bear case

The +1,511% rerating is the gap. Natural gas feedstock dependency creates emissions-policy risk; if any major hyperscaler ESG team blocks SOFC inclusion in green-power frameworks, the secondary-customer pipeline thins. Oracle concentration is real.

Trigger to upgrade / downgrade

Upgrade trigger: second hyperscaler GW-scale signing (Meta, Google, Microsoft). Downgrade trigger: any GW-scale customer pulling SOFC for emissions reasons.

The trade

*- **Entry zone**: $230-255 - **Stop**: $210 (2×ATR; below prior consolidation) - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 print July 30; Oracle Jupiter milestone updates - **Conviction**: 7/10*

Layer 7 · MOD — Modine Manufacturing

Thesis

Modine pivoted from automotive thermal to data-centre liquid cooling — DC cooling revenue +78% last quarter, FY26 guide raised to 20-25% top-line growth, new Data Centers segment created April 1 under Art Laszlo. Tier B because the stock is +89% YTD at ~35x EV/EBITDA pro forma and the May 26 Q4 print is a binary catalyst. A on a clean print + multiple compression.

What it does + financial print

Layer 7 thermal management — Modine makes liquid-cooling and HVAC infrastructure for data centres. The AI thesis is direct-to-chip liquid cooling: as rack thermal density goes from 30 kW (air-cooled limit) to 130 kW (NVIDIA NVL72 class), every hyperscaler is forced into liquid cooling, and Modine has been one of the early commercial-scale players with the Airedale platform.

Q3 FY26: Data Center cooling revenue +78% YoY. FY26 sales growth guide raised to 20-25%. Data-centre cooling is now 25% of revenue, up from 4% prior. The Franklin Wisconsin plant expansion adds >$100M of additional capacity. Dedicated Data Centers segment was carved out April 1 with Art Laszlo as president — operational separation telegraphs management confidence. Live price $271.26 (May 15) — note the -7.15% drop on May 15 from sector profit-taking.

Bull case

What promotes to Tier A is the May 26 Q4 print confirming the 20-25% guide is achievable, combined with order-book disclosures showing the Franklin plant is sold out. The thesis is structural — direct-to-chip cooling is the only way to handle current-generation AI rack thermals.

Gap / bear case

35x EV/EBITDA pro forma is the gap. The May 15 -7.15% drop signals positioning sensitivity. Modine still has substantial legacy automotive and HVAC business — segment-level dilution from those slower-growth legs caps the multiple.

Trigger to upgrade / downgrade

Upgrade trigger: clean May 26 Q4 print plus FY27 guide confirming DC cooling momentum. Downgrade trigger: any sequential slowdown in DC cooling segment revenue.

The trade

- ***Entry zone**: $255-275 (own positions through May 26 print; add on any post-print weakness)* - ***Stop**: $235 (technical; below recent swing low)* - ***Position size**: 1.5% of NLV* - ***Catalyst date**: Q4 FY26 print May 26 — within 8 days* - ***Conviction**: 7/10*

Layer 7 · POWL — Powell Industries

Thesis

Powell makes medium-voltage switchgear and electrical buildings — the substation-grade gear that powers data centre campuses. Q2 EPS and revenue both missed but new orders surged +97%

to $490M and backlog hit $1.8B (+33% YoY). A $400M+ mega data-centre order announced post-Q2 is the largest in Powell's history. Tier B because the stock trades at 58x earnings and the print itself missed.

What it does + financial print

Layer 7 electrical infrastructure. Powell sits between the utility-scale grid layer and the building-level UPS layer — they supply switchgear, motor control centres, packaged power buildings. AI data centre campuses need vastly more switchgear than conventional commercial real estate; a single 1 GW campus may consume 200+ medium-voltage switchgear assemblies.

Q2 FY26: EPS $1.25 (consensus miss -9.99%). Revenue $296.6M (miss). However: new orders +97% YoY to $490M, backlog $1.8B (+33% YoY). The post-Q2 May 4 announcement of a $400M+ mega data-centre order — Powell's largest ever — sent the stock +11-15% on May 5. Live price $292.65 (May 15).

Bull case

What promotes to Tier A is one clean print where the order book translates into beat-revenue. Backlog of $1.8B at the current revenue run rate is ~6 quarters of visibility; that should convert at some point. The $400M order alone is a multi-quarter revenue stream.

Gap / bear case

58x earnings is the gap, and the Q2 miss says order-book conversion is more uneven than the order-book growth implies. Execution risk on the mega-order is real — Powell is doubling capacity to absorb it.

Trigger to upgrade / downgrade

Upgrade trigger: Q3 print beats top and bottom line plus mega-order revenue conversion disclosed. Downgrade trigger: any execution-flag or order push-out.

The trade

*- **Entry zone**: $270-290 - **Stop**: $255 (2×ATR; below post-Q2 base) - **Position size**: 1% of NLV - **Catalyst date**: Q3 FY26 print early August - **Conviction**: 6/10*

Layer 7 · PWR — Quanta Services

Thesis

Quanta is the largest US electrical-infrastructure construction firm — grid build-out, substations, transmission, AI data centre electrical packages. Q1 adjusted EPS $2.68 versus $2.06 consensus — clean beat. FY26 raised to $13.55-14.25 adj EPS. Backlog $48.5B with Electric Power at $40.1B. Tier B because the stock has run +127% in a year and the valuation is rich.

What it does + financial print

Layer 7 grid + transmission construction. Quanta builds the high-voltage transmission lines, substations and electrical infrastructure that connect generators to load. AI data centre growth is grid-limited in core markets, and Quanta is the largest at-scale contractor. The Investor Day outlined a $2.4T infrastructure opportunity through 2030.

Q1 FY26: adjusted EPS $2.68 versus $2.06 consensus. Revenue $7.9B. FY26 adj EPS guide raised to $13.55-14.25 from $13.09 consensus. Backlog $48.5B total, with Electric Power at $40.1B. Quanta plans to roughly double mfg/fabrication footprint for modular DC infrastructure (Cubica modular substation product). Live price $769.99 (May 17). 1Y return +127%.

Bull case

What promotes to Tier A is a clean pullback of 10-15% from current that takes some valuation excess out, plus a second consecutive raise of FY26 guide on the Q2 print. The thesis is structural — every utility CIP plan in the US shows multi-year transmission upgrades; Quanta executes them.

Gap / bear case

+127% 1Y is the gap. Labour-scarcity (electricians) and copper cost inflation are the operational risks. Backlog conversion timing is the cyclical variable.

Trigger to upgrade / downgrade

Upgrade trigger: 10-15% pullback plus Q2 print confirming guide trajectory. Downgrade trigger: copper input cost spike or any flagged labour-availability constraint.

The trade

*- **Entry zone**: $700-740 - **Stop**: $665 (2×ATR; below 50-day MA proxy) - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 print early August - **Conviction**: 7/10*

Layer 7 · STRL — Sterling Infrastructure

Thesis

Sterling smashed Q1 — revenue +92%, EPS +120%, E-Infrastructure (DC + semi fab) segment +174%. FY26 guide raised to EPS $18.40-19.05 and revenue $3.70-3.80B (+51%). Backlog $3.80B (+78% YoY). Tier B for one reason: PE 74.55 is extreme-overbought territory and the stock is at 52-week-high range. Best-in-class print, worst-in-class entry.

What it does + financial print

Layer 7 site-preparation and civil infrastructure for data centres and semi fabs — earthwork, foundations, electrical infrastructure pads. E-Infrastructure is the AI-DC pure-play segment; mega-projects like the TSMC Arizona fabs, large hyperscaler DC campuses, etc all go through site-preparation contractors before vertical construction.

Q1 FY26: adjusted EPS $3.59 versus $2.19 consensus (+64% surprise). Revenue $825.7M, +92% YoY. E-Infrastructure +174% YoY tied to data centre + semi fab buildouts. FY26 raised to EPS $18.40-19.05, revenue $3.70-3.80B (+51%). Signed backlog $3.80B (+78% YoY). The CEC acquisition closed. Live price $842.00 (May 17). Stock surged 51% on the May 5 print.

Bull case

What promotes to Tier A is a healthy 15-20% pullback that flushes the parabolic move plus a second beat-and-raise. The structural setup — DC and semi-fab site work is genuinely capacity-constrained — is best-in-Layer-7.

Gap / bear case

PE 74.55 and a vertical chart is the gap. Concentration in DC and semi-fab capex cycle creates beta to any AI-capex revision. Stock at 52-week high makes entry hazardous.

Trigger to upgrade / downgrade

Upgrade trigger: 15-20% pullback plus Q2 print confirming guide. Downgrade trigger: any sequential decline in E-Infrastructure segment revenue.

The trade

*- **Entry zone**: $700-760 (only on meaningful pullback) - **Stop**: $650 (2×ATR; below technical breakout) - **Position size**: 1% of NLV (small until pullback confirms) - **Catalyst date**: Q2 print early August - **Conviction**: 6/10*

Layer 7 · VICR — Vicor Corporation

Thesis

Vicor makes the 800V-to-48V and 48V-to-PoL DC-DC power conversion modules that are core to NVIDIA's Rubin-era 800V power architecture. Q1 EPS $0.44 beat by 18.92%, revenue +20.2% YoY, book-to-bill >2, backlog +70% QoQ. Tier B because the price wasn't surfaced (needs_manual_check) plus heavy insider selling reported mid-May. Fundamentals Tier A.

What it does + financial print

Layer 7 power conversion. Vicor's tech is the modular vertical-power-delivery architecture — Gen 5 vertical converters that sit directly under or alongside the AI accelerator package, delivering current at extreme density. The NVIDIA Rubin design win is the structural anchor: 800V bus to the rack, then Vicor modules step it down to 48V and then to PoL on the GPU.

Q1 FY26: revenue $112.97M, +20.2% YoY. EPS $0.44, beat by 18.92%. Gross margin 55.2%. Book-to-bill >2; backlog $300.6M, +70% QoQ. Q2 guide $126M; FY26 ~$570M. Gen 5 vertical and 800V-to-48V converters target a $5B+ AI compute power market by 2027. Stock +17.88% on Q1 print. 1Y return 413%. Insider selling reported in mid-May. Market cap ~$11.62B implies ~$260 area.

Bull case

What promotes to Tier A is a verified live price plus an NVIDIA Rubin volume-ramp disclosure that anchors the 2026-2027 revenue ramp. The tech is genuinely best-in-class in vertical AI power.

Gap / bear case

Heavy insider selling mid-May is a real signal — when insiders are selling into a 413% 1Y move, position sizing matters. Customer concentration on NVIDIA Rubin timing is binary; any pushout would hurt.

Trigger to upgrade / downgrade

Upgrade trigger: NVIDIA Rubin volume confirmation plus insider-selling pause. Downgrade trigger: any pushout on Rubin 800V rollout timing.

The trade

*- **Entry zone**: Pending price verification — target ~$250 area on broker check - **Stop**: 2×ATR below entry once verified - **Position size**: 1% of NLV - **Catalyst date**: Q2 print late July; NVIDIA Rubin product disclosures Computex / GTC - **Conviction**: 7/10*

Layer 7 · VRT — Vertiv Holdings

Thesis

Vertiv is the integrated AI-rack thermal and power infrastructure leader — liquid cooling, busway, modular UPS, prefabricated AI factory enclosures. Q1 adjusted op margin +430 bps to 20.8%; backlog doubled to $15B+. FY26 raised +30% organic with 51% earnings growth. Tier B because the stock has rerated significantly and the valuation premium leaves limited margin for execution surprise.

What it does + financial print

Layer 7 integrated AI-DC infrastructure. Vertiv plays across thermal management (CRAH/CRAC, liquid cooling), power infrastructure (PDU, busway, UPS) and the modular prefabricated AI factory build-out. The thesis is that hyperscalers are buying prefab modular DC at scale because the speed-to-power-on cycle is significantly faster than custom build, and Vertiv has the most integrated portfolio.

Q1 FY26: revenue $2.65B. Adjusted EPS +83% YoY; diluted EPS +136% YoY. Adjusted op margin 20.8% (+430 bps). FY26 guide raised to $13.5-14.0B (+30% organic) with 51% earnings growth. Backlog doubled to $15B+ — that's roughly 12-18 months forward revenue. Recent +12.76% day on earnings beat. Live price $339.00 (May 8).

Bull case

What promotes to Tier A is one full quarter of post-rerate consolidation plus a Q2 print confirming the 51% earnings growth trajectory. The structural setup — liquid cooling adoption is essentially

mandatory for current-generation AI racks — is among the cleanest Layer 7 stories.

Gap / bear case

The valuation premium is the gap. Supply-chain pressure on liquid-cooling components (CDUs, manifolds, secondary loops) is real and could pressure margin. Hyperscaler customer concentration is the cyclical exposure.

Trigger to upgrade / downgrade

Upgrade trigger: clean consolidation plus Q2 print confirming guide. Downgrade trigger: any flagged hyperscaler capex push-out on cooling deployments.

The trade

*- **Entry zone**: $310-335 - **Stop**: $290 (2×ATR; below post-rerate consolidation) - **Position size**: 1.5% of NLV - **Catalyst date**: Q2 print late July - **Conviction**: 7/10*

Part 5 — Tier C · Watch / Pass

Sorted by AI stack layer ascending; within each layer by conviction rank. 31 names. Genuine AI exposure but case is early, speculative, or contingent on a specific catalyst we are tracking.

Layer 0 · CRWV — CoreWeave

Live: $106.92 USD · **Earnings (last):** Q1 26 — rev beat / EPS miss; guide cut · **Next earnings:** Aug 2026 · **Conviction:** 3/10

Bottleneck context

CoreWeave is a GPU-as-a-service hyperscaler-lite — datacenter operator running NVIDIA H100/H200/GB200 capacity on contract for Microsoft, Meta and OpenAI. Sits at the very top of the stack as a buyer of NVIDIA silicon and a seller of compute hours. The “moat” is allocation priority from NVIDIA plus signed multi-year take-or-pay contracts. The vulnerability is that it has neither power nor land nor unique IP — it is balance-sheet leverage on rented capacity.

Why Tier C

Earnings have missed five consecutive prints. Q1 26 revenue grew +106% YoY but the net loss widened to $740M, FCF break-even was pushed out, capex was raised, and management cut forward guide. This is a dramatic deterioration from prior rankings where CRWV sat in Bucket A on momentum and customer-list optics. The thesis has shifted from “AI hyperscaler proxy” to “leveraged GPU lessor with deteriorating unit economics.” Single-customer concentration (Microsoft >60%) compounds the risk.

Watch trigger

A clean FCF inflection print — positive operating FCF with capex normalized — plus customer diversification beyond Microsoft would promote CRWV back to Tier B.

Position guidance

No position. Wait for FCF break-even signal; do not chase rebounds driven by NVIDIA halo trades.

Layer 4 · 033640 — Nepes

Live: KRW 32,950 · **Earnings (last):** Q1 26 needs manual check · **Next earnings:** Q2 26 print pending · **Conviction:** 3/10

Bottleneck context

Nepes is a Korean KOSDAQ-listed back-end packaging specialist with exposure to fan-out wafer-level packaging (FOWLP) and panel-level packaging. In theory it sits at Layer 4 (advanced packaging) alongside Hanmi, ASMPT and Amkor. In practice it is a small-cap supplier with limited HBM exposure and limited English-language disclosure.

Why Tier C

Price sources disagree materially — KRW 20.1k on one feed vs KRW 32.95k on another — indicating either a corporate action or a stale quote, and the disclosure quality flag is on `needs_manual_check`. Beyond the data hygiene problem, FOWLP/panel-level is a slow-moving thesis vs the urgent HBM advanced-packaging story dominating 2026. Limited liquidity and limited English coverage mean upside surprise is unlikely to be capturable.

Watch trigger

Confirmed price source plus a clean Q2 print showing material HBM-adjacent revenue would warrant a Tier B look.

Position guidance

No position. Data quality alone disqualifies.

Layer 4 · 042700 — Hanmi Semiconductor

Live: KRW 369,000 · **Earnings (last):** Q1 26 — miss; SK Hynix share loss accelerating · **Next earnings:** Aug 2026 · **Conviction:** 4/10

Bottleneck context

Hanmi is the legacy TC (thermo-compression) bonder supplier into SK Hynix HBM lines. For two years it traded as the single most direct HBM picks-and-shovels name in Korea. The thesis was: SK Hynix wins HBM, Hanmi wins the bonder socket, Hanmi compounds at a 50%+ rev CAGR.

Why Tier C

Thesis broken. ASMPT confirmed share take in TC bonding at SK Hynix; the HBM4 16-Hi roadmap requires hybrid bonding where Hanmi has no incumbent position. The stock dropped -10% in a single session on the confirmation. The Hanwha Semitech patent dispute is the only remaining bullish optionality and it is binary and slow. With consensus EPS now being revised down quarter-on-quarter, the multiple compression is not yet finished.

Watch trigger

A favourable Hanwha patent settlement plus a credible hybrid-bonding roadmap from Hanmi would warrant re-engagement. Absent that, watch for an SK Hynix share-recovery surprise.

Position guidance

No position. If holders are still long from higher levels, this is a “trim into strength” name, not a buy-the-dip name.

Layer 4 · 110990 — DIT Corp

Live: KRW 15,790 · **Earnings (last):** Q1 26 needs manual check · **Next earnings:** Pending · **Conviction:** 2/10

Bottleneck context

DIT is a small-cap KOSDAQ packaging-equipment / test-handler vendor. v3 universe had it tagged as a Korean back-end picks-and-shovels candidate but the actual product line is poorly disclosed in English-language sources.

Why Tier C

v3 already flagged DIT as SKIP. -37.83% on a 52-week basis confirms the market is not buying any thesis. Small-cap KOSDAQ liquidity, opaque product mix, and no near-term identifiable catalyst put this firmly in the watch-only bucket.

Watch trigger

A confirmed HBM tool design-win at SK Hynix or Samsung disclosed in either Korean financial press or company filings would re-open the case.

Position guidance

No position. If held, trim — the price action and disclosure quality both argue against patience.

Layer 4 · 489790 — Hanwha Vision (Semitech)

Live: KRW 86,500 · **Earnings (last):** Q1 26 needs manual check · **Next earnings:** Pending · **Conviction:** 1/10

Bottleneck context

Flag: this ticker is wrong. 489790.KS is Hanwha Vision — a security-camera and SMT chip-mounter manufacturer — NOT Hanwha Semitech, the private TC-bonder subsidiary of Hanwha Aerospace that has been the subject of HBM bonder commentary. Hanwha Semitech is not separately listed. Any thesis built around this ticker as an HBM packaging play is built on a misidentification.

Why Tier C

Auto-Tier C on misidentification. Even if Hanwha Vision the camera business is fundamentally sound, it does not belong in an AI-stack universe at Layer 4. Including it dilutes the framework.

Watch trigger

Re-enter the AI universe ONLY if Hanwha Semitech IPOs or is spun off as a separately listed entity. Treat any rally driven by HBM headlines on this ticker as a tape error.

Position guidance

No position. Exclude from coverage.

Layer 4 · IQE — IQE plc

Live: 46.00 GBX · **Earnings (last):** 2024 — rev +2.4%; losses widened to -£38M · **Next earnings:** May 20 FY25 results · **Conviction:** 3/10

Bottleneck context

IQE is the UK-listed compound-semiconductor epitaxial-wafer specialist (GaAs, InP, GaN). Theoretically a picks-and-shovels Layer 4 name into VCSELs (datacom optics), RF (handset PAs) and emerging power-GaN. The pitch is that AI optical interconnect demand for InP epi-wafers eventually shows up here.

Why Tier C

Loss-making, small-cap (sub-£200M market cap), low-liquidity, and the May 20 FY25 print is pending. Recent quarters showed losses widening to -£38M against modest top-line growth. RF/handset cyclicality dominates the P&L today; AI optical contribution is too small to move the numbers.

Watch trigger

A clean FY25 print showing InP/datacom revenue inflection — or a strategic disposal of the RF business — would warrant fresh look. Watch the May 20 release closely.

Position guidance

No position into print; speculative low-conviction post-print only on confirmation of optical mix shift.

Layer 4 · LWLG — Lightwave Logic

Live: \$13.65 USD · **Earnings (last):** Q1 26 rev +27% YoY but loss \$6.3M · **Next earnings:** Aug 2026 · **Conviction:** 2/10

Bottleneck context

Lightwave Logic develops electro-optic polymers for high-speed modulators — the active component in optical interconnects. In theory this is a Layer 4 photonics enabler that could disrupt lithium niobate and silicon photonics in 400G/800G/1.6T pluggables and CPO. In practice, after ~20 years of development, commercial revenue is still measured in single-digit millions.

Why Tier C

Pre-revenue at scale. The Q1 rev +27% YoY headline obscures the absolute size — revenue remains tiny against a \$6.3M quarterly loss. Cash burn vs commercial-traction gap is the binary risk: foundry capacity constraints are cited as the bottleneck to scaling. Auto-C on speculative-photonics filter. Stock trades like a momentum vehicle on every CPO headline regardless of LWLG's actual involvement.

Watch trigger

A signed, named-customer, multi-million-dollar modulator volume order (not an evaluation or MOU) with delivery in current fiscal year.

Position guidance

No position. If sized as lottery-ticket exposure, hold size under 0.5% NAV and accept full loss potential.

Layer 4 · POET — POET Technologies

Live: $20.57 USD · **Earnings (last):** Q1 26 EPS miss · **Next earnings:** Aug 2026 · **Conviction:** 3/10

Bottleneck context

POET produces optical engines and interposers for high-speed datacom — pluggable transceivers and CPO building blocks. Optical interposer platform is the differentiated IP. Layer 4 photonics enabler with NVIDIA/AI-DC narrative tailwind.

Why Tier C

Pre-revenue at scale — Q1 revenue was $503K against a $12.3M net loss. Recent capital raise was dilutive. Stock is momentum-driven; the Lumilens $50M EIO engine order is real but back-end-weighted and execution risk is high. Insider activity and equity issuance patterns are not consistent with a self-sustaining commercial ramp. The valuation has decoupled from the financial reality.

Watch trigger

Demonstrated quarterly revenue >$10M with positive gross margin from named hyperscaler shipment volumes — not engineering samples.

Position guidance

No position. Trade-only candidate for momentum players; not a portfolio name.

Layer 4 · SIVE — Sivers Semiconductors

Live: SEK 55.70 · **Earnings (last):** Q4 FY25 loss SEK -23.67M · **Next earnings:** May 29 (Q1 FY26) · **Conviction:** 2/10

Bottleneck context

Sivers makes InP lasers and photonics ICs for datacom and 5G. Theoretical CPO play: indium-phosphide laser sources are the bottleneck input into co-packaged optics. Sub-$2B market cap, Stockholm-listed, illiquid.

Why Tier C

Speculative auto-C. Negative TTM EPS, persistent operating losses (SEK -23.67M last quarter), sub-$2B mcap and the May 29 print is the gating event. The InP/CPO thesis is real at the industry level but Sivers is one of several second-tier vendors competing for design-in slots dominated by Coherent, Lumentum and Marvell-Inphi.

Watch trigger

A confirmed CPO design-win at a tier-1 hyperscaler with material revenue contribution disclosed in subsequent earnings.

Position guidance

No position. Watch the May 29 print for any commercial inflection; speculate only on tier-1 design-win confirmation.

Layer 4 · SOI — Soitec

Live: EUR 148.05 · **Earnings (last):** Rev -10.2% YoY; NI -23.5% · **Next earnings:** May 27 FY release · **Conviction:** 3/10

Bottleneck context

Soitec manufactures engineered SOI wafers (FD-SOI, RF-SOI, POI) — specialty substrates used in RF, automotive and emerging power applications. Layer 4 enabler with a niche-monopoly profile in its sub-segments. The AI narrative pull is photonics-SOI and RF-SOI into 5G/optical front-end.

Why Tier C

Severe valuation/fundamentals divergence. YTD +447% against revenue -10.2% YoY and net income -23.5%. The stock has rerated on AI-themed flows while the underlying business is contracting. Flipped from Tier B to Tier C on this gap. The May 27 FY release is the gating catalyst — into the print this is a coin flip.

Watch trigger

A clean FY print with explicit AI/photonics revenue contribution call-out plus FY26 guide reaffirming or raising. Absent that, gravity wins.

Position guidance

No position into print. Post-print, only re-engage on revenue inflection — do NOT chase a "guide-in-line" rally.

Layer 5 · 4004 — Resonac Holdings

Live: JPY 17,200 · **Earnings (last):** Mixed Q4 · **Next earnings:** Aug 11 · **Conviction:** 4/10

Bottleneck context

Resonac is the merged Showa Denko / Hitachi Chemical entity — global leader in CMP slurries, photoresist ancillaries, and back-end packaging materials (die-attach films, EMC). Genuine Layer

5 materials franchise with structural AI/HBM packaging leverage. Strong long-term thesis on advanced-packaging material intensity.

Why Tier C

Tactical, not structural. Sharp -7.97% intraday pullback signals positioning unwind. Normalized P/E of 76.81 is stretched even for a quality compounder. The August 11 earnings is the next catalyst — far away, no near-term news flow to support the multiple. Flipped from Tier B to Tier C on stretched valuation plus dead-zone calendar.

Watch trigger

A pullback to JPY 14,000-15,000 range, OR a strong August print confirming advanced-packaging material revenue mix accelerating.

Position guidance

No new position. Existing holders can hold but should consider trimming into strength.

Layer 5 · 5201 — AGC Inc.

Live: JPY (price not surfaced) · **Earnings (last):** Q1 FY26 OP +49% YoY on FX · **Next earnings:** Q2 FY26 · **Conviction:** 4/10

Bottleneck context

AGC is the Japanese glass-and-chemicals conglomerate — LCD glass substrates, EUV mask blanks (joint venture with Hoya), specialty chemicals. The Layer 5 thesis is mask-blank franchise plus rising AI-display content. Diversified business model dilutes the AI signal.

Why Tier C

Q1 beat looked solid on the headline (+49% YoY OP) but the beat was FX-driven, not unit-volume driven. Management held FY26 guide unchanged despite the Q1 beat — this is either ultra-conservative or a topping signal, and either way it removes the raise-the-guide catalyst. Live price not surfaced reliably (`needs_manual_check`). Flipped from B to C on the unchanged-guide signal plus data quality issue.

Watch trigger

A guide raise at Q2, OR a confirmed price plus clean LCD-glass pricing inflection.

Position guidance

No position. Re-evaluate on Q2 print.

Layer 5 · 5333 — NGK Insulators

Live: JPY (price not surfaced) · **Earnings (last):** FY guide JPY 85B below consensus 89.6B · **Next earnings:** Q1 FY27 · **Conviction:** 3/10

Bottleneck context

NGK Insulators makes ceramic substrates, semiconductor wafer-handling ceramics, and electronics components. AI relevance comes from the semiconductor-ceramic franchise (electrostatic chucks, susceptors) — Layer 5 materials supplier into the wafer-fab equipment chain.

Why Tier C

FY guide came in below consensus (JPY 85B vs 89.6B expected) — a soft-guide signal. Material auto-cycle exposure adds cyclical drag. The JPY 31B buyback is shareholder-friendly but does not change the demand picture. Quality franchise but the cycle is against it near-term.

Watch trigger

Guide upgrade at next print, OR cycle bottom confirmation in auto/industrial.

Position guidance

No position. Quality watch-only name.

Layer 5 · 5802 — Sumitomo Electric Industries

Live: JPY 9,900 · **Earnings (last):** FY26 EPS beat 18%; NI +91% YoY · **Next earnings:** Aug 2026 · **Conviction:** 4/10

Bottleneck context

Sumitomo Electric makes optical fibre, cable, and electronic components — the only Japanese vertically integrated optical fibre/cable player with InP wafer assets. Genuine AI optical-buildout leverage at Layer 5. FY26 EPS beat 18% and NI +91% YoY are strong headline numbers.

Why Tier C

The fundamentals are solid but the situation is messy. Auto headwinds drag on the legacy wire/harness business. Sell-side PT range is unusually wide — JPY 9,800-16,700 — signalling no consensus on the AI optical-leverage thesis. Intraday -2.85% suggests positioning unwind. Flipped from B to C on the spread of views plus tactical weakness.

Watch trigger

Clean print showing optical-fibre/InP segment revenue mix accelerating, OR PT range tightening upward as sell-side coalesces around the AI thesis.

Position guidance

No position. High-quality watch-only; revisit after Aug print.

Layer 5 · 6920 — Lasertec Corporation

Live: JPY 43,770 · **Earnings (last):** Q4 FY26 EPS/rev below consensus · **Next earnings:** Aug 2026 · **Conviction:** 4/10

Bottleneck context

Lasertec holds the actinic EUV mask inspection monopoly — sole-source supplier of the inspection tools mask shops must use to qualify EUV mask quality. Genuine Layer 5 monopoly franchise that should be Tier A on structure alone.

Why Tier C

Q4 print missed both EPS and revenue. -16.4% reaction confirmed the disappointment. Order timing in this business is famously lumpy — single-tool orders can swing quarters by 20%. HBM4 mask commentary from management was vague and is flagged `needs_manual_check`. Flipped B to C on the print disappointment; structurally this is a Tier A monopoly, tactically this is in the penalty box.

Watch trigger

Stabilization in order book with 2-3 sequential quarters of clean order flow, OR explicit HBM4 mask-inspection design-in disclosure.

Position guidance

No position. Best-in-class structural name in tactical penalty box — patient watch-only. Strong B candidate on a recovery print.

Layer 5 · PLAB — Photronics Inc.

Live: price not surfaced (needs manual check on Yahoo PLAB) USD · **Earnings (last):** Q1 FY26 — rev $225.1M +6.1% YoY, EPS $0.74; consensus n/a · **Next earnings:** 2026-05-28 (Q2 FY26; EPS est $0.53) · **Conviction:** 3/10

Bottleneck context

Photronics sits at the photomask layer immediately upstream of wafer fab. Masks are the physical pattern carrier for lithography — every chip needs one — so the franchise has structural moat against the IDM/foundry duopoly that uses third-party masks. The bottleneck PLAB participates in is high-end logic masks (sub-7nm) and the AI-specific mix shift toward more complex reticles per design tape-out. Capex is doubling to $330M in 2026 to support that pivot. The catch:

PLAB has no HBM photomask exposure — masks for memory are mostly produced in-house by Samsung/Hynix/Micron — so the AI bull case here is logic-only, and even within logic, the highest-margin EUV masks are increasingly captive at TSMC and Intel.

Why Tier C

Three flags. First, no HBM exposure means the dominant capacity-constrained piece of the AI stack bypasses PLAB entirely. Second, capex doubling weighs on near-term EPS just as the Q2 estimate ($0.53) shows sequential decline from Q1's $0.74. Third, top-line growth is only ~6% YoY — slower than any peer at the same stack layer. The AI photomask pivot is real but PLAB is at best a derivative beneficiary, not a primary one, and a price not surfaced in our refresh means we cannot even underwrite valuation.

Watch trigger

Q2 print May 28 — needs (a) EPS beat materially above the $0.53 estimate AND (b) management commentary explicitly quantifying AI logic mix as % of revenue. Both required to promote.

Position guidance

No position. Reassess only after May 28 print.

Layer 6 · 005290 — Dongjin Semichem

Live: KRW 52,900 · **Earnings (last):** Reference 2026-05-13 release — last quarter NI KRW 66.9B (up from 31.1B); FY24 rev +7.5%, NI +21.6%; consensus n/a · **Next earnings:** none scheduled within 60d · **Conviction:** 4/10

Bottleneck context

Dongjin makes CMP slurries, photoresists and wet chemicals for Korean memory fabs — primarily SK Hynix and Samsung. The bottleneck adjacency is real: HBM stack height growth (8Hi → 12Hi → 16Hi) increases CMP steps per die and pushes specialty-chemistry consumption per wafer. Within Korea this gives Dongjin a structural tailwind. But Dongjin is also a textbook commodity-chemistry name — peer Soulbrain competes head-on in HF/etchants, and pricing concessions are routine when memory cycles roll.

Why Tier C

The May 16 print produced a sharp -8.8% intraday drop (KRW 58k → 52.9k), signalling either disappointment in the read-through to HBM volume or broader Korea semi positioning unwind. The May 13 earnings catalyst is already past so the next 60-day window is empty. Cyclicality is the headline risk — every Korean memory-materials name trades as a high-beta proxy for the HBM cycle, and at current levels Dongjin offers neither a clean entry nor a near-term catalyst.

Watch trigger

Next quarterly print needs to show YoY NI growth re-accelerating above 30% AND a stabilising / rising stock base before promotion.

Position guidance

No position. Soulbrain or direct Korean memory exposure (000660) is a cleaner expression of the same thesis.

Layer 6 · 2408 — Nanya Technology

Live: TWD 311.50 · **Earnings (last):** Q1 2026 — EPS NT$8.41 vs -NT$0.63 prior; rev NT$49.1B; NI NT$26.1B; major swing to profit · **Next earnings:** ~2026-07-13 · **Conviction:** 4/10

Bottleneck context

Nanya is Taiwan's pure-play DRAM producer — DDR4 and some DDR5, no HBM. The thesis is commodity-DRAM supply tightness as Samsung, Hynix and Micron divert wafer capacity to HBM, leaving the conventional DRAM pool undersupplied. Q1 confirmed the cycle inflection — a NT$26.1B swing to profit on a 53% margin print. But Nanya does not own a bottleneck node; it benefits from one being created next door, which is fundamentally a cycle trade rather than a structural moat trade.

Why Tier C

The setup is cyclically positive but the durability is unclear. Once HBM capacity decisions stabilise (likely 2027), conventional DRAM pricing typically corrects fast. Stock has already re-rated meaningfully — the question is whether the next print extends the margin gain or shows it as a one-quarter spike. No HBM optionality, yield-execution overhang, and a long wait (July) for the next data point.

Watch trigger

July print needs to show GM holding > 50% AND management raising the cycle forward — both required for a B-tier upgrade.

Position guidance

No position. The cleaner cycle play is Micron in Tier A.

Layer 6 · 3436 — SUMCO Corporation

Live: JPY 3,578 · **Earnings (last):** Q1 FY26 (2026-05-12) — rev JPY 101.4B -1% YoY, OP loss JPY -5.3B (vs +6.0B prior), NI loss JPY -8.5B; missed badly · **Next earnings:** Aug-ish (not scheduled within 60d) · **Conviction:** 2/10

Bottleneck context

SUMCO is the world's #2 silicon wafer supplier (behind Shin-Etsu) — every chip starts on its substrate. Bottleneck adjacency is real for advanced 300mm wafers feeding AI logic and HBM, and the bull narrative had been that AI wafer demand would compensate for the weak 200mm legacy book. The Q1 print broke that thesis: even with the +88.9% 1M sentiment-driven rally going into the print, the company swung to an operating loss as non-advanced wafer drag overwhelmed AI uplift.

Why Tier C

Capital-cycle headwinds are now in the numbers, not just narrative. Stock is priced for an AI wafer revenue inflection that has not arrived. Japanese small/mid-cap silicon-wafer producers carry the additional risk of being caught between Shin-Etsu's scale advantage and Korean/Taiwanese alternatives. The Q1 OP loss surprise after a +147.6% 6M run-up is the worst possible technical setup.

Watch trigger

Two consecutive quarters of positive operating income AND explicit 300mm AI mix disclosure above 40% — only then promote.

Position guidance

No position. Avoid until at least Q3 FY26 print confirms a real turn.

Layer 6 · 357780 — Soulbrain

Live: price not surfaced (manual check pending; earnings releasing today) KRW · **Earnings (last):** FY24 — rev KRW 863.4B +2.3%; earnings KRW 118.4B -9.2% · **Next earnings:** 2026-05-18 (TODAY) · **Conviction:** 4/10

Bottleneck context

Soulbrain produces HF etchants and other wet chemicals concentrated in HBM and advanced logic — the closest Korean analogue to Entegris in the Tier A list. The HBM stack-height trade reads through directly: more layers = more etch / clean steps = more chemistry per wafer. Soulbrain is the more HBM-pure name vs Dongjin's broader memory-materials mix.

Why Tier C

FY24 earnings fell -9.2% YoY despite the HBM tailwind, suggesting either pricing concessions to SK Hynix or share loss to competitors. Today's Q1 release is the binary catalyst — if numbers reaccelerate it warrants a Tier B re-look, but until the print lands and the price surfaces for valuation work, this stays C.

Watch trigger

Today's release shows Q1 NI growth > 20% YoY AND price clears prior consolidation — promote on confirmation.

Position guidance

No position. Wait for today's print and then a clean entry.

Layer 6 · 3711 — ASE Technology

Live: TWD 555.00 · **Earnings (last):** Q1 2026 (2026-04-29) — rev NT$173.7B vs NT$168.5B est; EPS NT$3.08 vs NT$2.73 (+12.8% beat) · **Next earnings:** not within 60d · **Conviction:** 5/10

Bottleneck context

ASE is the world's largest OSAT and the largest non-captive advanced-packaging house — fan-out, flip-chip, CoWoS-adjacent ATM lines. Advanced-packaging guide is doubling to $3.2B for 2026, a strong fundamental signal. The bottleneck context: as TSMC fills its CoWoS internal capacity, overflow demand for chiplet integration and 2.5D packaging spills to ASE. Q1 was a clean beat — and on raw fundamentals ASE looks like an A-tier name.

Why Tier C

Two structural caveats keep ASE in C. First, TSMC is bringing more advanced packaging in-house — CoWoS at TSMC + COUPE optical packaging — which caps the upside ASE can capture as the AI mix shifts to highest-margin assemblies. Second, the EMS (electronics manufacturing services) half of the business is seasonally soft and dilutive to AI mix. The batch view flipped this from B to C on those structural caps; we agree. Strong fundamentals are not enough when the share of the highest-value pie is contracting.

Watch trigger

Either (a) management discloses CoWoS-equivalent advanced packaging revenue surpassing $1B/quarter OR (b) ASE wins an explicit NVIDIA / AMD chiplet program. Either flips to Tier B.

Position guidance

No position. AMKR (also Tier C here) and direct CoWoS exposure via Tier A names cover the same theme more cleanly.

Layer 6 · 402340 — SK Square

Live: KRW 1,104,000 · **Earnings (last):** holding co — value tracks SK Hynix ~20% stake · **Next earnings:** not within 60d · **Conviction:** 3/10

Bottleneck context

SK Square is the listed holding company that owns ~20% of SK Hynix plus other minority stakes (semiconductor design, advertising). It exists as a NAV-discount vehicle — when SK Hynix re-rates, SK Square typically follows with a delay and at a discount. The Hynix bottleneck (HBM monopoly with NVIDIA pull-through) is the entirety of the thesis here.

Why Tier C

SK Square crossed the symbolic KRW 1M “emperor stock” threshold in May, hitting a 52-week high of KRW 1.149M, then pulled back -5% on Kospi foreign selling. The structural problem: NAV-discount mechanics introduce additional governance overhang on top of Hynix’s own risk, and the entire position is redundant with direct 000660 ownership. The batch flipped this from B to C precisely for that reason — there is no scenario in which SK Square outperforms direct SK Hynix on a risk-adjusted basis except via discount narrowing, which is governance-dependent.

Watch trigger

Discount to NAV narrows below 30% on confirmed buyback or restructuring announcement.

Position guidance

No position. Hold SK Hynix directly via 000660.

Layer 6 · 4186 — Tokyo Ohka Kogyo (TOK)

Live: price not surfaced (needs manual check on Yahoo 4186.T) JPY · **Earnings (last):** Q1 2026 — rev surprise +5.90%, EPS beat consensus; guide neutral · **Next earnings:** 2026-05-13 (release reference; verify on next report) · **Conviction:** 4/10

Bottleneck context

TOK is one of three Japanese photoresist majors (with JSR and Shin-Etsu Chemical) and one of the few EUV photoresist suppliers globally. EUV photoresist is the chemistry layer that enables sub-7nm patterning — without it, the AI logic node ramp stops. Real bottleneck, real moat. Feb 2026 partnership with Irresistible Materials extends the EUV roadmap, and Q1 revenue beat by 5.9% confirms underlying franchise health.

Why Tier C

The problem is competitive intensity at the EUV layer itself — JSR, Shin-Etsu and Fujifilm all serve the same TSMC/Samsung/Intel customer set, and customer share is closely guarded. With the price not surfaced in our refresh, we cannot underwrite valuation, and a neutral guide direction suggests management does not expect the next leg to come from EUV mix in the next quarter. Decent franchise, no near-term catalyst, no verified price — exactly the C-tier setup.

Watch trigger

Next print shows EUV resist disclosed as > 25% of revenue AND price surfaces below ~JPY 11,000 for an acceptable entry.

Position guidance

No position. Verify price first; reassess on next earnings.

Layer 6 · 6488 — GlobalWafers

Live: TWD 822.0 · **Earnings (last):** Q1 2026 — EPS TWD 3.97 beat 2.2%; rev declined and GM compressed on ramp costs · **Next earnings:** not within 60d · **Conviction:** 4/10

Bottleneck context

GlobalWafers is the #3 silicon wafer player globally (after Shin-Etsu and SUMCO). The bottleneck adjacency is identical to SUMCO's — every chip starts on a wafer — and GlobalWafers is building greenfield 300mm capacity in Texas and Missouri to capture AI demand inside the CHIPS Act umbrella. Q1 showed the trade-off: EPS beat but gross margin compressed and revenue declined, because greenfield ramp costs hit before the AI volume arrives.

Why Tier C

Stock is up 142% over the past year on a wafer-cycle recovery narrative, even as Q1 GM took a step down. The bull-vs-bear hinge is whether the AI wafer pull-through materialises before the new capacity creates oversupply. That timing question makes this a B/C swing depending on data — and at +142% trailing return with margin compression visible, the asymmetry tilts negative. The same thesis is cleaner via Shin-Etsu Chemical (Tier B in our universe) which has higher-margin franchises (photoresist, encapsulation) on top of wafers.

Watch trigger

GM stabilises sequentially AND management discloses AI mix > 30% of advanced wafer revenue.

Position guidance

No position. Re-look only if stock corrects 20%+ from current levels.

Layer 6 · 8150 — ChipMOS Technologies

Live: TWD 73.7 · **Earnings (last):** Q1 2026 — EPS TWD 0.71 vs TWD 0.52 (+36.5% beat); rev $216.4M +25.4% YoY · **Next earnings:** not within 60d · **Conviction:** 4/10

Bottleneck context

ChipMOS is a niche OSAT focused on DDIC (display driver IC) and memory test. DDIC supports the AI-adjacent display ecosystem (server-room monitors, automotive ADAS displays, OLED for premium tier devices) but is not itself AI-bottlenecked. Memory test is where the AI read-through exists — HBM stack growth requires more known-good-die verification, and ChipMOS sits on that flow. Q1 beat by 36.5% reflected DDIC and memory test recovery, both real.

Why Tier C

Stock is up 183% over the past year, so the cycle recovery is already substantially priced. Two structural risks: (a) DDIC is cyclical with smartphone and TV demand, both of which are weak; (b) memory test pricing pressure from HBM siphoning OSAT capacity to higher-margin lines can actually hurt the conventional memory test book. ChipMOS captures the AI theme indirectly, at a smaller scale than AMKR or ASE, with more cyclical baggage.

Watch trigger

Two consecutive quarters of EPS beat with explicit HBM test win disclosure.

Position guidance

No position. AMKR or direct Advantest exposure dominates this expression.

Layer 6 · AEHR — Aehr Test Systems

Live: USD 105.72 · **Earnings (last):** Q3 FY26 — EPS -$0.05 (beat -$0.07); rev $10.3M (miss $10.8M); AH -5.67% · **Next earnings:** not within 60d · **Conviction:** 5/10

Bottleneck context

Aehr makes wafer-level burn-in systems — the equipment that stress-tests chips before they ship. The AI angle is real and specific: hyperscale AI ASICs (custom inference chips for Google, AWS, Meta) require burn-in to validate sub-1ppm defect rates, and Aehr's WaferPak system is one of two industry-standard solutions. Bookings of $92M for H2 FY26 vs prior $60-80M guide confirm a step-function order increase, anchored by a $41M follow-on burn-in order from a lead hyperscale customer for AI ASICs.

Why Tier C

The reason Aehr is not Tier B yet: revenue is still down 44% YoY and gross margin compressed to 36.5% from 42.7%. Lumpy order patterns mean Q3 revenue lagged the booking surge significantly, and Aehr has historically been a feast-or-famine name. Stock is +144% in the last month and hit a 52-week high May 11 — entry now means buying at the order-momentum peak. The setup is asymmetric in the wrong direction: any quarterly miss between now and FY27 ramp visibility will get punished.

Watch trigger

FY27 outlook from management explicitly above $200M revenue run-rate.

Position guidance

No position. Strong watch-list candidate; would prefer to re-enter after a 20%+ pullback or on FY27 guide.

Layer 6 · AMKR — Amkor Technology

Live: USD 70.35 · **Earnings (last):** Q1 2026 — EPS $0.33 vs $0.22 (+50% beat); rev $1.68B record (+27% YoY) · **Next earnings:** not within 60d · **Conviction:** 5/10

Bottleneck context

Amkor is the #2 non-captive OSAT (behind ASE), with disproportionate advanced packaging exposure including HDFO (high-density fan-out) for datacenter CPU AI accelerators. Q1 was a record print — communications +42% YoY, computing +19% with record AI DC revenue — and management guided HDFO datacenter CPU programs to ramp in Q2 with $2.5-3.0B capex for advanced packaging expansion. The bottleneck context is genuine: TSMC's CoWoS capacity tightness pushes some advanced packaging work to Amkor.

Why Tier C

Two issues. First, consensus rating is Hold with an average PT of $54.68 — that implies 22% downside from $70.35 even as fundamentals beat. The street is implicitly signalling that the Q1 record is the peak. Second, stock is +336% over the past year, valuation has detached from earnings, and capex intensity weighs on FCF near-term. The setup is identical to ASE: strong AI bottleneck participation, but valuation has front-run the fundamentals.

Watch trigger

PT consensus revises to within 10% of spot AND Q2 print delivers another double-digit beat.

Position guidance

No position. Wait for the inevitable consolidation or sector pullback.

Layer 6 · VECO — Veeco Instruments

Live: USD 60.26 · **Earnings (last):** Q1 2026 — EPS $0.14 missed $0.27; rev $158.34M missed $162.13M; AH +21.76% on backlog · **Next earnings:** not within 60d · **Conviction:** 4/10

Bottleneck context

Veeco makes deposition, etch and laser annealing tools used in advanced packaging, silicon photonics and InP (indium phosphide) laser manufacture. Compound semi revenue is projected +50% in 2026 on InP laser adoption for optical interconnect — that is a real bottleneck adjacency given the CPO / silicon photonics wave we have flagged elsewhere in the master document. >$250M in new multi-customer tool orders gives multi-year backlog visibility.

Why Tier C

Q1 missed both EPS and revenue consensus — only the order announcement saved the stock. Stock is +209% past year and +108% YTD; valuation is stretched at a 52-week high. The backlog-driven re-rating is faith-based until conversion to revenue actually shows in print numbers. Until the next earnings translates backlog into top line, this remains a C-tier “show me” name even though the InP / CPO bottleneck thesis is intellectually attractive.

Watch trigger

Next print shows backlog conversion with rev > $200M AND GM expansion sequentially.

Position guidance

No position. Strong watch candidate alongside AEHR.

Layer 7 · NVTS — Navitas Semiconductor

Live: USD 20.0 · **Earnings (last):** Q1 2026 — rev $8.6M (beat $8.18M); EPS -$0.04; +18% sequential rev growth · **Next earnings:** not within 60d · **Conviction:** 4/10

Bottleneck context

Navitas makes GaN (gallium nitride) and SiC (silicon carbide) power semiconductors. The AI-specific angle is the NVIDIA Rubin 800V power delivery architecture — at GTC 2026 Navitas demoed an 800V-to-6V power delivery board for Rubin, and Baird raised its PT $9 → $20 on this thesis. If the 800V architecture wins as the standard for next-gen AI racks, GaN/SiC vendors capture a high-margin layer that did not exist in pre-AI data centers.

Why Tier C

Real concerns. NVTS is still unprofitable (-$0.04 EPS) on a tiny $8.6M revenue base — the design-win narrative is much bigger than the financial reality, and one program slip from NVIDIA collapses the thesis. Insider selling has been reported. The stock is up 88% YTD and is held at +140% by some long names — the asymmetry is now negative. Direct AI exposure exists but in token-revenue form. Limited operating leverage shows up only in 2027+ if everything goes right.

Watch trigger

Q2 print shows revenue > $10M AND management confirms first dollar of Rubin-related revenue.

Position guidance

No position. Hold any existing position with a tight stop; do not initiate new.

Layer 7 · POWI — Power Integrations

Live: USD 74.99 · **Earnings (last):** Q1 2026 — non-GAAP EPS $0.25 beat by 11%; rev $108.3M missed top-line (but +2.6% YoY) · **Next earnings:** not within 60d · **Conviction:** 3/10

Bottleneck context

Power Integrations sells AC-DC and DC-DC controller ICs into industrial, consumer, communications and automotive end markets. The industrial segment is growing +23% YoY on renewables and battery storage, now 41% of sales. That growth is genuine but it is not directly tied to AI data center power delivery — that bottleneck flows through NVTS, VICR (Tier B) and a few specialty silicon names. POWI is best understood as an industrial-cycle name with peripheral AI mention.

Why Tier C

Limited direct AI data center exposure is the headline reason — peers NVTS and VICR are the clean expressions of the 800V power delivery thesis. Q1 missed top-line consensus (beat only on EPS), and the +2.6% YoY growth rate is not high enough to argue for re-rating. The company is a good operator with a clean balance sheet, but it is not a bottleneck name in our master frame. The C-tier flag is structural, not tactical.

Watch trigger

Disclosure of an explicit AI data center program win (e.g., NVIDIA or hyperscaler design-in).

Position guidance

No position. Default pass.

Layer 7 · WOLF — Wolfspeed Inc.

Live: USD 38.55 · **Earnings (last):** Q3 FY26 — EPS -$3.26 vs -$0.56 consensus; rev $150M vs $209.76M consensus (-28%); emerged Chapter 11 · **Next earnings:** not within 60d · **Conviction:** 3/10

Bottleneck context

Wolfspeed is a pure-play silicon carbide (SiC) vertically integrated producer — substrate, epi, device, module. The AI / power-electronics bottleneck is real if you believe SiC ends up in AI rack-level power conversion, but the dominant SiC demand pull is electric vehicles, where the demand picture has materially softened. The Mohawk Valley fab is the asset; the question has always been whether ramp execution beats the cash burn timeline.

Why Tier C

Just emerged from Chapter 11 — total debt cut ~70%, maturities pushed to 2030, cash interest -60%. Stock is +7.26% AH post-Q3 and has recovered to $38.55 near 52-week high. That is the positive read. The negative read: Q3 revenue still missed Street consensus by 28%, EPS loss of -$3.26 was almost 6x worse than -$0.56 consensus, and the operational thesis remains unproven even with the balance sheet repaired. Internal flag moved from WATCH-DISTRESS 3/10 to WATCH-RECOVERY but still high-risk turnaround.

Watch trigger

Two consecutive quarters of revenue beating prior quarter AND Mohawk Valley utilisation disclosed above 60%.

Position guidance

No position. Upgrade-watch only — do not initiate until earnings confirm operational stabilisation. Speculative re-look candidate for Q1/Q2 2027.

Appendix A — Coverage Audit

- **Universe size:** 98
- **Tier A:** 18
- **Tier B:** 49
- **Tier C:** 31
- **Names covered (post-assembly):** 98 of 98

Dedupes resolved

Ticker	Kept	Dropped	Reason
PLAB	05b_tier_c_part2.md	05a_tier_c_part1.md	longer/more complete (same tier)
005290	05b_tier_c_part2.md	05a_tier_c_part1.md	longer/more complete (same tier)
2408	05b_tier_c_part2.md	05a_tier_c_part1.md	longer/more complete (same tier)
3436	05b_tier_c_part2.md	05a_tier_c_part1.md	longer/more complete (same tier)
357780	05b_tier_c_part2.md	05a_tier_c_part1.md	longer/more complete (same tier)
3711	05b_tier_c_part2.md	05a_tier_c_part1.md	longer/more complete (same tier)

Names that fell through and were stubbed in assembly

Each of the following was missing from the deep-dive section files; a placeholder profile was generated from `tiers_final.json` metadata and placed in its assigned tier. Full long-form writeups are queued for the next refresh cycle.

Ticker	Name	Layer	Tier	Note
3037	Unimicron Technology	4	Tier B	added in assembly — full snapshot pending
4062	Ibiden	4	Tier B	added in assembly — full snapshot pending
6315	TOWA Corporation	4	Tier B	added in assembly — full snapshot pending
AAOI	Applied Optoelectronics	4	Tier B	added in assembly — full snapshot pending
AXTI	AXT Inc	4	Tier B	added in assembly — full snapshot pending
CIEN	Ciena Corporation	4	Tier B	added in assembly — full snapshot pending
COHR	Coherent Corp	4	Tier B	added in assembly — full snapshot pending
CRDO	Credo Technology	4	Tier B	added in assembly — full snapshot pending

Appendix B — Data Refresh Provenance

- **Source files:** `data/batch_1.json` through `data/batch_4.json` (4 batches × ~24-25 tickers each).
- **Refresh date:** 2026-05-18.

- **Re-rank methodology:** `data/tiers_final.json` — generated by applying tier discipline to the fresh batch outputs, capturing 11 material flips from the batch recommendations.

Tickers flagged needs_manual_check

Tickers below carried a price or earnings data variance in the batch refresh that warrants manual verification before sizing. Listed here for the QA pass:

- **489790** — Hanwha Vision (Semitech)
- **110990** — DIT Corp
- **033640** — Nepes
- **6315** — TOWA Corporation
- **5201** — AGC Inc.
- **5333** — NGK Insulators
- **6981** — Murata Manufacturing
- **PLAB** — Photronics Inc.
- **SNPS** — Synopsys
- **GLW** — Corning Incorporated
- **APH** — Amphenol Corporation
- **6920** — Lasertec Corporation
- **4186** — Tokyo Ohka Kogyo (TOK)
- **357780** — Soulbrain
- **FORM** — FormFactor
- **VICR** — Vicor Corporation
- **4063** — Shin-Etsu Chemical
- **8150** — ChipMOS Technologies