KGI

CIO Insights

Cusson Leung, Chief Investment Officer, International Wealth Management

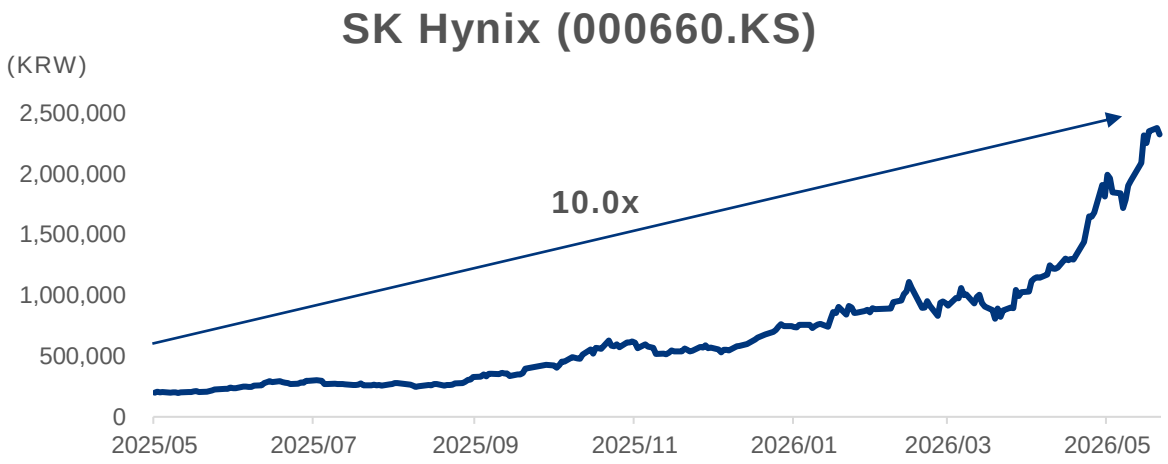
“Memory” Is Still Too Good to Forget

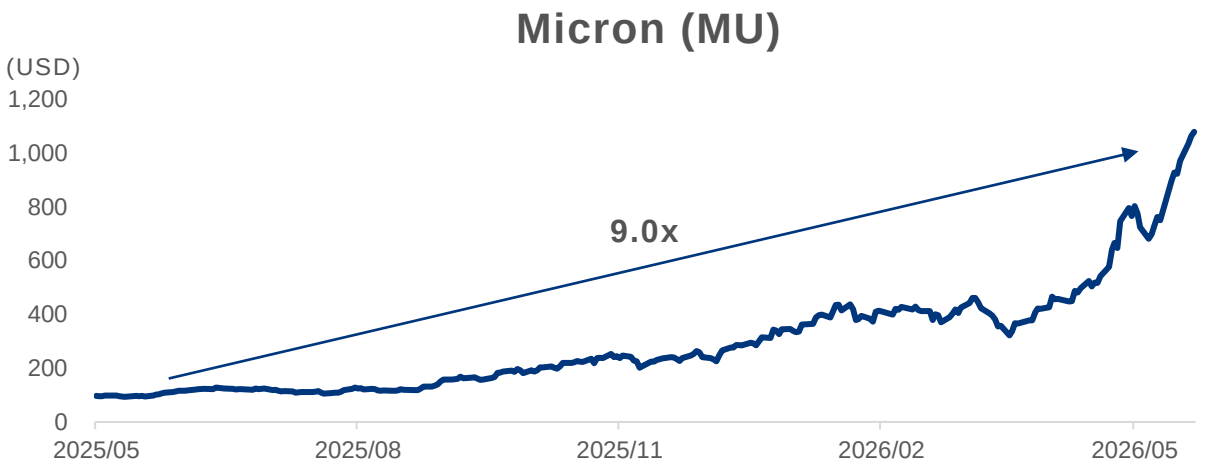
Diversify Into Other Ai-Bottleneck Supply Chain in Case the Music Stops

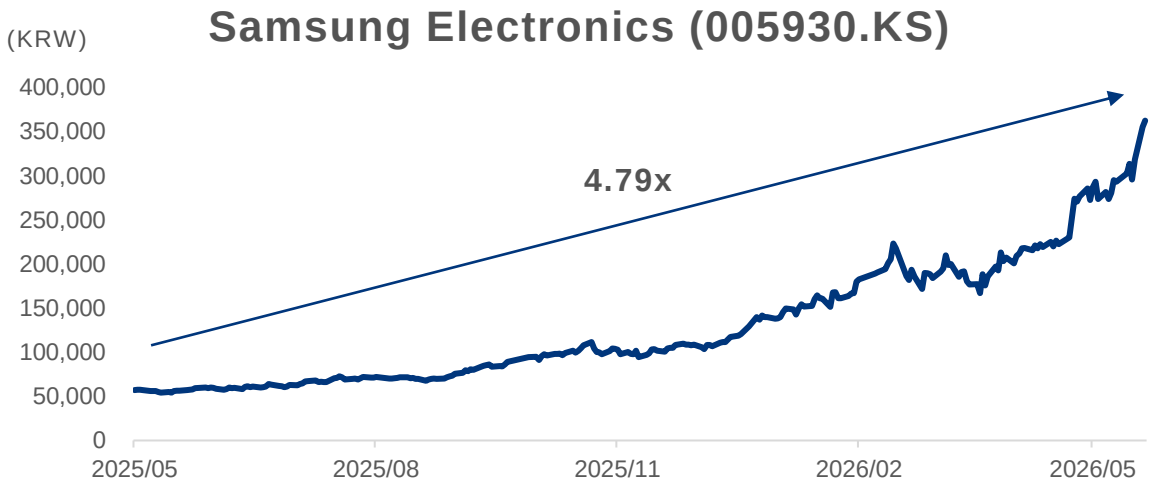
5 June 2026

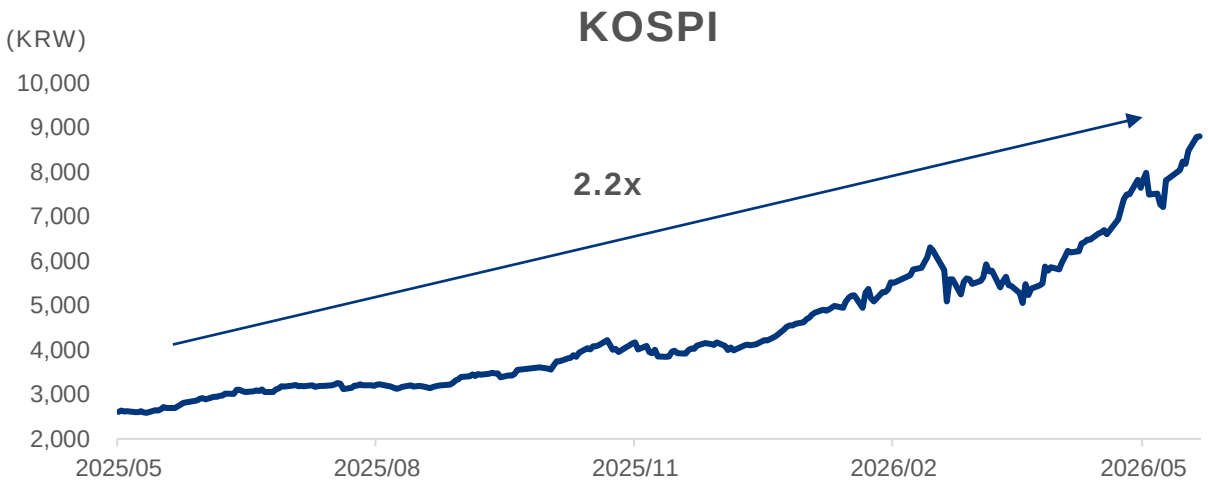
GPU and HBM Are Inseparable

HBM (High Bandwidth Memory) Is Like Having 10 Assistants Passing Ingredients to a Chef at Once







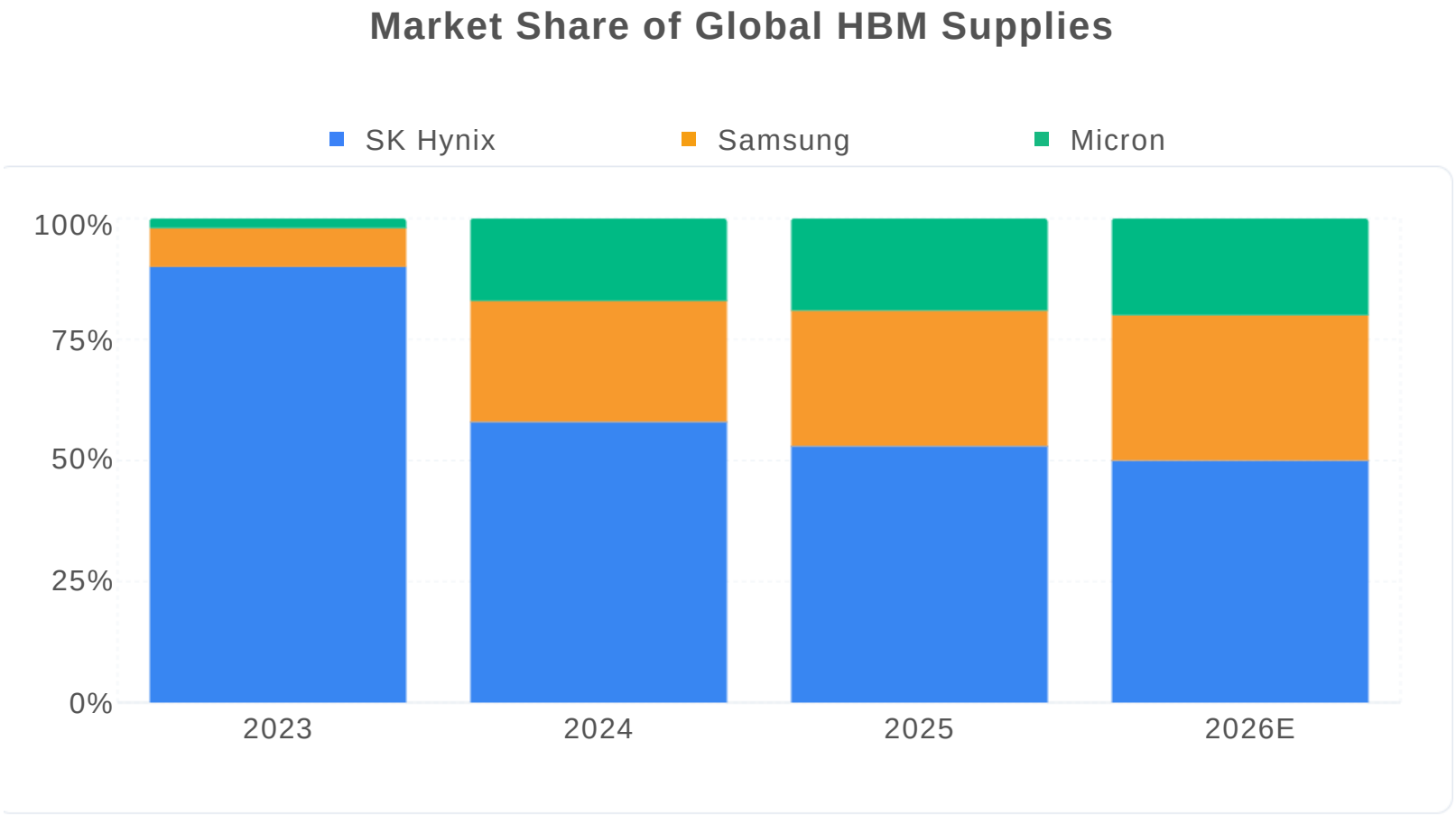


Source: Bloomberg, KGI

Only Three Companies in the World Can Make HBM

SK Hynix Is Still the World Largest Supplier Despite Market Share Gain From Samsung and Micron

- Producing one HBM chip consumes roughly three times the wafer area of a standard DRAM chip, which means that every HBM chip made is effectively three standard memory chips that were never made. This is why the broader memory market — the chips in your phone, laptop, and TV — has been experiencing shortages and price increases. The entire industry has pivoted its manufacturing lines toward HBM, leaving conventional memory supply constrained.

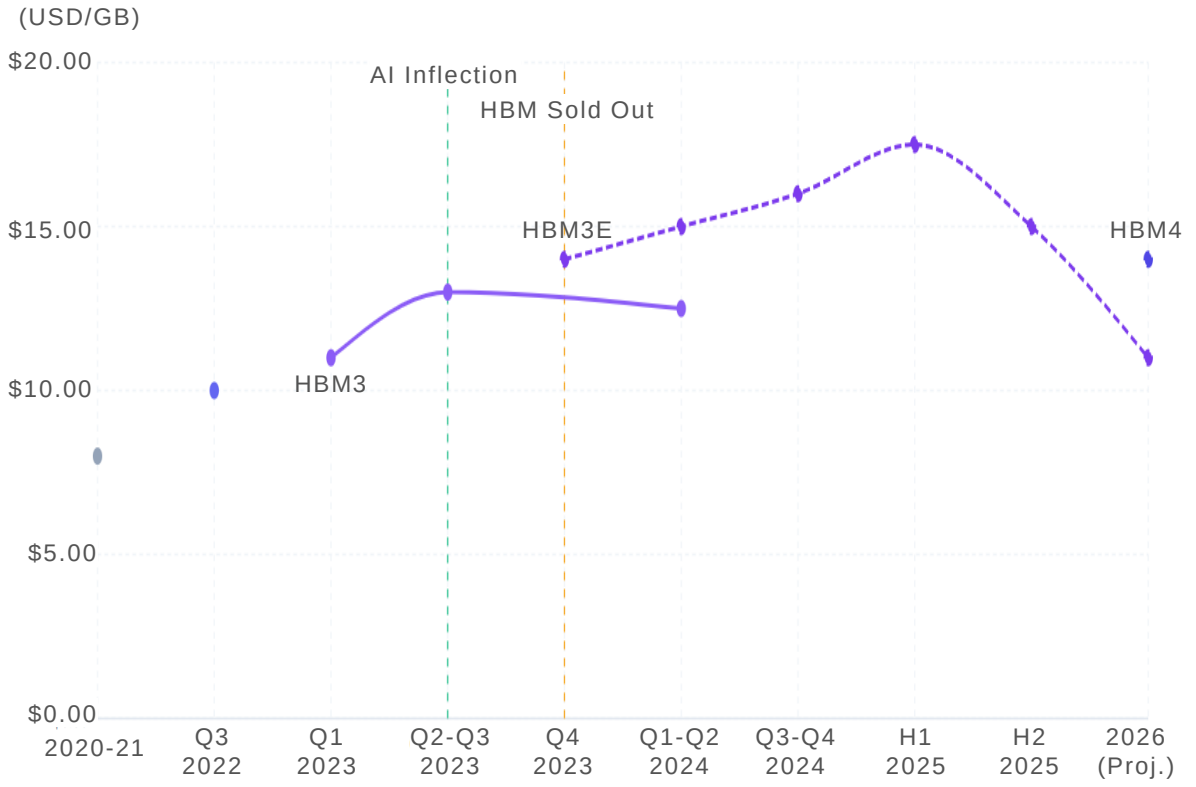


Source: KGI

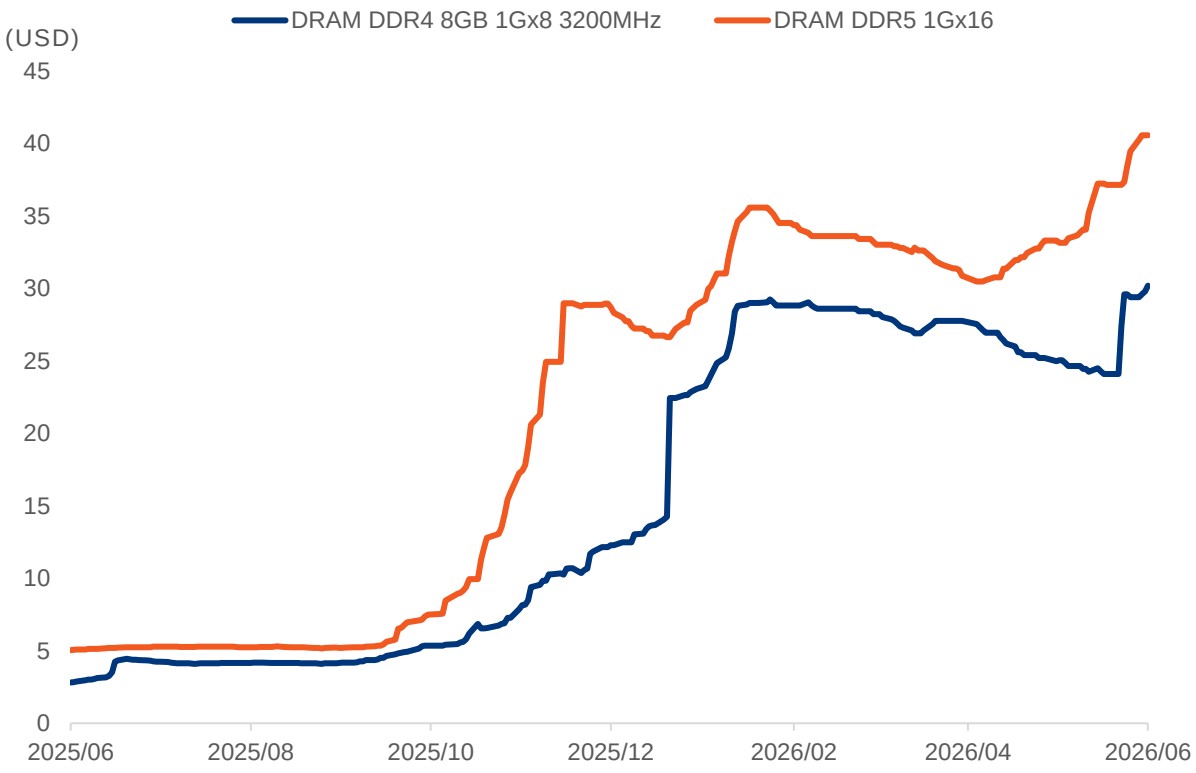
Short-Term Pause in Memory Prices

Does It Mean Peaking of Share Price Momentum?

HBM Spot Price



DRAM Spot Price



Source: Bloomberg, Silicon Analyst, KGI

Supply Constraints Is the Only Guardrail for Prices

Three Players Have Committed US$100bn Of Capex But Supply Ramp Up Needs 18 To 24 Months

- All three major memory producers have announced historic capital expenditure programs. SK Hynix (000660.KS) is constructing an entirely new semiconductor city in Yongin, South Korea — a campus so large it will eventually house four advanced fabs and is expected to cost over $94 billion when fully built out by 2047.
- Samsung Electronics (005930.KS) is ramping its massive Pyeongtaek campus in Korea — the world's largest semiconductor facility — and is simultaneously building a $17 billion fab in Taylor, Texas.
- Micron (MU.US), the only US-based player, is investing $25 billion in its fiscal year 2026 capital expenditure alone, its largest ever, and has signed its first-ever five-year customer supply agreement, signaling a fundamental shift from spot-market selling to long-term contracted revenue.
- The critical caveat for investors is timing: expansion of production capacity takes at least 18-24 months. The capacity being committed today will not produce chips until 2027 at the earliest, and in large volumes not until 2028. This creates a window of continued supply tightness — and elevated pricing — through at least the next 18 to 24 months.

Source: KGI

The Memory Trade Has Become Crowded

We Are Not Sure When Everyone Will Forget the “Memory” – It Is Beyond a Fundamental Call

- Retail participation at high level (if not at its peak). When a sector's investment thesis becomes widely understood and retail investors begin participating at scale, it is often a sign that the easy money has already been made. In 2026, the memory trade has reached exactly this point. South Korea's stock market has surged nearly 96% year-to-date, driven almost entirely by Samsung Electronics (005930.KS) and SK Hynix (000660.KS). A new US-listed ETF called the Roundhill Memory ETF (DRAM) attracted over $10 billion in assets within just 50 days of its April 2026 launch — one of the fastest ETF growth stories in history.
- In South Korea, regulators approved the country's first-ever single-stock leveraged ETFs in May 2026, with 16 products tied exclusively to Samsung Electronics (005930.KS) and SK Hynix (000660.KS) listing on the Korea Exchange. On their debut day alone, these products generated $6.9 billion in trading volume. While the underlying demand for HBM is real and structural, the risks are equally real: memory is the most cyclical corner of the semiconductor industry, with past downturns seeing chip prices fall 50–60% within 18 months.
- The market liquidity and the current market sentiment is too strong to be ignored. The fundamental story for HBM is real but the risk is also real. It would be too early for us to call it a day on memory. Instead, we guide investors to focus on the next AI bottleneck rather than chasing the crowd.

Source: KGI

The Bottleneck Migration

The Game of Finding the Weakest Link

- The development of AI is not a smooth, linear progression. It is a shifting-constraint problem. In any complex system — like a massive computer cluster or a global factory network — there is always one component that is the slowest, the tightest, or the most expensive. That component is the bottleneck. It dictates the speed of the entire system.

- As soon as engineers and billions of dollars of capital resolve that bottleneck, the limit immediately migrates to the next adjacent component. We have already seen this play out in phases: Phase 1 was raw computing power (GPUs). Phase 2 was advanced chip packaging (TSMC's CoWoS). Phase 3, which is peaking in 18-24 months, is memory bandwidth (HBM).

- As the memory industry commits $100 billion to expand production, the HBM shortage will eventually ease. But this will not make AI development infinitely fast. Instead, the bottleneck is shifting again. We believe the three physical-world constraints that are emerging as the next critical chokepoints are: **Optical networking**, **Thermal management**, and **Power grid infrastructure.**

Source: KGI

Bottleneck One: Optical Networking – the Speed of Light

Connecting 100,000 GPU Together Requires Moving Data in Light Waves Instead of Copper Wires

- Building a modern AI supercomputer is not like building a regular server. You cannot just put 100,000 GPUs in a room and turn them on. They must all talk to each other constantly, sharing billions of data points every millisecond to train a single model. If the connection between them is slow, the GPUs sit idle, wasting expensive computing power. This is what engineers call the "communication wall."

- To break this wall, the industry is transitioning from copper cables (which are too slow and generate too much heat over distance) to optical networking. Instead of electricity, data is converted into light waves and sent through fiber-optic cables. This requires a highly specialized component called an **“optical transceiver”**.

- As clusters scale to 100,000-plus GPUs in 2026, hyperscalers are rapidly upgrading from 800G to 1.6T (Terabit) transceivers. The critical material bottleneck here is a rare semiconductor called **“Indium Phosphide (InP)”**, which is used to make the microscopic lasers inside these transceivers. Because only a handful of specialized companies can manufacture these advanced optical components, lead times are stretching, making optics the next major chokepoint in cluster construction.

Source: KGI

Bottleneck Two: Liquid Cooling – Air Is No Longer Enough

Cooling Down the Sprint Is as Essential as the Sprint Itself

A single NVIDIA Blackwell GPU generates up to 1,200 watts of heat — roughly the same as a small electric space heater. Imagine 72 of these chips packed into a single server rack. The heat density is so extreme that blowing air over them, the way the laptop fan works is simply no longer sufficient. The industry has moved to liquid cooling: running chilled water through pipes directly attached to the chips, the same basic principle as the radiator in a car.

Source: KGI

Bottleneck Three: The Power Grid

Half of the US's Planned AI Data Centers in 2026 Have Been Delayed Due to Grid Connection Limits

- A large AI data center consumes as much electricity as a small city. To connect to the national power grid, it needs a piece of equipment called a "**Large Power Transformer**" — a massive, custom-built device that steps down the high-voltage electricity from transmission lines to the lower voltages that buildings and computers can use.

- These transformers are not off-the-shelf products. They are custom-engineered for each installation, weigh hundreds of tonnes, and are assembled largely by hand by skilled workers winding miles of copper wire around enormous steel cores. Before the AI boom, lead times for these transformers were around 30 to 50 weeks. **Today, they have stretched to between 128 weeks and four years**.

- The reason is a combination of surging demand from AI data centers, electric vehicle charging infrastructure, and renewable energy projects — all competing for the same limited manufacturing capacity. The steel used in transformers — a specialized product called Grain-Oriented Electrical Steel (GOES) — is itself in short supply, with only a handful of mills in the United States, Europe, and Japan capable of producing it.

Source: KGI

Moving From Silicon to the Physical Infrastructure

Nothing Is Untouched Along the AI Supply-Chain but They May Not Be as Crowed as “Memory”

- Optical Networking: The transition to 1.6T transceivers creates immense pricing power for laser diode leaders like **Lumentum** (LITE.US) and **Coherent** (COHR.US), and high-precision manufacturer **Fabrinet** (FN.US). In Taiwan, **LandMark Optoelectronics** (3081.TW) is a critical chokepoint as the primary foundry for the Indium Phosphide (InP) epi-wafers used to grow these lasers.

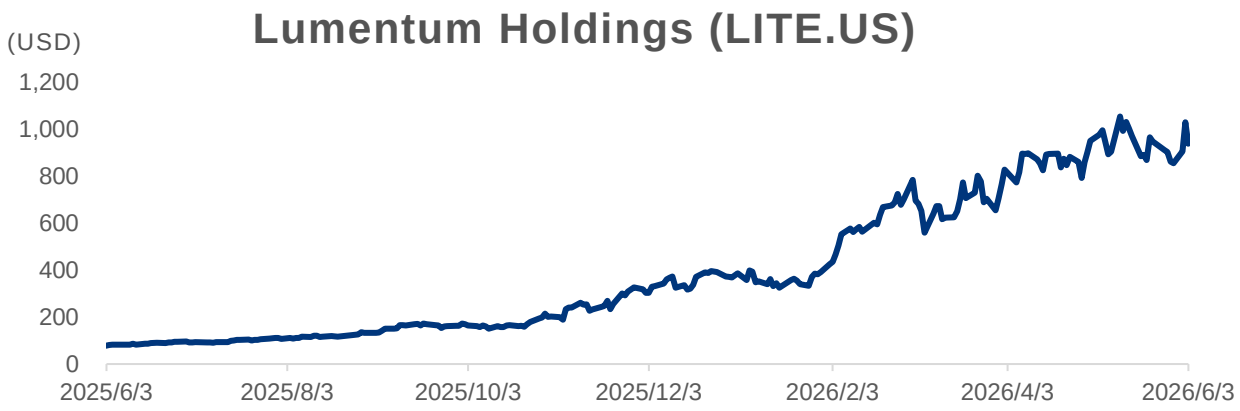
- Liquid Cooling: Direct-to-chip cooling demands zero-leak mechanical tolerances. While US-based **Vertiv** (VRT.US) dominates system design, Taiwanese thermal giants **Asia Vital Components** (AVC, 3017.TW) and **Auras Technology** (3324.TW) are the leading global suppliers of the advanced cold plates and liquid cooling modules required for NVIDIA's GB200 systems.

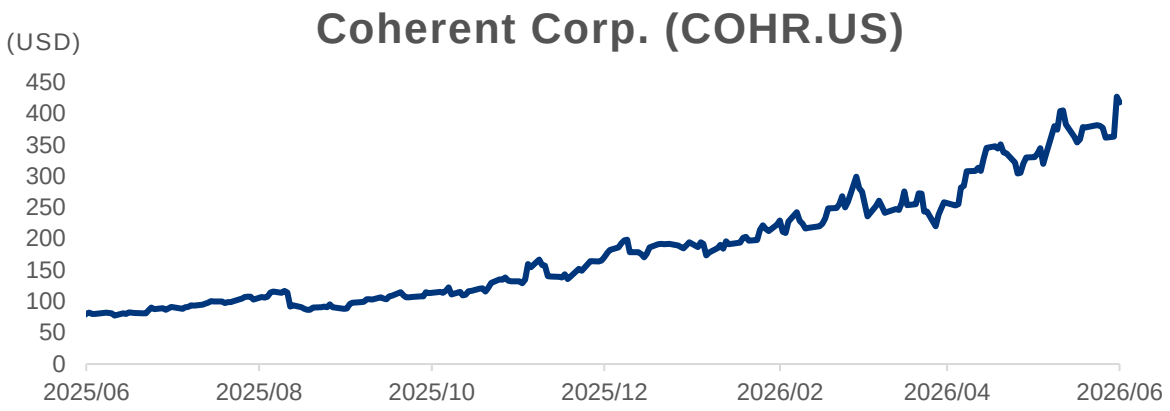
- Grid Hardware: Connecting gigawatt-scale data centers is blocked by 4-year transformer lead times. Alongside global giants like **Eaton** (ETN.US) and **Siemens Energy** (ENR.DE), Taiwan's **Fortune Electric** (1519.TW) has emerged as a high-growth play, actively exporting custom high-voltage transformers directly to US utilities to bypass domestic supply constraints

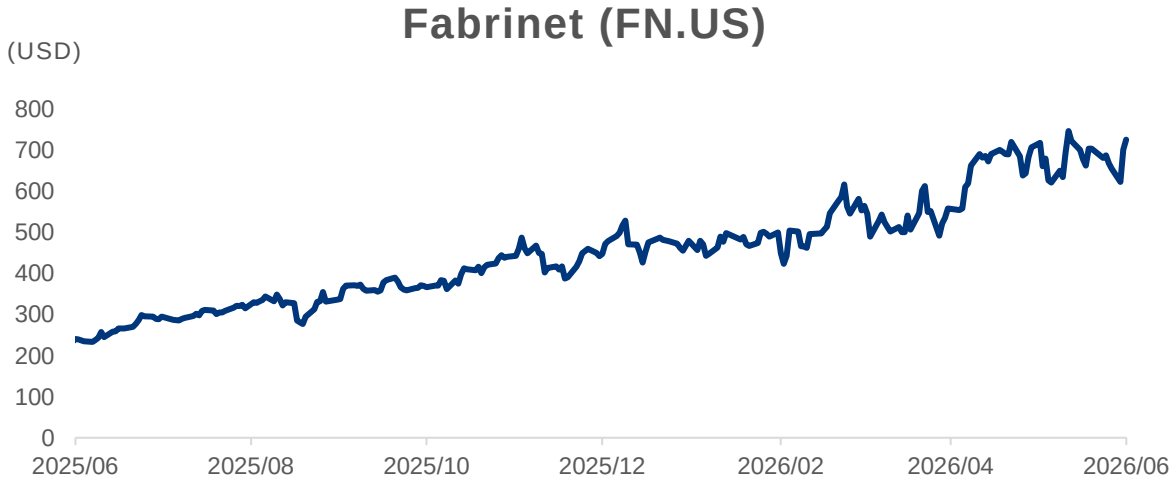
Source: KGI

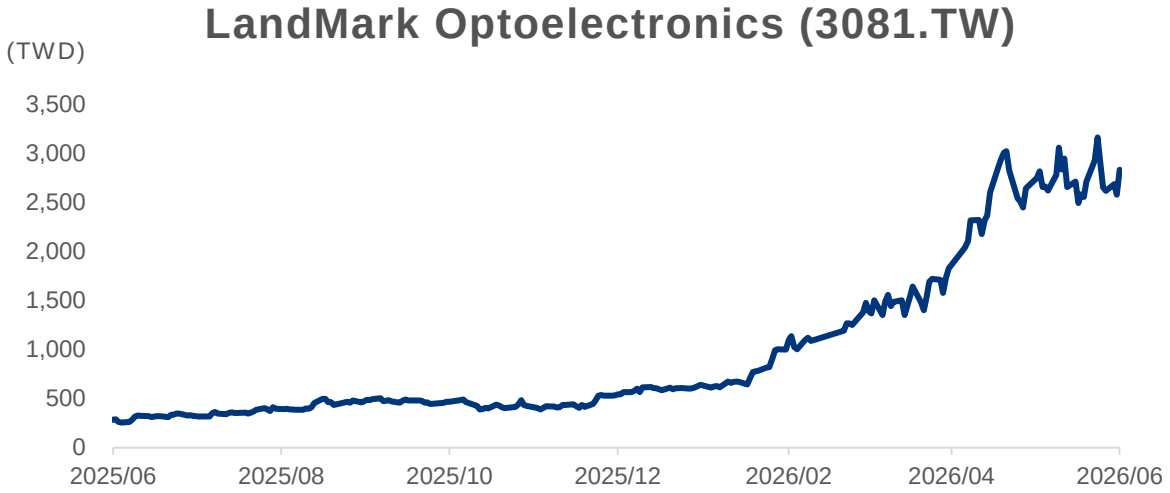
Optical Networking

Their Share Price Performances Has Been the Closest to the Memory Stocks





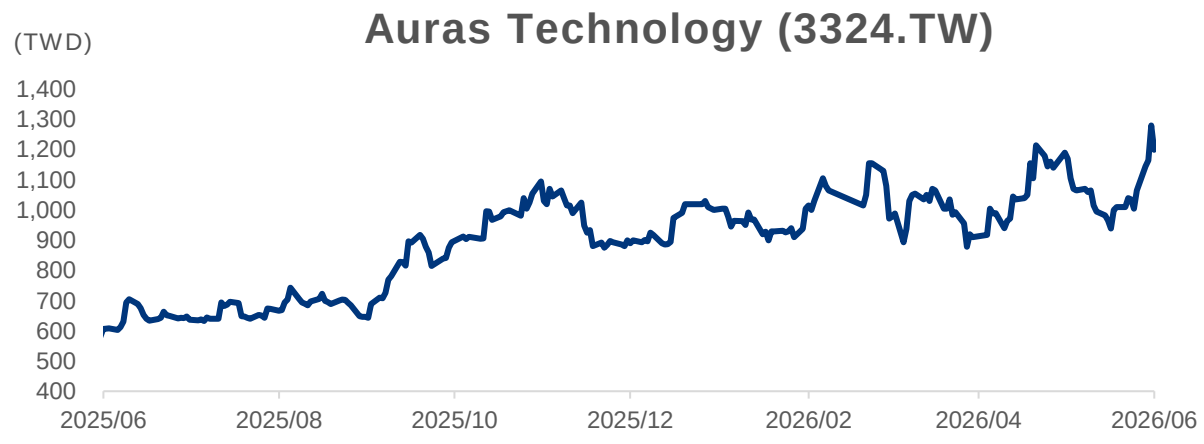




Source: Bloomberg, KGI

Liquid Cooling

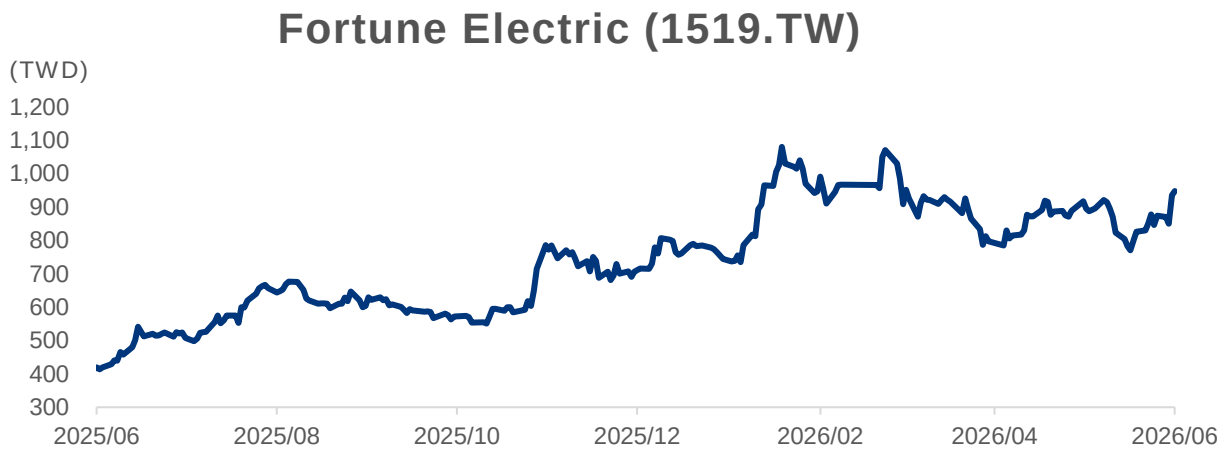
“Temperature” of the Share Prices Are Still Under Control

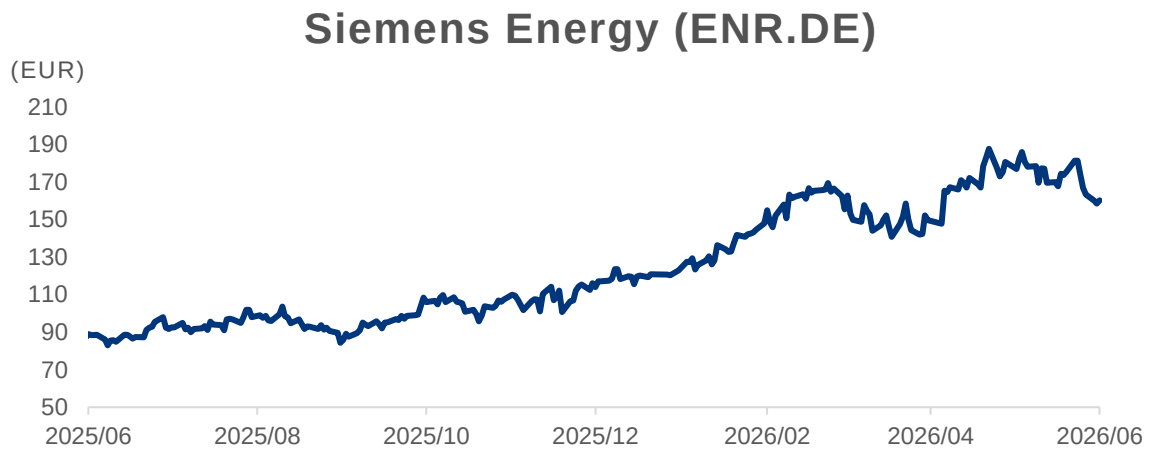


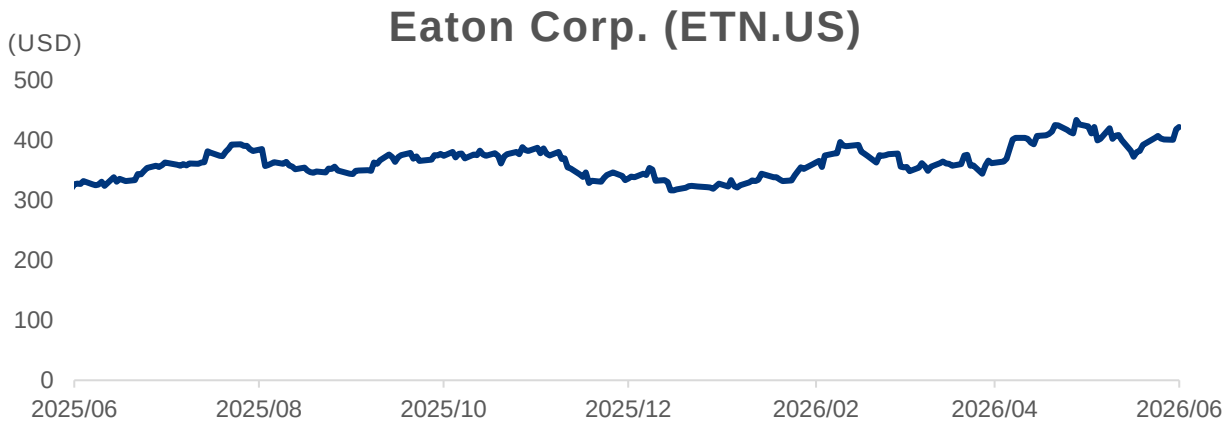
Source: Bloomberg, KGI

Grid Hardware

More Opportunity in Powering up the Electricity Grid







Source: Bloomberg, KGI

Disclaimer and Important Notice

Hong Kong: The information contained in the document herein is confidential and is not intended for general public distribution or for use by any person or entity located or residing in any jurisdiction which restricts the distribution of such information by KGI Asia Limited ("KGI") or any affiliate of KGI. Re-distribution of the document herein and any part thereof by any means is strictly prohibited. Such information shall not be regarded as an offer, invitation, solicitation or recommendation to invest in or sell any securities or investment products to any person or entity in any jurisdiction. The above information (including but not limited to general financial and market information, news services, market analysis and product information) is for general information and reference purpose only and may not be reproduced or published (in whole or in part) for any purpose without the prior written consent of KGI Asia Ltd. Such information is not intended to provide investment advice and should not be relied upon in that regard. You are advised to exercise caution, and if you are in any doubt about such information, you should seek independent professional advice.

You are advised to exercise caution and undertake your own independent review, and you should seek independent professional advice before making any investment decision. You should carefully consider whether investment is suitable in light of your own risk tolerance, financial situation, investment experience, investment objectives, investment horizon and investment knowledge.

No representation or warranty is given, whether express or implied, on the accuracy, adequacy or completeness of information provided herein. In all cases, anyone proposing to rely on or use the information contained herein should independently verify and check the accuracy, completeness, reliability and suitability of the information. Simulations, past and projected performance may not necessarily be indicative of future results. Information including the figures stated herein may not necessarily have been independently verified, and such information should not be relied upon in making investment decisions. None of KGI, its affiliates or their respective directors, officers, employees and representatives will be liable for any loss or damage of any kind (whether direct, indirect or consequential losses or other economic loss of any kind) suffered or incurred by any person or entity due to any omission, error, inaccuracy, incompleteness or otherwise, or any reliance on such information. Furthermore, none of KGI, its affiliates or their respective directors, officers, employees and representatives shall be liable for the content of information provided by or quoted from third parties.

Complex Products refers to an investment product whose terms, features and risks are not reasonably likely to be understood by a retail investor because of its complex structure. Investors should exercise caution in relation to complex products. Investors may lose the entire amount or more than the invested amount. For complex products with offering documents or information not reviewed by the Hong Kong Securities and Futures Commission (SFC), investors should exercise caution regarding the offer. For complex products described as SFC-approved, such approval does not imply official endorsement, and SFC recognition does not equate to a recommendation or assurance of the product's commercial viability or performance. Past performance data, if provided, is not indicative of future performance. Some complex products are only available to professional investors. Before making any investment decisions, investors should review the offering documents and other relevant information to understand the key nature, features, and risks of the complex products. Independent professional advice should be sought, and investors should have sufficient net assets to bear the potential risks and losses associated with the product. Members of the KGI group and their affiliates may provide services to any companies and affiliates of such companies mentioned herein. Members of the KGI group, their affiliates and their directors, officers, employees and representatives may from time to time have a position in any securities mentioned herein.

Bond investment is NOT equivalent to a time deposit. It is NOT protected under the Hong Kong Deposit Protection Scheme. Bondholders are exposed to a variety of risks, including but not limited to: (i) Credit risk - The issuer is responsible for payment of interest and repayment of principal of bonds. If the issuer defaults, the holder of bonds may not be able to receive interest and get back the principal. It should also be noted that credit ratings assigned by credit rating agencies do not guarantee the creditworthiness of the issuer; (ii) Liquidity risk - some bonds may not have active secondary markets and it would be difficult or impossible for investors to sell the bond before its maturity; (iii) Interest rate risk – When the interest rate rises, the price of a fixed rate bond will normally drop, and vice versa. If you want to sell your bond before it matures, you may get less than your purchase price. Do not invest in bond unless you fully understand and are willing to assume the risks associated with it. Please seek independent advice if you are unsure.

All investments involve risks. The prices of securities fluctuate, sometimes dramatically. The price of a security may move up or down, and may become valueless. It is as likely that losses will be incurred rather than profit made as a result of buying and selling securities. Prices of securities and fund units may go up as well as down and past performance information presented is not indicative of future performance. Investors should read the relevant fund's offering documents (including the full text of the risk factors stated therein (in particular those associated with investments in emerging markets for funds investing in emerging markets)) in detail before making any investment decision. You are advised to exercise caution and undertake your own independent review, and you should seek independent professional advice before making any investment decision. You should carefully consider whether investment is suitable in light of your own risk tolerance, financial situation, investment experience, investment objectives, investment horizon and investment knowledge.

Singapore: This document is provided for general information and circulation only, and is not an offer or a solicitation to deal in any securities or to enter into any legal relations, nor an advice or a recommendation with respect to any financial products mentioned herein. It does not have regard to the specific investment objectives, financial situation and the particular needs of any recipient hereof. You should ensure that you understand the risk(s) involved and should independently evaluate particular investments and consult an independent financial adviser before making any investment decisions. All information and opinions contained herein is based on certain assumptions, information and conditions available as at the date of this document and may be subject to change at any time without notice.