Semiconductor sector

2H26 strategy – Investment opportunities & iteration challenges

Key message

1. Semiconductor, power, optical & thermal (SPOT) components represent the key bottlenecks in AI chip product iterations.
2. Upward revisions to 2027F GPU & TPU demand are gaining momentum; CPUs emerge as another focal point.
3. Rising AI chip design complexity driving up testing demand & ASP.
4. AI spillover effects extend into non-AI sectors; Nvidia (US) production issues remain a short-term overhang.

Impact

Semiconductor, power, optical & thermal components (SPOT) represent the key bottlenecks in AI chip product iterations. Market confidence in the authenticity of AI demand is now well established, and the key debates instead center on: (1) whether AI revenue growth can keep pace with accelerating capex; and (2) whether physical constraints—such as optical integration, thermal management, and power delivery—can be effectively addressed.

Upward revisions to 2027F GPU & TPU demand are gaining momentum; CPUs emerge as another focal point. According to third-party data, Anthropic's (US) annual recurring revenue (ARR) surpassed OpenAI's (US) in April and has reached breakeven. The rapid adoption of Anthropic's Claude and Google's (US) Gemini is being reflected in upward revisions to TPU demand. Among TSMC's (2330 TT, NT$2,410, OP) CoWoS customers for next year, Nvidia (US), Broadcom (US), and MediaTek (2454 TT, NT$4,245, OP) are seeing more pronounced demand growth, underscoring robust GPU and TPU demand. Notably, AMD (US) has recently sharply revised up its CoWoS demand forecasts for next year, doubling it to 210k units. In parallel, the rise of agentic AI is driving upward revisions to server CPU demand, as CPUs account for over 40% of workloads in agentic AI inference. With major CSPs (excluding Meta (US)) expected to double CPU procurement in 2026F, we estimate a CPU supply gap of 15-20%, with shortages potentially extending into 2028F.

Rising AI chip design complexity driving up testing demand & ASP. Accelerating product iterations and increasingly complex chip designs—characterized by higher pin counts and growing chip shipment volumes—are driving strong growth in testing demand, benefiting both testing equipment and interface suppliers. In addition, TSMC's move to outsource chip-probe (CP) testing to OSATs to optimize internal capacity is prompting upward revisions to OSAT capex. Beyond GPUs, AI infrastructure is expanding into data transmission, with co-packaged optics (CPO) emerging as a key solution to raise transmission speeds. That said, optical integration increases process complexity and weighs on yields, making higher test channel counts and electro-optical co-testing critical to reducing defect rates.

AI spillover effects extend into non-AI sectors; Nvidia (US) production issues remain a short-term overhang. Strong AI demand is absorbing industry resources, leading to broad crowding-out effects. The dynamics seen with memory demand are now evident across analog ICs, mature process nodes, and passive components. Given the lack of structural expansion in memory supply, shortages are likely to persist at least through end-2027F. Demand is not the constraint—capacity is. Meanwhile, due to SK Hynix's (KR) HBM4 challenges, CoWoS re-tape-outs, and thermal design changes, we expect Nvidia's Vera Rubin shipments to fall meaningfully below initial early-year expectations. Product mix shifts may also lead to downward revisions to previously optimistic ASP assumptions across Nvidia's supply chain in 2Q26F, or even early 3Q26F.

Stocks for Action (focused on CPUs, TPUs, GPUs, testing & memory)

TSMC remains our top pick, supported by its technology leadership and attractive valuation. The firm's capex is expected to stay elevated over the next several years, and we are also optimistic about the sales growth prospects of equipment vendors. Taking technology iterations and valuations into account, Hon Precision (7769 TT, NT$6,470, OP) and Lam Research (US) are our preferred equipment names. Other recommended stocks include Winway (6515 TT, NT$8,085, OP), MPI (6223 TT, NT$6,090, OP), MediaTek, Alchip Technologies (3661 TT, NT$4,180, OP), Global Unichip (3443 TT, NT$4,845, OP), Chipbond Technology (6147 TT, NT$211.5, OP), Nanya Technology (2408 TT, NT$452.5, OP), Macronix (2337 TT, NT$155.5, OP), AMD, ASML (NL), and Texas Instruments (US).

Risks

Over-expansion of AI facilities and demand; global economic recession.

2H26F semiconductor sector outlook

Summary

1. GPUs and TPUs will continue to drive AI chip shipments growth in 2027F, while the rise of agentic AI is catalyzing strong demand for server CPUs.

2. Greater optical integration and increasing AI chip design complexity are expected to raise testing demand and ASPs; electro-optical co-testing will emerge as a key focus in 2027F.

3. The impact of AI demand is extending beyond AI-centric sectors, with memory and analog IC markets undergoing structural shifts, and other spillover effects likely.

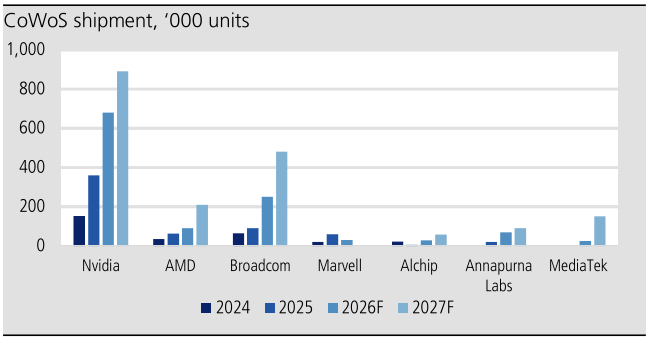
Positive signal #1 – Upward revisions to advanced packaging demand, with CPO, SoIC & CoPoS next in line

We remain optimistic in our outlook on AI semiconductor demand, although growth will no longer be driven solely by Nvidia GPUs. From an end-demand perspective, Anthropic's Claude has a relatively high share in agent-based applications and the enterprise segment, while Google's Gemini has steadily encroached on OpenAI's website traffic share. This trend suggests AI model competition is shifting away from OpenAI-dominated training demand, and toward broader inference and agent deployment. The rapid market share expansion of Claude and Gemini is being directly reflected in robust TPU demand.

This shift is also evident in ongoing upward revisions to CoWoS demand. In the last six months, 2027F CoWoS demand forecasts for Nvidia, AMD, Broadcom, and MediaTek have all been meaningfully revised up. Broadcom's and MediaTek's CoWoS demand forecasts have been raised by 60% and 90% to 480k and 150k units, respectively—well above Nvidia's 10% upward revision to 890k units—driven by greater upside from TPU demand. While Nvidia faces limited competition in the training segment, the firm's share within CSP deployments may gradually erode as adoption of ASICs accelerates. Meanwhile, AMD's 2027F CoWoS demand at TSMC has also been sharply revised up, doubling to 210k units, largely reflecting strong demand for MI450 and MI455 chips.

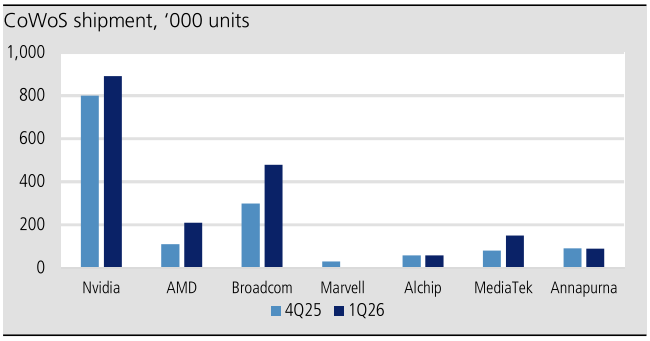
Separately, the rise of agentic AI is driving a notable recovery in general server demand, given that CPUs account for more than 40% of workloads in agentic AI inference. As major CSPs (excluding Meta) are expected to double CPU procurement in 2026F, we estimate a CPU supply gap of 15-20%, with tightness potentially persisting through 2028F. As TSMC's capacity may be insufficient to fully meet AMD's Venice CPU requirements, AMD is likely to rely more heavily on ASE Technology (ASEH; 3711 TT, NT$680, OP) and Spil (TW; unlisted) to secure sufficient packaging capacity.

Figure 1: 2025-27F CoWoS shipment forecasts by client



Source: KGI Research estimates

Figure 2: Changes in 2027F CoWoS shipments, 4Q25 vs. 1Q26



Source: KGI Research estimates

Positive signal #2 – Testing supply chain to outperform other subsectors

King Yuan Electronics (KYEC; 2449 TT, NT$337.5, OP) and ASEH have raised their 2026 capex by 28% and 24% to NT$50bn and US$8.5bn, respectively, representing 35% and 55% YoY growth amid strong AI demand. As TSMC outsources CP testing to OSAT providers to optimize internal capacity allocation, and demand related to US fab buildouts increases, we expect upside to OSAT capex. Overall, the simultaneous expansion of capex by TSMC and OSAT vendors is expected to support demand for back-end equipment, and we believe the testing segment will outgrow packaging, as the rapid iteration of AI chip designs and increasing design complexity are driving both longer test times and higher testing difficulty, boosting testing demand and prices. We believe testing equipment vendors and test-interface manufacturers will stand out as major beneficiaries.

The test interface industry has transitioned from a relatively mature segment that fluctuated with overall semiconductor shipment volumes, and with products often regarded as consumables, into a growing industry driven by rising complexity in AI and HPC chips. Increases in I/O and pin counts, package sizes, and power consumption have contributed to greater testing difficulty and longer testing time. As far as probe cards are concerned, advanced AI and HPC chips require higher pin counts, finer pitch and improved signal integrity & reliability, which in turn have pushed up both the ASP of a probe card and technical entry barriers. For test sockets, larger AI and HPC chips featuring higher power consumption and more stringent test conditions also contribute to rising ASP and consumption volumes. As a result, even though the recovery in the semiconductor sector is uneven, the test-interface industry is benefiting from AI demand and is positioned to pull off stronger sales growth in relation to the traditional cyclical pattern.

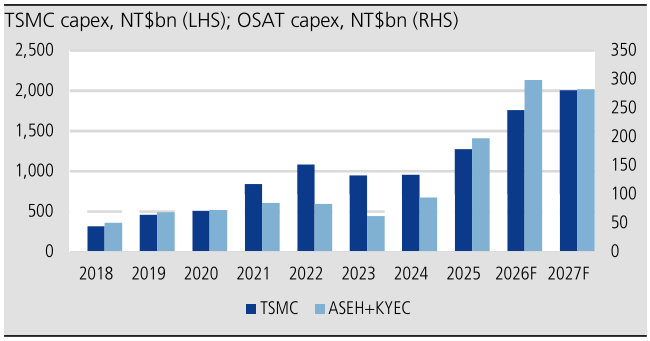
We believe that Taiwanese test interface vendors will be the primary beneficiaries of the higher demand for testing AI chips. While the broader test-interface market is fragmented among suppliers in the US, Japan, South Korea, and Europe, Taiwanese companies have established stronger customer loyalty while acquiring technical certifications within the supply chains of AI ASIC, GPU, CPU and networking firms. We estimate that Taiwanese players have captured approximately 44% of the AI test interface market versus a 16% share in the overall test interface market. As the weighting of AI applications in the test-interface market rises, Taiwanese companies are expected to benefit not only from end-market demand growth, but also from product mix improvements and market share gains, allowing them to outpace the industry average in terms of sales growth.

In the probe card segment, as process nodes advance to 3nm and below, traditional probing technologies face increasing challenges in achieving fine pitch, low contact force and effective pad damage control, thus driving MEMS probe card adoption. In the test socket segment, higher pin counts, larger package sizes and greater power consumption have raised requirements for contact stability, signal integrity, chip durability and thermal performance. As ASPs and gross margins of probe cards and test sockets have risen, optimization of product mix is expected to become the key driver of earnings growth for Taiwanese vendors from 2H26F through 2027F. We expect Taiwanese companies with established certifications among high-end AI and HPC customers, strong MEMS manufacturing capabilities and fast delivery execution, such as MPI and WinWay, to gain market share.

In addition, investment in AI infrastructure has expanded from GPU to data transmission and interconnect technologies. Co-packaged optics (CPO), as a key solution for raising bandwidth and transmission efficiency, will become a major focus in the next phase of industry upgrades. However, the introduction of optical components has increased the complexity of packaging and testing processes, adding to challenges in yield management. Compared with traditional electrical testing, electro-optical co-testing requires simultaneous validation of both optical and electrical signals, resulting in additional testing steps and higher technical barriers. This makes testing a critical factor in reducing defect

and failure rates. We expect that as CPO and optical integration gain traction, demand for related testing equipment and test interface solutions will expand.

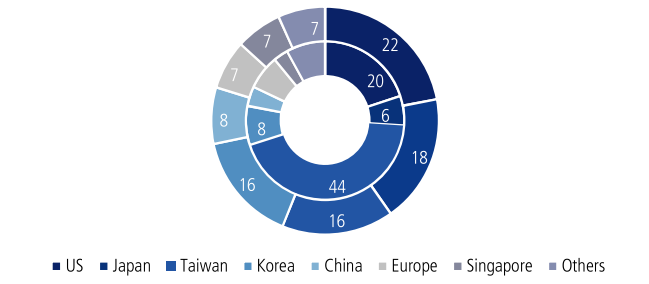
Figure 3: Capex by TSMC & OSAT providers



Source: Bloomberg; KGI Research

Figure 4: The market for AI chip testing is dominated by Taiwanese providers

Total market share, percent (outer); AI market share, percent (inner)



Source: KGI Research estimates

Figure 5: CPO testing process

Test Insertion	Insertion 1	Insertion 2	Insertion 3	Insertion 4	CPO FT		
				Joint Development	Mass Production		
Test Content	PIC	EIC-PIC	Optical Engine	ASIC + Optical Engine	ASIC + Optical Engine	ASIC	Optical Engine
	Wafer Level (Foundry)	Wafer Level (Foundry)	DIE/Chip Level (OSAT)	Package-Level(OSAT)			
				O/E	E/E	E/E	E/E
HON.PREC Test Equipment	NA			ATE Tester + ATC (3KW) Handler	ATE Tester + ATC (2KW) Handler	ATE Tester + ATC (1KW) Handler	ATE Tester + ATC (1KW) Handler

Source: Company data

Positive signal # 3 – Structural changes in memory & analog IC demand

Memory supply shortage to extend to 2027F

We expect memory demand to grow substantially in 2026-27F, driven by capex expansion by North American CSPs. We believe that general server demand will rise more than 30% YoY in 2026F. However, due to supply constraints for key components such as memory and CPUs, actual shipment growth is likely to be closer to 20% YoY. In addition, the adoption of high-density memory modules, including 96GB and 128GB modules, will pick up, prompting growth in server DRAM content by approximately 30% YoY and total server DRAM bit demand by 46% YoY. Looking ahead to 2027F, we expect general server shipment growth of more than 20% YoY, which will be the most important driver of memory demand.

On the supply side, the memory industry is facing bottlenecks, including limited availability of advanced cleanrooms and tight supply of equipment for advanced processes. We estimate DRAM wafer input will grow by a respective 9% and 10% YoY in 2026-27F, while NAND flash wafer inputs are projected at a 2% decline and 4% growth YoY in the same period. Although new capacity from facilities such as SK Hynix's Yongin fab, Micron's (US) ID1 and Tongluo fabs, Nanya Tech's Fab 5A, CXMT's (CN) Shanghai fab, and YMTC's (CN) Fab 3 will likely come online from 2H26F through 2027F, most of these fabs are still at the early stage of buildout and will contribute only incremental supply gains in the near term. In addition, the advanced DRAM capacities of Samsung (KR) and SK Hynix require a

significant number of EUV equipment, but equipment supply from ASML is constrained. At the same time, memory makers are prioritizing resources for DRAM and HBM products, which limits the pace of NAND flash capacity expansions.

Figure 6: 2021-27F global wafer inputs for DRAM

kwpm	2021	2022	2023	2024	2025	2026F	2027F	2028F
Wafer input	**1,495**	**1,593**	**1,363**	**1,645**	**1,916**	**2,095**	**2,304**	**2,615**
Samsung	584	653	527	614	655	690	783	818
SK Hynix	356	393	352	416	528	590	639	708
Micron	355	353	278	314	333	358	376	414
Nanya Technology	71	68	54	56	58	65	73	105
Winbond	26	21	24	24	24	24	31	45
PSMC	47	43	27	39	42	45	45	45
CXMT	50	54	90	173	260	298	326	403
JHICC	6	9	10	10	17	26	30	30
YMTC							1	49
YoY (%)		**7**	**(14)**	**21**	**16**	**9**	**10**	**13**
Samsung		12	(19)	16	7	5	13	4
SK Hynix		10	(10)	18	27	12	8	11
Micron		(0)	(21)	13	6	8	5	10
Nanya Technology		(4)	(21)	2	4	12	13	44
Winbond		(19)	14	0	(2)	1	29	44
PSMC		(9)	(37)	44	8	7	0	0
CXMT		7	69	91	51	14	10	23
JHICC		40	14	0	70	54	14	0
YMTC								3,800

Source: TrendForce; KGI Research estimates

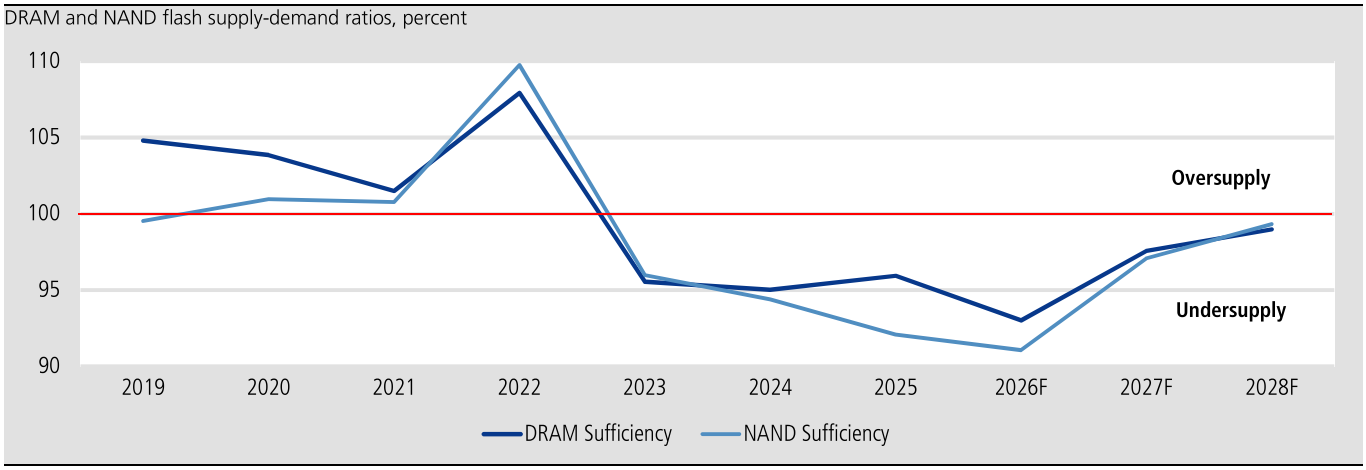
Figure 7: 2021-27F global wafer inputs for NAND flash

kwpm	2021	2022	2023	2024	2025	2026F	2027F	2028F
Wafer input	**1,527**	**1,698**	**1,397**	**1,428**	**1,354**	**1,333**	**1,381**	**1,531**
Samsung	574	636	489	473	414	348	310	333
Kioxia/Sandisk	496	475	395	443	400	409	443	498
SK Hynix	195	293	234	215	233	234	230	230
Micron	170	169	134	143	130	130	130	139
YMTC	66	98	120	125	148	180	219	269
PSMC	3	5	4	4	5	6	8	8
Winbond	6	7	8	13	13	14	19	28
Macronix	11	13	12	11	8	9	17	22
SMIC	5	4	3	3	4	5	5	5
YoY (%)		**11**	**(18)**	**2**	**(5)**	**(2)**	**4**	**11**
Samsung		11	(23)	(3)	(12)	(16)	(11)	7
Kioxia/Sandisk		(4)	(17)	12	(10)	2	8	12
SK Hynix		50	(20)	(8)	8	1	(2)	0
Micron		(1)	(21)	7	(9)	0	0	7
YMTC		47	23	4	18	22	22	23
PSMC		50	(22)	14	31	14	33	0
Winbond		12	7	70	(2)	8	41	46
Macronix		14	(8)	(4)	(27)	6	100	29
SMIC		(20)	(25)	0	42	6	11	0

Source: TrendForce; KGI estimates

We estimate that DRAM and NAND flash will each suffer a supply shortfall of 8-9% in 2026F. In 2027F, with demand continuing to grow strongly and supply expansions constrained, DRAM and NAND flash are expected to see respective supply shortfalls of 2.4% and 2.9%. Overall, the continuation of tight supply conditions into 2027F is becoming the consensus view across the industry.

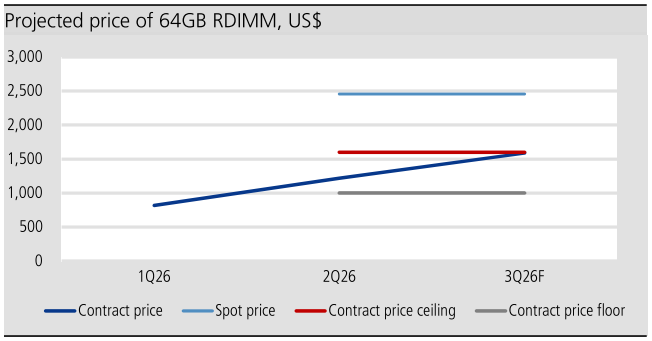
Figure 8: Memory supply-demand ratio



Source: TrendForce; KGI Research estimates

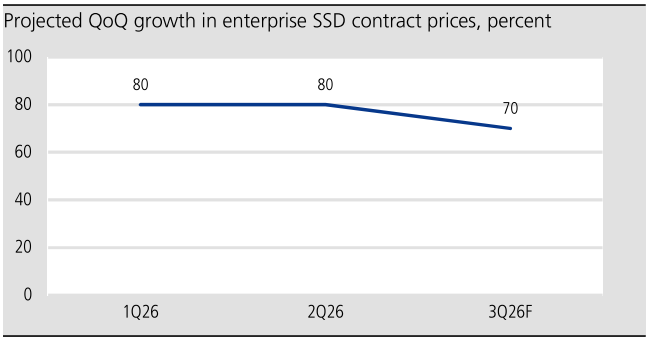
With supply tight, customers have been eager to secure long-term agreements (LTAs) with memory makers, and we note that most major US CSPs have already closed deals with their suppliers and the trend is spreading to OEMs. Based on our understanding, an LTA for RDIMM generally comes with a specific price band, and the floor price is only below the benchmark of a 2Q26 contract by a slim margin, while the ceiling price is higher by approximately 30-40%. As the supply volume through LTAs is still insufficient to meet demand, excess demand must be fulfilled through the spot market. At present, spot prices of 64GB RDIMMs are more than double that of contract prices, indicating upside for prices. Enterprise SSD prices increased by approximately 80% in both 1Q-2Q26, and memory makers are expected to maintain aggressive pricing in 3Q26F, with price hikes above 70% QoQ. In contrast, consumer products such as SODIMM, LPDDR, and client SSDs are less likely to match the price gains seen in server-related products, as producers may struggle to pass through additional costs.

Figure 9: 3Q26F RDIMM official & market price estimates



Source: KGI Research estimates

Figure 10: 3Q26F enterprise SSD contract price estimates



Source: KGI Research estimates

Analog semiconductor data center sales to be the bright spot

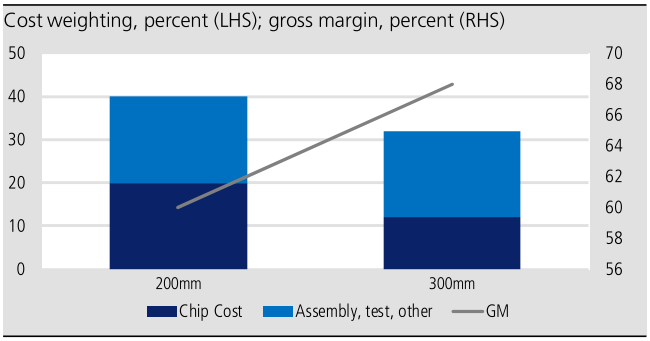
The analog semiconductor industry has undergone a prolonged inventory adjustment period since 2022. Although demand for consumer electronics is susceptible to fluctuations in memory prices, we expect demand for automotive and industrial applications to sustain positive growth in 2H26F. The sales weighting of data center applications is likely to increase from a low base of 5-10%, driven by expanding AI demand. Based on our supply chain checks, data center-related products command price premiums of approximately 20-30% over general industrial and automotive offerings. Meanwhile, due to the cannibalization effect from memory demand, 8-inch wafer capacity is undersupplied, forcing analog semiconductor vendors to either raise prices or accelerate migration to 12-inch wafers. Against this backdrop, we are upbeat about the prospects of Texas Instruments, as the firm has established a first-mover advantage in 12-inch capacity relative to peers, and has maintained a higher weighting of in-house production. This enables the firm to mitigate the price volatility associated with outsourced 8-inch fabs. As the firm's 12-inch fabs enter mass production in 2027F, overall depreciation expenses are expected to decline steadily, reinforcing its cost advantage. We believe Texas Instruments is well positioned within the competitive landscape, with long-term profitability likely to outperform peers.

Figure 11: Data center sales weightings of major analog semiconductor vendors

	2025	2026F	Comment
MPWR	25	32	AI-related revenue reached 32.7% in 1Q26, and the full-year AI revenue growth forecast has been raised from 50% to 85%.
ADI	17	22	ATE and data center businesses are expected to achieve double-digit annual growth.
Infineon	9	14	AI power revenue is projected to reach €1.5bn / €2.5bn in 2026/27, representing approximately 9.3/13.6% of total revenue.
TXN	9	11	Demand for AI-specific power solutions is expected to increase significantly during 2H26-2027.
MCHP	6	8	
ON	4	6	AI data center revenue is expected to double in 2026, while the related TAM is projected to grow at a 40% CAGR.
STM	0	4	AI data center revenue is expected to exceed US$500mn / US$1bn in 2026/27, accounting for approximately 3.5/6.1% of total revenue.
NXPI	2	4	Data center revenue is expected to increase from approximately US$200mn in 2025 (1.6% of revenue) to over US$500mn in 2026 (3.5% of revenue).

Source: Company data; KGI Research

Figure 12: Migration to 12-inch wafers will significantly reduce the unit cost of chips



Source: KGI Research estimates

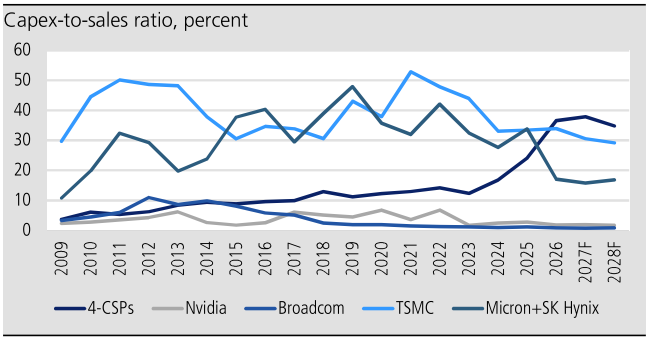
Potential reversal in CSP capital intensity; near-term noise in Nvidia production

Driven by AI investment, CSP capital intensity has risen rapidly and now exceeds that of semiconductor companies. While near-term attention is focused on the strong demand for GPUs, TPUs, networking equipment, and data center buildouts, risks may begin to emerge looking into 2H26–1H27F. If CSP revenue growth fails to accelerate, or if AI-related revenues are insufficient to sustain ongoing capex expansions, concerns around returns on investments, depreciation burdens, and valuation sustainability could intensify. Accordingly, whether CSP capital intensity begins to fall will be a key indicator in assessing the durability of the AI investment cycle beyond 2H27F. Should revenue fall short of expectations, valuations across the broader AI supply chain could come under pressure.

Separately, near-term uncertainties are emerging within the Nvidia supply chain. Factors including SK Hynix's HBM4-related challenges, CoWoS re-tape-outs, and changes in thermal design are expected to weigh on Vera Rubin shipment volumes, which we believe will fall short of market expectations. We now estimate Vera Rubin chip shipments at approximately 2 million units this year. However, after accounting for yield and lead time constraints, effective downstream availability may be limited to between 50–60%, compared with our earlier estimate of 3 million units at the start of the year. In addition, a declining weighting of Vera Rubin within product mixes, alongside a shift toward less favorable configurations, may undercut supply chain expectations. We expect ASP forecasts for certain suppliers to face downward revisions as early as 2Q26F or early 3Q26F.

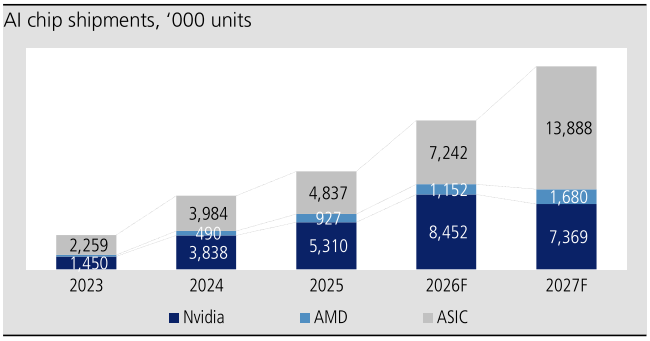
Overall, while Nvidia-related AI demand is fundamentally strong, near-term product mix shifts represent a key risk to supply chain earnings expectations.

Figure 13: Capital intensity of CSPs vs. semiconductor companies



Source: Bloomberg; KGI Research

Figure 14: 2023–27 AI chip shipment forecasts



Source: KGI Research estimates

TSMC & CoWoS

Following several rounds of upward capacity revisions, we expect TSMC's CoWoS monthly capacity to increase from approximately 120k wafers by end-2026F to 170k wafers by the end-2027F. While there are rumors suggesting upside to as much as 200k wafers—driven by factors such as Phase II expansion of AP7 and the possible conversion of certain 6-inch fabs exiting production into CoWoS-related capacity—we have yet to obtain sufficient evidence to verify these developments. As such, we maintain a more conservative capacity outlook.

We expect the strongest growth in CoWoS demand next year to come from Broadcom, MediaTek, and Nvidia, reflecting robust demand for GPUs and TPUs. In addition, since March, we have observed a notable upward revision to AMD's 2027F CoWoS demand outlook. We now estimate AMD's demand will double to approximately 210k wafers next year, driven by strong demand for MI450 and MI455 chips.

Nvidia

We expect Nvidia to secure approximately 60% and 50% of TSMC's CoWoS capacity in 2026-27F, respectively. Compared with six months ago, our 2027 CoWoS demand forecast for Nvidia has been revised up by close to 10% to 890k wafers (albeit slightly lower than our estimate three months ago), reflecting strength in overall AI demand. However, the downside to Vera Rubin shipments is expected to be largely offset by stronger orders for Blackwell.

We estimate Nvidia's Blackwell and Vera Rubin chip shipments at approximately 5.4mn and 2.0mn units, respectively, implying combined YoY growth of 59%. While overall demand for Nvidia's AI chips is robust, lower Vera Rubin shipments may lead to less favorable product mixes than previously expected. During its latest earnings call, Nvidia highlighted surging demand for CPUs, with a total addressable market estimated at approximately US$200bn. The company expects CPU-related revenue to reach US$20bn this year, including Vera, Grace, and CPUs integrated within AI systems. While part of this revenue reflects bundled CPUs within AI racks, the implied scale is already comparable to the annual server CPU revenue of Intel (US) and AMD. Currently, packaging for Vera CPUs is primarily handled by TSMC and Amkor (US).

TPU

Compared with six months ago, we have raised our 2027F CoWoS demand forecasts for Broadcom and MediaTek by 60% and 90%, respectively, to 480k and 150k wafers. This upward revision reflects upside to TPU demand, driven by the strong growth of workloads by Claude and Gemini.

In addition, recent supply chain checks indicate that Broadcom's 2028 v.9 project (Pumafish) has been canceled, with an upgraded version of Sunfish expected to serve as a transitional solution ahead of the v.10 rollout. While it is uncertain whether Humufish will become the primary TPU platform in 2028, its development is likely to support MediaTek's TPU market share. We expect Intel's EMIB-T to ramp from 4Q27F, earlier than our previous estimate, Humufish's shipment timing and product mix are also likely to be better than prior expectations. Assuming ASP growth of 200%—driven by a higher number of I/O dies and adoption of the N2 process—and shipment growth of 50%, alongside the addition of a second ASIC customer, we forecast MediaTek's AI ASIC revenue to grow by 66% YoY to US$50bn in 2028F, accounting for approximately 67% of total revenue.

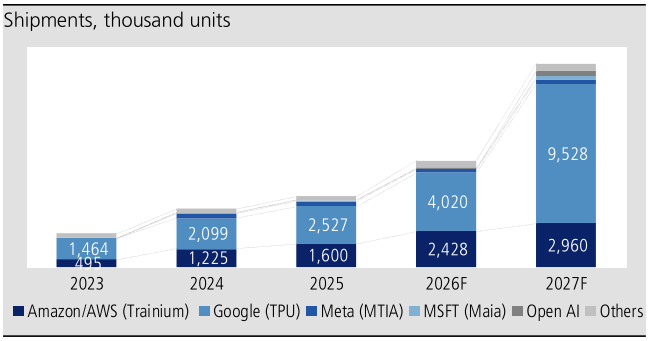
OSAT

KYEC and ASEH raised 2026F capex plans by 28% and 24% in 1Q26, to NT$50bn and US$8.5bn, respectively, implying YoY growth of 35% and 55%, and reflecting sustained strength in AI-driven demand. ASEH's on-substrate sales are expected to benefit from TSMC's expansion of CoWoS capacity, while also serving as a key CoWoS partner for AMD's new Venice CPUs. Meanwhile, KYEC is well positioned to benefit from longer testing times for AI GPUs and increasing demand for ASIC burn-in testing. In addition, as TSMC outsources CP testing to OSAT providers to optimize internal capacity allocation, we believe capex for OSAT players has upside. With foundries facing constraints in bumping capacity—and both CoWoS and EMIB-T requiring mixed-bumping to accommodate different die heights in heterogeneous packages—OSAT companies may also indirectly benefit from the expansion of EMIB-T production.

ASIC industry – Demand shifting toward inference

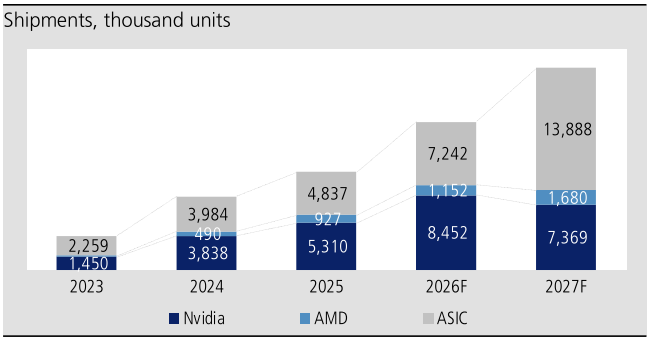
As global AI compute demand shifts from model training toward commercial deployment—driven by inference workloads and the rise of agentic AI—we expect inference to become the dominant segment of AI compute demand. Historically, model training has relied primarily on general-purpose GPUs. However, inference introduces variable costs with each user interaction, increasing the need for low latency, high throughput, and customization. In this context, ASICs—offering superior cost efficiency and the ability to be tailored to specific CSP requirements—are well positioned to capture this demand sweet spot. With hyperscalers such as Google and Amazon (US) accelerating their in-house silicon strategies, we forecast ASIC chip shipments to grow by 50.1% and 35.3% in 2026-27, respectively, compared with moderating growth in general-purpose GPUs at 49.7% and 91.8%. Among key beneficiaries, Alchip is expected to see strong sales growth driven by the ramp-up of AWS' (US) Trainium 3, with revenues likely to increase QoQ in 2H26F. We anticipate a valuation re-rating alongside an acceleration in revenue and earnings growth.

Figure 15: ASIC chip shipment forecast



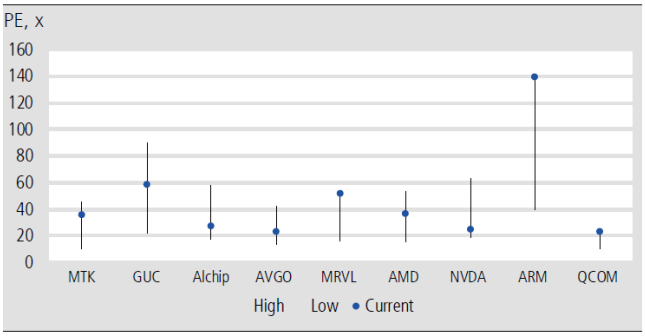
Source: KGI Research estimates

Figure 16: ASIC shipments approaching GPUs, but still lagging in output value



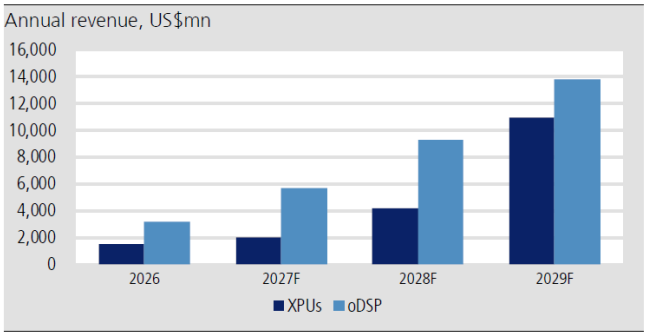
Source: KGI Research estimates

Figure 17: 2023–26F PE of AI-related stocks



Source: Bloomberg; KGI Research
Note: Current PE is based on 2027 forecast EPS

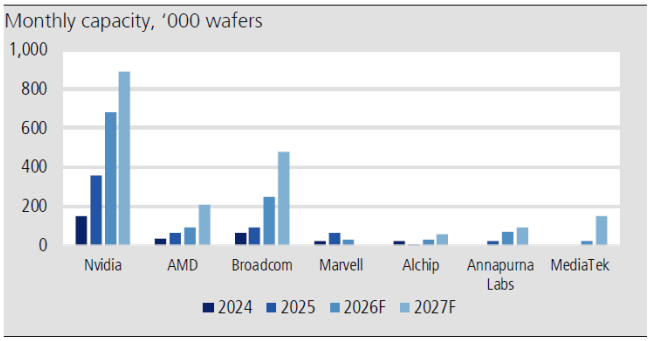
Figure 18: Strong growth in Marvell's oDSP & ASIC segments



Source: KGI Research estimates

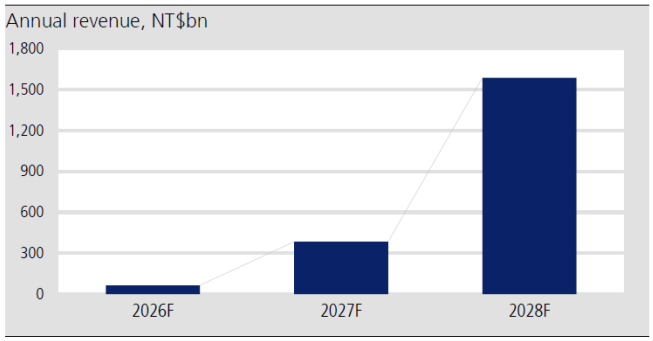
Finally, despite ongoing tightness in CoWoS capacity, the ability of ASIC players to execute shipments is dependent on backend packaging rather than front-end wafer fabrication. We also observe that CSPs are accelerating the validation of alternative advanced packaging solutions beyond TSMC. Companies such as ASEH, Amkor, KYEC, and Intel are well positioned to benefit from the spillover effects of rising advanced packaging demand, which should support upward revisions to ASIC shipment volumes. Among these, MediaTek's collaboration with Intel on EMIB-T for Google's Humufish (v.8e) is expected to generate approximately US$50bn in revenue by 2028F, with upside potential.

Figure 19: 2025–27F TSMC CoWoS client capacity allocations



Source: Bloomberg; KGI Research estimates

Figure 20: Projected revenue from TPU products for MediaTek



Source: KGI Research estimates

All the above named KGI analyst(s) is SFC licensed person accredited to KGI Asia Ltd to carry on the relevant regulated activities. Each of them and/or his/her associate(s) does not have any financial interest in the respectively covered stock, issuer and/or new listing applicant.