

The Governed Compute Order

**NVIDIA, China and Europe after the first
AI chip war**

A strategic investor brief on market concentration, export controls, HBM, energy, sovereign infrastructure, and Europe's narrow path to leverage.

INVESTOR BRIEF

Investor Brief, No. 02

The Governed Compute Order

NVIDIA, China and Europe after the first AI chip war

A strategic investor brief on market concentration, export controls, HBM, energy, sovereign infrastructure, and Europe's narrow path to leverage.

Publication date	Late August 2026
Publisher	RefleXio Intelligence
Series	Investor Briefs, No. 02
Version	1.0, post-NVIDIA Q2 FY27 update
Prepared by	RefleXio Intelligence Editorial Desk

Independent intelligence on AI, compute and the geopolitics of the stack.



Contents

Contents	2
Key numbers at a glance.....	4
1. Executive summary	5
2. Central thesis	6
3. What changed after April 2026	7
3.1 NVIDIA crossed into a new revenue regime	7
3.2 Export control moved from a wall to a valve	8
3.3 China moved further toward system-level substitution.....	8
3.4 Europe moved from program architecture toward procurement scale	8
4. Market structure and the principal players.....	9
4.1 NVIDIA.....	10
4.2 AMD.....	10
4.3 Huawei and the Chinese domestic stack	10
4.4 TSMC and the fabrication center of gravity.....	11
4.5 ASML and Europe's upstream leverage	11
4.6 HBM suppliers	11
5. Export controls and the political price of compute	11
6. The physical bottlenecks the market still underprices.....	12
6.1 Memory bandwidth and packaging	12
6.2 Networking and scale architecture.....	12
6.3 Power, grid connection and cooling	13
6.4 Capital and financing	13
7. Europe's position in the governed compute order	13
7.1 Europe's real assets	13
7.2 Europe's structural weaknesses	14
7.3 Why Chips Act 2.0 must be different.....	14
7.4 AI Factories and Gigafactories.....	14
7.5 Europe's investable strategy.....	15
8. Company and segment scorecard.....	15
9. Scenario framework through 2029.....	17
10. Investor indicator dashboard	18
10.1 Decision rules	19
11. What this means for investors now.....	19

12. Key signals for the next 90 days.....	20
13. Conclusions.....	20
14. Methodology and scope	21
14.1 Purpose	21
14.2 Source hierarchy.....	21
14.3 Evidence classification.....	21
14.4 Limitations	22
15. References	22
16. Disclaimer.....	23

Key numbers at a glance

Ten figures frame the updated thesis. Company figures use the latest reported quarter available on August 28, 2026. Market and policy figures retain the source date shown in the references.

Metric	Latest evidence	Strategic reading	Source
NVIDIA Q2 FY27 revenue	\$96.2B, up 106% year over year	The physical AI cycle is still accelerating	COMPANY FILING NVIDIA Q2 FY2027 results, Aug 26 2026
NVIDIA Data Center revenue	\$89.0B, up 117% year over year	The revenue pool remains highly concentrated	COMPANY FILING NVIDIA Q2 FY2027 results, Aug 26 2026
NVIDIA Q3 FY27 guide	\$108.0B ±2%	No China Data Center compute assumed	COMPANY FILING NVIDIA Q2 FY2027 results, Aug 26 2026 (guidance)
NVIDIA gross margin	75.0% GAAP	Pricing power remains exceptional	COMPANY FILING NVIDIA Q2 FY2027 results, Aug 26 2026
AMD Q2 2026 revenue	\$11.5B	Alternative scale is growing but remains far smaller	COMPANY FILING AMD Q2 2026 results, Aug 4 2026
Micron fiscal Q3 2026 revenue	\$41.46B	Memory is no longer a supporting character	COMPANY FILING Micron fiscal Q3 2026 results, Jun 24 2026
ASML Q2 2026 sales	€9.3B	Europe owns a critical upstream choke point	COMPANY FILING ASML Q2 2026 results, Jul 15 2026
Data center electricity	485 TWh in 2025, about 950 TWh in 2030	Grid access is becoming a compute allocation mechanism	OFFICIAL STATISTIC IEA, Key Questions on Energy and AI, 2026
EU AI Factories	19 factories plus 13 antennas	Deployment access is broadening	OFFICIAL STATISTIC EuroHPC JU, AI Factories, accessed Aug 28 2026
EU Gigafactory call	Up to 7 sites, more than €30B investment	Europe is finally moving toward scale	OFFICIAL STATISTIC European Commission, AI Gigafactories call, Jul 30 2026

Table 1. Key numbers at a glance.

1. Executive summary

The first phase of the AI chip war was organized around denial. Washington attempted to slow China's access to frontier accelerators and manufacturing equipment. China responded by increasing domestic substitution, redesigning systems around available silicon and treating compute as strategic infrastructure. That phase did not end. It became more complex.

The second phase is organized around governed access. The direction of travel is not one-way. Washington rescinded the AI Diffusion Rule in May 2025 (U.S. Department of Commerce, 2025) before moving to case-by-case licensing in January 2026, which is why the current regime is better read as a valve than as a wall. The United States is no longer relying on a single prohibition architecture. In January 2026, the Bureau of Industry and Security moved selected H200, MI325X, and comparable exports to China into case-by-case review under defined security conditions (U.S. Department of Commerce, 2026). NVIDIA nevertheless excluded China Data Center compute from its August guidance. Policy has become more conditional while commercial planning remains conservative.

The market has also become more concentrated. NVIDIA reported \$96.2 billion of quarterly revenue and \$89.0 billion from Data Center for the quarter ended July 26, 2026 (NVIDIA, 2026b). One direct customer represented 16% of revenue, per the 10-Q (NVIDIA, 2026a), while customer location disclosures show how little geographic labels reveal about final deployment. This is not simply vendor dominance. It is a financing, supply allocation, and customer concentration regime.

China's response is increasingly systemic. Huawei's published roadmap couples Ascend processors with UnifiedBus, Atlas SuperPoDs, and a domestic software environment (Huawei, 2025a, 2025b, 2026). The purpose is not to win a benchmark with one device. It is to compensate for component constraints through larger systems, more aggressive parallelism, local availability, and policy alignment. China is converting a frontier gap into a deployment architecture. Alibaba's August 2026 results, which pair full-stack AI with its strongest cloud growth in twenty-two quarters, show the commercial side of the same substitution (Alibaba Group, 2026).

Europe occupies the most uncomfortable and potentially valuable position. It lacks a frontier accelerator champion at NVIDIA scale, depends on U.S. software and Asian fabrication, and faces slower grid connections than its ambitions require. Yet Europe controls ASML, hosts leading research through imec and CEA-Leti, operates EuroHPC infrastructure, and has now launched a tender for up to seven AI Gigafactories. Its credible strategy is not autarky. It is selective sovereignty built around lithography, trusted deployment, energy discipline, regulated demand, and procurement scale.

Bottom line. The market is underpricing the transition from a chip contest to a governed compute order. Value accrues to companies and jurisdictions that can coordinate silicon, memory, packaging, networking, power, software, finance, and permission. Peak FLOPS still matter. The right to deploy them, at scale, and under acceptable rules, matters more.

The emerging governed compute order

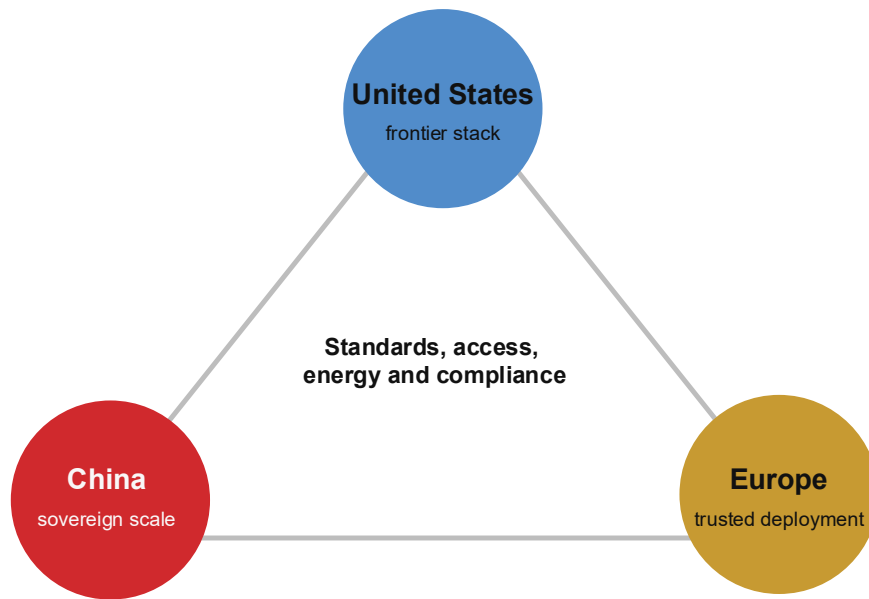


Figure 1. The governed compute order. RefleXio analysis, August 2026.

Source note RefleXio Intelligence analysis.

2. Central thesis

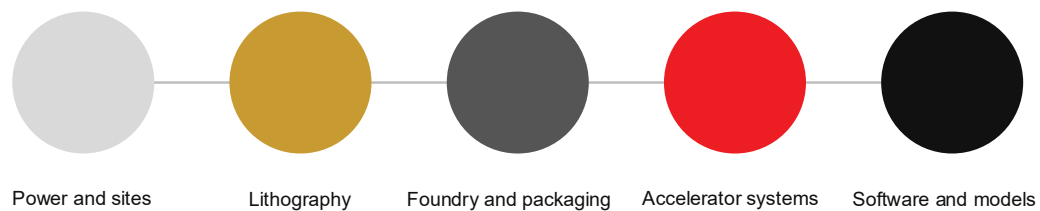
AI infrastructure is now a regulated industrial system. The winning unit is not the accelerator. It is the deployable cluster whose components can be financed, imported, powered, cooled, programmed, and legally operated. This shifts the valuation lens from product comparison toward ecosystem coordination.

Thesis. NVIDIA retains the strongest commercial stack, but its moat is increasingly exposed to policy decisions, customer concentration, and infrastructure constraints outside the chip itself. China is building an alternative stack whose advantage is availability and alignment. Europe can create leverage only where it aggregates demand and turns its upstream assets into deployment power.

- The United States controls the frontier commercial stack through accelerator design, hyperscaler demand, capital markets, and export jurisdiction.
- China uses state alignment, domestic procurement, systems engineering, and software adaptation to reduce the penalty of weaker individual components.
- Taiwan and South Korea remain indispensable through leading-edge fabrication, advanced packaging and HBM.
- Europe owns lithography leverage and a credible trusted-compute market, but it does not yet own the full accelerator-to-cloud chain.

- Power availability, grid queues and cooling are now market variables rather than engineering footnotes.

Where strategic control accumulates



The moat is no longer a chip. It is the ability to coordinate every constrained layer.

Figure 2. Strategic control accumulates across the stack, not inside one chip. Reflexio analysis.

Source note Reflexio Intelligence analysis. Layer control is a qualitative Reflexio judgment, not a measured share.

3. What changed after April 2026

3.1 NVIDIA crossed into a new revenue regime

NVIDIA's Q2 FY27 result removed any serious argument that the infrastructure cycle had already peaked. Revenue more than doubled year over year, up 106% to \$96.2 billion. Data Center reached \$89.0 billion, and gross margin returned to 75.0%. The company guided the following quarter to \$108.0 billion (NVIDIA, 2026b) while continuing to assume no China Data Center compute revenue. The contradiction is strategically important. China can be commercially absent from guidance while remaining central to policy, product design, and investor risk.

Quarterly revenue scale, selected AI hardware companies

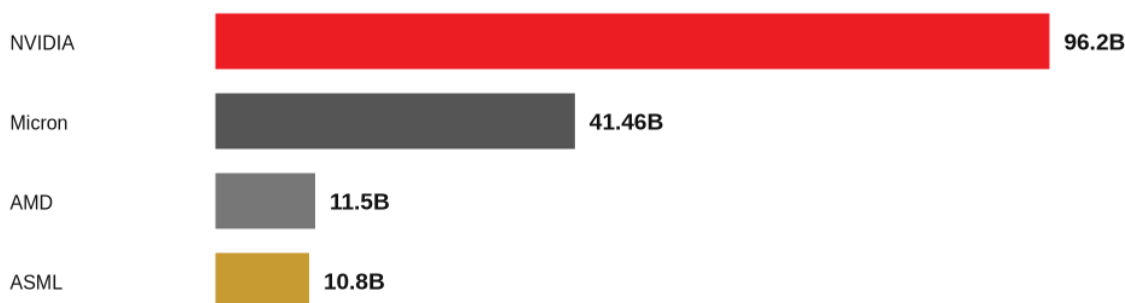


Figure 3. Selected latest reported quarterly revenue in US dollars. ASML reports in euros; its €9.3 billion figure is shown converted to US dollars at a EUR/USD rate of approximately 1.16 as of July 15, 2026. The fiscal periods differ across the four companies, so the bars compare scale, not like-for-like quarters.

3.2 Export control moved from a wall to a valve

The January 2026 BIS rule introduced case-by-case review for H200, MI325X, and comparable products when specified security requirements are met. This does not normalize unrestricted sales. It converts access into a policy valve. Washington can calibrate volume, customer, location, and security conditions without abandoning the strategic objective of limiting China's frontier capability.

For investors, this increases optionality and uncertainty at the same time. A license can unlock revenue, but it can also be delayed, narrowed, or reversed. Product roadmaps must now account for regulatory latency as though it were another manufacturing lead time.

3.3 China moved further toward system-level substitution

Huawei's public roadmap extends from Ascend 950PR and 950DT to the 960 and 970 series. Its Atlas 950 SuperPoD is designed to scale to 8,192 NPUs, while Atlas 960 is planned around 15,488 NPUs. These are vendor claims and should not be treated as independent benchmark results. They are still strategically meaningful because they reveal the chosen response to lithography, memory, and yield constraints. China intends to build larger coordinated systems around silicon it can obtain and control.

3.4 Europe moved from program architecture toward procurement scale

By August 2026, EuroHPC listed 19 AI Factories and 13 AI Factory Antennas (EuroHPC JU, 2026a). The Commission's July tender seeks up to seven AI Gigafactories and aims to unlock more than €30 billion in combined investment (European Commission, 2026a; EuroHPC JU, 2026b). The change is not only numerical. Factories broaden access and services. Gigafactories are intended to train and operate frontier-scale models. Europe is attempting to connect public compute, private capital and industrial demand.

Europe moves from factories to gigafactories

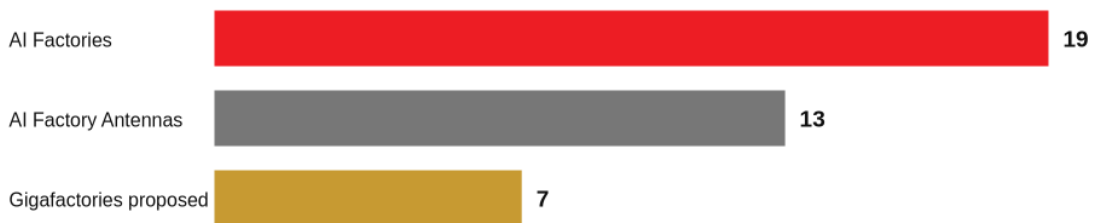


Figure 4. The European compute program now spans factories, antennas and proposed gigafactories.

Source note Reflexio Intelligence analysis, drawing on European Commission and EuroHPC JU data cited in this section.

4. Market structure and the principal players

The market should be read as a set of constrained layers. Each layer has a different concentration profile, investment cycle, and geopolitical exposure. The strongest companies are those positioned at a bottleneck or able to coordinate several bottlenecks at once.

Layer	Principal players	Current source of power	Primary vulnerability
Accelerators	NVIDIA, AMD, Huawei, hyperscaler ASIC teams	Architecture, software and system design	Export policy, customer concentration and supply allocation
Foundry	TSMC, Samsung, Intel Foundry, SMIC	Process technology, yield and customer trust	Capital intensity, geopolitics and equipment access
Lithography	ASML	EUV monopoly and installed base	Export controls, supplier depth and geopolitical pressure
HBM and memory	SK hynix, Samsung, Micron	Bandwidth, stacking, qualification, and supply discipline	Cycle risk, packaging dependencies and customer concentration
Advanced packaging	TSMC, ASE, Amkor, Intel, Samsung	CoWoS-class capacity, substrates and integration	Long lead times and execution complexity
Networking	NVIDIA, Broadcom, Marvell, Arista	Scale-up and scale-out bandwidth	Custom silicon substitution and architecture shifts
Cloud demand	Microsoft, Amazon, Alphabet, Meta, Oracle	Capital, customers and model distribution	Return on invested capital, power and public resistance

Layer	Principal players	Current source of power	Primary vulnerability
European compute	EuroHPC, national hosts, sovereign cloud providers	Trusted access and public procurement	Fragmentation, utilization and non-European silicon

Table 2. Market structure and the principal players.

4.1 NVIDIA

NVIDIA remains the reference system company. Its advantage combines accelerator cadence, NVLink, networking, rack-scale design, libraries, developer familiarity, and the ability to secure supplier capacity. NVIDIA's published roadmap extends the strategy from a device family toward a six-chip Rubin platform. That is a stated roadmap, not shipped product evidence. The company's economic evidence is stronger than any competitor's narrative. At the same time, \$96.2 billion of quarterly revenue makes the base harder to compound, and the 10-Q shows concentration among a small number of direct customers.

The China problem is now structural rather than temporary. NVIDIA can grow without China in the short term, as its guidance demonstrates. It cannot ignore a market large enough to finance an alternative ecosystem. Every restricted sale reduces current revenue, but every permanent exclusion also encourages local software, model optimization, and procurement to migrate away from CUDA.

4.2 AMD

AMD is the most credible merchant alternative in the Western stack. Q2 2026 revenue reached \$11.5 billion, with management reporting that Data Center revenue more than doubled year over year (AMD, 2026a, 2026b). EPYC provides a strategically useful CPU position while Instinct and Helios expand the accelerator and rack-scale offer. The company benefits from buyer demand for a second source, but it still faces a software and deployment gap relative to NVIDIA.

The BIS decision to review MI325X-class exports case by case places AMD inside the same policy perimeter as NVIDIA. That is a validation of technical relevance and a reminder that commercial opportunity in China is governed rather than open.

4.3 Huawei and the Chinese domestic stack

Huawei is best understood as an ecosystem integrator under constraint. Its public claims emphasize NPU count, interconnect, memory capacity, and unified system operation. The objective is to make local models and applications run at acceptable economics on domestically aligned infrastructure, even if the best individual accelerator remains behind the U.S. frontier.

The critical uncertainty is manufacturing. Public disclosures do not provide the independent yield, volume, and HBM evidence required for a clean comparison. The right analytical response is neither dismissal nor credulity. Huawei should be scored higher on sovereignty, deployment access, and state-backed demand than on independently verified performance and manufacturing transparency.

4.4 TSMC and the fabrication center of gravity

TSMC remains the manufacturing center of gravity for the frontier stack. Its Q2 2026 materials guided Q3 revenue to \$39.0 billion to \$40.2 billion and gross margin to 65.5% to 67.5% (TSMC, 2026b). The company's earlier 2026 outlook described AI accelerator revenue as a high-teens share of 2025 sales and forecast close to 30% U.S. dollar revenue growth for 2026 (TSMC, 2025, 2026a). This is evidence that the AI cycle is entering foundry planning, not merely accelerator order books.

The strategic risk is obvious. A large share of the world's frontier logic still depends on Taiwan. U.S., Japanese and European expansion improves geographic resilience, but it does not quickly reproduce the density of suppliers, engineering practice, and yield learning present in Taiwan.

4.5 ASML and Europe's upstream leverage

ASML is Europe's strongest semiconductor asset because it occupies a genuine global choke point. Q2 2026 sales reached €9.3 billion, about \$10.8 billion, net income €2.9 billion and gross margin 54.0%. Management raised its 2026 sales outlook to €43 billion to €45 billion (ASML, 2026) and linked the improvement to AI-related Logic and Memory investment.

This leverage is powerful but incomplete. ASML enables leading-edge capacity. It does not determine where the resulting chips are packaged, installed or offered as cloud services. Export restrictions can protect allied advantage while also encouraging substitution efforts and limiting ASML's addressable market. Europe owns an indispensable tool but not the orchestra.

4.6 HBM suppliers

HBM has moved from component detail to strategic allocation. SK hynix reported record Q2 2026 results and the start of HBM4 mass shipments (SK hynix, 2025, 2026). Micron reported \$41.46 billion of fiscal Q3 2026 revenue (Micron, 2026b) and entered high-volume HBM4 production for NVIDIA's Vera Rubin platform (Micron, 2026a). The figures show that memory economics are being reset by AI bandwidth demand.

The transition to HBM4 also binds memory more tightly to logic and packaging. Qualification cycles, base-die choices, thermal behavior, and advanced packaging capacity create switching costs. Memory vendors gain pricing power, while system vendors gain another reason to lock supply through long-term collaboration.

5. Export controls and the political price of compute

Export controls are often modeled as a binary variable. In practice, they operate as a changing set of thresholds, end-user rules, licensing presumptions, foreign-direct-product provisions, and enforcement priorities. The January 2026 change illustrates this. Access to selected advanced chips may be possible, but only through a review system designed to preserve U.S. discretion.

Policy mechanism	Commercial effect	Strategic effect	Investor implication
Performance threshold	Forces product segmentation	Slows frontier diffusion	Creates inventory and roadmap risk
Case-by-case license	Allows selective revenue	Preserves policy leverage	Adds regulatory lead time
Entity restrictions	Blocks named buyers and affiliates	Targets military and strategic users	Raises customer diligence cost
Equipment controls	Limits fab expansion and upgrades	Attacks manufacturing capability	Affects ASML and global tool vendors
End-use enforcement	Restricts deployment purpose	Turns compute into governed infrastructure	Makes traceability commercially valuable

Table 3. Export controls and the political price of compute.

The policy paradox is that restrictions can work tactically and fail strategically. They can delay access to leading technology while accelerating investment in alternatives. The key question is not whether China can reproduce the U.S. stack immediately. It is whether restricted markets become large enough to support a second standard that is good enough, available and politically preferred.

RefleXio judgment. The optimal U.S. policy is unlikely to be maximum denial or unrestricted trade. It is controlled access that preserves a performance gap, increases visibility into deployment and prevents allied markets from defaulting to a Chinese stack. The investment risk lies in abrupt calibration changes, not only in the direction of policy.

6. The physical bottlenecks the market still underprices

6.1 Memory bandwidth and packaging

Model performance and inference economics increasingly depend on moving data rather than simply adding arithmetic units. HBM4, advanced substrates and packaging capacity determine how much useful compute can be assembled and delivered. A nominal accelerator order is not a deployable system until memory, packaging, networking, and power components arrive together.

6.2 Networking and scale architecture

The transition from single accelerators to rack-scale and multi-rack systems shifts value toward interconnect. NVIDIA's strategy integrates scale-up and scale-out networking. Broadcom and

Marvell benefit from custom accelerator and connectivity demand. Arista benefits from Ethernet scale-out. The competitive question is whether open networking can reduce system lock-in quickly enough to offset NVIDIA's integrated advantage.

6.3 Power, grid connection and cooling

The IEA's updated 2026 projection places data center electricity use at roughly 485 TWh in 2025 and 950 TWh by 2030 (International Energy Agency, 2025, 2026). AI-focused data center use grows faster than the total and is expected to triple over the period. Efficiency per task is improving rapidly, but demand growth and more energy-intensive agentic workloads overwhelm much of the gain.

Global data center electricity demand

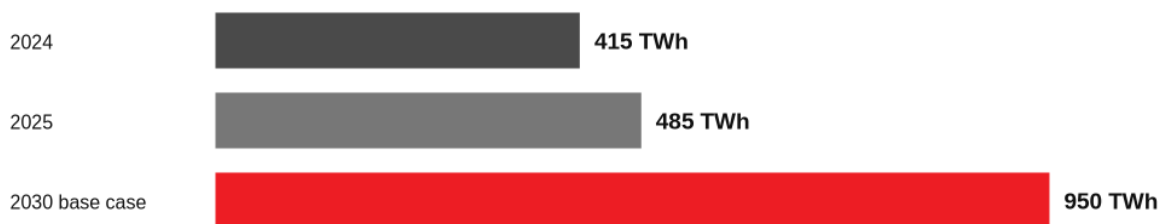


Figure 5. Global data center electricity demand.

This turns grid access into an allocation mechanism for compute. Projects can secure accelerators and still fail to obtain timely interconnection, transformers, cooling water or local permission. Europe is particularly exposed because its power markets and permitting remain fragmented. The practical unit of capacity is no longer chips ordered. It is megawatts energized.

6.4 Capital and financing

AI infrastructure increasingly combines vendor finance, hyperscaler commitments, private credit, sovereign funds, and long-term capacity contracts. This expands the market but can obscure who carries utilization risk. Financing structures should be analyzed with the same discipline applied to semiconductor inventories. A signed capacity agreement is not the same as profitable end demand.

7. Europe's position in the governed compute order

7.1 Europe's real assets

- ASML provides unique leverage through EUV lithography and a deep installed base.