AI Infrastructure: Eight Maps

A set of structural maps I built while working through the AI infrastructure value chain.

They are industry classification and a research framework, not investment advice. Company assignments reflect public information through mid-2026 and should be verified before use.

SUMMARY

The spine: one argument, five links

1. The funding structure changed. Through 2024 AI capex was paid from hyperscaler operating cash flow. In 2025-26 the marginal dollar comes from bond issuance, GPU-backed loans, convertibles, ABS/CMBS and private project finance. The payer base widened to balance sheets that are far weaker than Microsoft's.

2. So interest rates now touch the whole chain. A long-end that refuses to fall, and a Fed that may be pricing a hiking cycle, raise the cost of the marginal data center. The effect is not uniform.

3. Transmission is tiered by payer. Neoclouds and converted miners feel it first (debt-financed, pay Nvidia full price, no other business). Private developers next. Oracle after that. Microsoft/Meta/Amazon adjust guidance. Alphabet, self-funded and running its own TPUs, may buy capacity cheaply at the bottom.

4. The theme shifts from quantity to efficiency. Effective compute = chips × per-generation performance × cluster efficiency × utilization × algorithmic efficiency × interconnect efficiency. When the marginal dollar is expensive, the chain is paid to squeeze the last four terms.

5. That lands on the interconnect layer. Interconnect efficiency is where scale-up vs scale-out, CPO vs pluggables, Ethernet vs InfiniBand and NVLink vs UALink are fought. Who owns the standard decides who keeps the margin.

What would prove me wrong

The master falsifier sits downstream, not at the component makers: **any of the four hyperscalers cutting capex guidance**, or **a frontier lab raising a round below its annual infrastructure commitment**. If either happens, transmission starts at tier 1 and works down. If neither happens, the efficiency theme still holds but "rates touch the whole chain" must be re-examined. Figure 6 lists twelve indicators, earliest first.

AT A GLANCE

Maps	8
Payer tiers, by rate sensitivity	5
Falsifier indicators, by lead time	12
Live interconnect debates	5
Scaling domains	3

KEY FIGURES

Amazon 2026 capex	~$220B
Google capex guidance	$195-205B
Meta capex floor	$130-145B
Microsoft FY26 capex guide	~$190B
Hyperscaler-related issuance this year (record)	~$225B
OpenAI obligations to revenue	46:1

From Figures 5 and 6.

DATA CUT-OFF

Q2 2026 reports (August 2026)

ANALYST

Sonny Xi

Research notes

Authorship: framework, arguments and source reading are mine; the figures were drafted, translated and edited in collaboration with Claude (Anthropic). Any errors are mine.

How the eight figures fit together

Exhibit 1. What each figure maps and the question it answers

Figure	What it maps	Question it answers
1	The AI data center chain: upstream inputs, midstream data centers, downstream demand and funding, with power as the binding constraint	Who sells what to whom, and where the money comes from
2	Two ways to cut the communications sector: one Chinese sell-side team covering the whole chain vs the US split across three or four analysts	Why the same companies are valued with different languages
3	The interconnect layer in detail, arranged by the three scaling domains, with the five live debates numbered	Where the research battleground is
4	Scale-up, scale-out, scale-across: three nested layers, three physical ceilings, who owns each standard and who wins when a boundary moves	Which layer's boundary is moving
5	Downstream payers in five tiers by rate sensitivity, with funding source, cash flow, whether they pay Nvidia's margin, and the leading indicator for each	Who feels higher rates first
6	The falsifier dashboard: twelve indicators ordered by lead time, each with data source, threshold, who it hits first, current reading, and bear/bull readings	How I would know the framework is failing
7	Protocols and technologies (NVLink, UALink, Ethernet, UEC/MRC, InfiniBand, DWDM, SRv6): what each is, where it sits, why it matters, who owns it	Why open vs proprietary decides margins
8	From switch to transceiver: the signal path and its parts, then DSP pluggable vs LPO vs CPO side by side	Why the DSP is the center of debate 1

Reading conventions

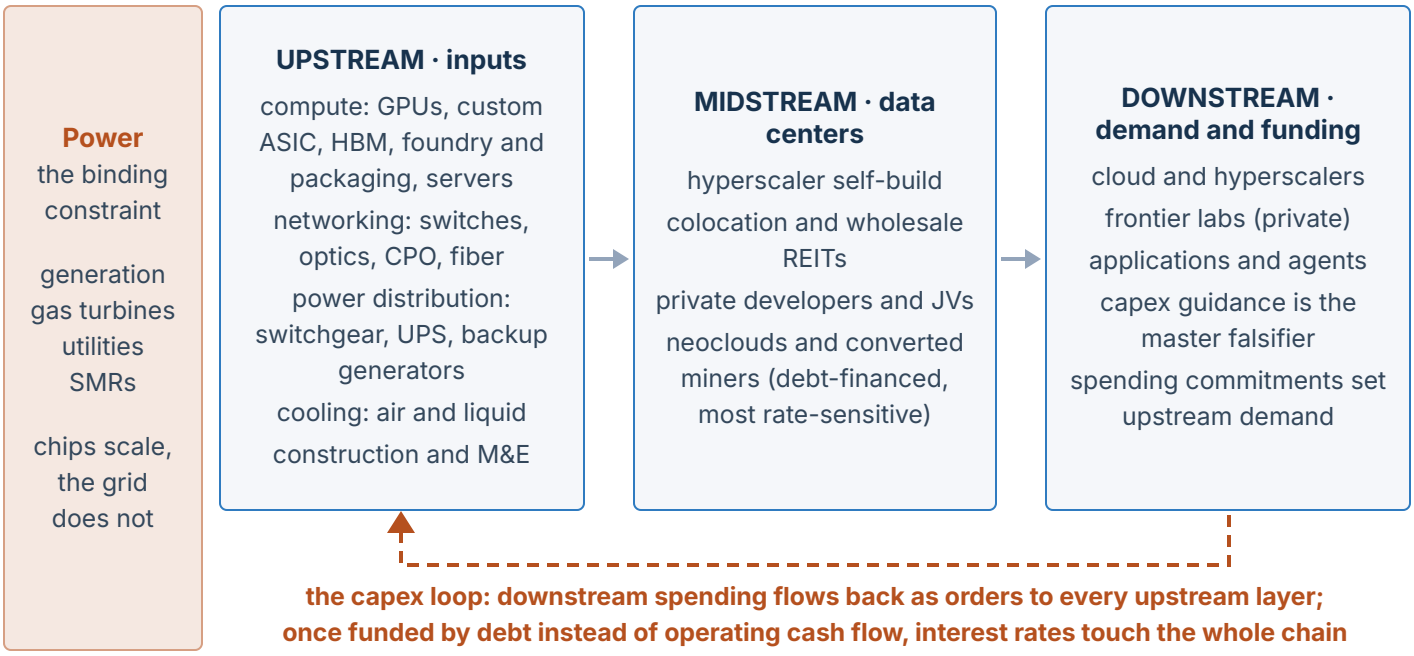
Boxes are roles, not companies. One company can occupy several boxes at once: Microsoft builds data centers (midstream), pays for them (downstream) and rents capacity from neoclouds (tier 1). A technology can also cross layers: CPO arrives first in scale-out switches and later inside the scale-up accelerator package. Where a company sits in a figure is the role being discussed, not a verdict on what the company is.

Bold navy marks a category or a key term; *gray italic* marks representative companies; plain text is explanation. **Red numbers** ①-⑤ mark the five debates and are consistent across figures. **Green "open"** means several vendors can supply; **red "proprietary"** means one vendor decides. Company names in gray are representative, not exhaustive, and carry no rating. Chinese companies are given in English transliteration (Innolight = 中际旭创, Eoptolink = 新易盛, YOFC = 长飞, TFC = 天孚).

Figure 1 · The AI data center chain

Upstream sells inputs, midstream turns inputs into compute capacity, downstream buys the capacity and pays for the whole chain. Power is drawn as a vertical bar on the left because it spans every layer: it is a constraint, not an input to one layer.

Exhibit 2. Upstream inputs → midstream data centers → downstream demand and funding



KEY TAKEAWAYS

- Upstream sells inputs, midstream turns them into capacity, downstream pays. Power runs across all three and is the constraint, not an input.
- The same company can sit in two boxes: hyperscalers build (midstream) and pay (downstream); Oracle and the neoclouds build with borrowed money.
- Capex guidance from the four hyperscalers is the single number that validates or falsifies the whole chain.

Cost reference (Tianfeng Securities, China data): IT equipment is about 90% of a new data center's build cost (servers and racks 59%, networking and security 17.5%); power is about 57% of operating cost.

Figure 1 · The AI data center chain detail

Exhibit 3. Figure 1 detail

Power and energy
the binding constraint across the chain

Generation
Constellation · Vistra · Talen

Gas turbines
GE Vernova · Siemens Energy

Regulated utilities
Dominion · Southern

Small modular reactors
NuScale · Oklo (early stage)

Chips scale. The grid does not. Grid interconnection queues, not wafer capacity, set the pace of the midstream.

Upstream · inputs

Compute and IT infrastructure
Accelerators, GPU / custom ASIC *Nvidia · AMD · Broadcom · Marvell*
HBM, high-bandwidth memory *SK Hynix · Micron · Samsung*
Foundry and advanced packaging *TSMC (CoWoS)*
AI servers *Dell · Super Micro · HPE*

Networking and interconnect (detail in Figure 3)
Switches *Arista · Cisco*; switch ASIC *Broadcom · Nvidia*
Optical transceivers *Coherent · Lumentum · Innolight · Eoptolink*
Co-packaged optics (CPO) *Nvidia · Broadcom · TSMC COUPE*
Fiber and cable *Corning · Prysmian · YOFC*

Power distribution and backup
Switchgear and UPS *Eaton · Vertiv · Schneider · ABB · Powell*
Backup generators *Caterpillar · Cummins*

Cooling and liquid cooling
Cooling equipment *Vertiv · Modine · Munters · Trane*
Liquid cooling pure plays *CoolIT · JetCool · LiquidStack (private)*

Construction and M&E
M&E contractors *EMCOR · Comfort Systems · Quanta*
General contractors *Turner · DPR · Rosendin (private)*

→

Midstream · data centers

Hyperscaler self-build
Microsoft · Alphabet · Amazon · Meta
Both the midstream builder and the downstream payer.

Colocation and wholesale REITs
Equinix · Digital Realty · Iron Mountain

Developers, mostly private
Vantage · QTS · CyrusOne · Aligned · Switch
Stargate (JV) · DigitalBridge (asset manager)

Neoclouds and converted miners
CoreWeave · Nebius · IREN · Core Scientific · Cipher · TeraWulf · Applied Digital
Debt-financed, the most rate-sensitive layer.

Where to put Google, Oracle and Microsoft is a judgment call: they build (midstream) but their primary identity is the payer (downstream). Here they stay in "self-build" with the payer role shown on the right.

→

Downstream · demand and funding

Cloud and hyperscalers
Microsoft · Alphabet · Amazon · Meta · Oracle
Alibaba · Tencent (partly decoupled)
Capex guidance is the master falsifier for the whole chain.

Frontier labs, almost all private
OpenAI · Anthropic · xAI · Mistral
Their spending commitments set demand several layers up.

Applications and agents
Enterprise software
Palantir · ServiceNow · Datadog
Devices, robotics, autonomous driving
Furthest from the network, but where monetization is finally decided.

The capex loop
Spending decisions on a handful of downstream balance sheets flow back as orders to every upstream layer. Once payers move from operating cash flow to debt and equity issuance, interest rates begin to touch the entire chain.

Figure 2 · Two ways to cut the sector

Left: how a Chinese sell-side "communications" team covers the chain, one team end to end. Right: how US sell-side splits the same companies across three or four analysts, with switch silicon and optical DSPs assigned to semiconductors. Red labels mark the strength of each group's link to AI infrastructure.

Exhibit 4. A Chinese sell-side telecom team vs the US sell-side split

CHINA · one telecom team covers the whole chain

Upstream
optical chips, components, fiber preform, base stations

▼

Midstream A · equipment and optics
telecom equipment, optical transceivers (the profitable layer)

▼

Midstream B · carriers and IDC
three carriers hold ~57% of IDC; third-party data centers

▼

Downstream
internet companies, media, consumers, enterprise networks

US · the same companies split across three or four teams

Telecom Services
AT&T, Verizon, cable
AI link: weak

Networking & Optical
Arista, Cisco, Ciena, Coherent, Lumentum
AI link: strong

Communications Infrastructure
towers, data center REITs, Lumen
AI link: medium

Semiconductors
Broadcom, Marvell: switch ASIC, DSP
AI link: strong

Index classification is a third cut:
GICS puts Arista and Cisco in IT, Coherent in Electronic Components, Broadcom in Semis

KEY TAKEAWAYS

- In China one telecom team covers the chain end to end; in the US the same companies are split across three or four analysts, and switch silicon goes to semis.
- US carriers are nearly irrelevant to AI infrastructure; data center REITs sit in real estate or Communications Infrastructure, apart from optical networking.
- Index classification (GICS) and sell-side coverage only partly overlap, so the same name is valued in different languages.

Three differences: (1) China covers the chain with one team; the US splits it and gives switch silicon to semis. (2) The US carrier group is almost unrelated to AI infrastructure. (3) US data centers usually sit in real estate or Communications Infrastructure, separate from optical networking. Boundaries shift every year; this is the common shape.

Figure 2 · Two ways to cut the sector detail

Exhibit 5. Figure 2 detail

China sell-side: one team covers the whole chain

Upstream

Optical chips, optical components, fiber preform *TFC · Accelink · YOFC*

Base stations and components

Power and liquid cooling are sometimes put here too; they are really the data center physical layer, not telecom upstream.

Midstream A: equipment and optics

Telecom equipment *Huawei · ZTE · Ruijie*

Optical transceivers *Innolight · Eoptolink · HG Genuine*

Transceivers are a layer of their own, and the most profitable one in this group.

Midstream B: carriers and IDC

Carriers *China Mobile · China Telecom · China Unicom* (the three hold roughly 57% of China's IDC capacity)

Third-party data centers *GDS · VNET · Range · Chindata*

This is why "midstream = traditional telecom companies" is true in China and not in the US.

Downstream

Internet companies, media, consumers, enterprise networks

This is the downstream of consumer communications. The downstream of AI infrastructure is a different set of payers (Figure 1).

Telecom Services AI link: weak

AT&T · Verizon · T-Mobile · Frontier

Cable operators *Comcast · Charter*

Analysts watch subscribers, ARPU, FTTH capex. US carriers do not own data centers.

Networking & Optical AI link: strong

Switches *Arista · Cisco · Juniper*

Optical transport *Ciena · Nokia/Infinera*

Optical transceivers *Coherent · Lumentum · Applied Opto*

Connectivity ICs *Credo · Astera Labs* (sometimes semis)

Several banks now brand this "Communications & AI Infrastructure".

Communications Infrastructure AI link: medium

Towers *American Tower · Crown Castle · SBA*

Data centers *Equinix · Digital Realty · DigitalBridge*

Long-haul fiber *Lumen* (repackaging backbone as data center interconnect)

Some banks put data center REITs under real estate instead.

Semiconductors AI link: strong

Switch ASIC *Broadcom Tomahawk · Nvidia Spectrum/Quantum*

Optical DSP and SerDes *Marvell · Broadcom*

Custom ASIC *Broadcom · Marvell*

The chain is cut here: a call on Broadcom or Marvell crosses into the semis team's language and valuation methods.

IT Hardware AI link: medium

Servers *Dell · Super Micro · HPE*; optics contract manufacturing *Fabrinet* sometimes lands here.

Index classification differs from sell-side teams

Since 2018 the GICS Communication Services sector is mostly Alphabet, Meta, Netflix and the carriers. Cisco, Arista and Ciena sit in IT under Communications Equipment; Coherent and Lumentum under Electronic Components; Broadcom under Semiconductors. Index buckets and sell-side coverage only partly overlap.

Figure 3 · The interconnect layer in detail

Columns are the three domains US sell-side now uses, ordered by link distance. Rows run from who pays down to upstream components. Red numbers mark the five live debates. This is the research battleground.

Exhibit 6. Arranged by the three scaling domains

	Scale-up inside the rack · 1-3 m · copper	**Scale-out** across racks · 30 m to 2 km · optics	**Scale-across** across campuses · 10-100+ km · coherent
Fabric / protocol	④ **NVLink (proprietary) vs** UALink / scale-up Ethernet (open)	② **Ethernet + RoCE (open) vs** InfiniBand (Nvidia)	**Ethernet + coherent DWDM** cross-campus training is new demand
Switching	**NVLink switch (Nvidia)** scale-up Ethernet switch (Broadcom, Arista)	① **pluggable-optics switch vs CPO switch** Arista, Cisco, Nvidia; Broadcom, Marvell silicon	**routing and optical transport** Ciena, Nokia/Infinera, Cisco/Acacia
Physical link	**copper backplane, DAC** ③ AEC pushes copper to 5-7 m (Credo, Marvell)	**pluggables 800G / 1.6T, CPO engines, LPO** Coherent, Lumentum, Innolight, Eoptolink	**coherent pluggables 400ZR / 800ZR** fiber: Corning, Prysmian, YOFC
Silicon in the link	**SerDes and retimers** Broadcom, Marvell, Credo, Astera	**optical DSP (what CPO and LPO remove)** ⑤ lasers: EML / CW / VCSEL, InP capacity	**coherent DSP** Ciena, Marvell, Acacia

who pays sits above all three: cluster architecture (Nvidia full rack vs custom accelerator + Ethernet) decides every cell

KEY TAKEAWAYS

- Three domains, ordered by link distance: copper inside the rack, optics across racks, coherent optics across campuses.
- Cluster architecture is chosen at the top (Nvidia full rack vs custom accelerator + Ethernet); every row below is derived from that choice.
- The five debates all reduce to one question: who owns the standard, and therefore who keeps the margin.

THE FIVE DEBATES

1 **CPO vs pluggables:** pace and scope of substitution (switches first, accelerators later). 2 **Ethernet vs InfiniBand:** the driver is custom-silicon penetration. 3 **AEC vs short-reach optics:** how far copper's reach can be pushed. 4 **NVLink vs UALink:** whether scale-up opens up. 5 **Laser capacity:** CW demand migration and InP oversupply, where Coherent and Lumentum ratings diverge.

Debates ②④⑤ live in the "fabric and protocol" and "silicon inside the link" rows, which is why those two rows matter most. Company assignments reflect knowledge through May 2026; industry classification only.

Figure 3 · The interconnect layer in detail detail

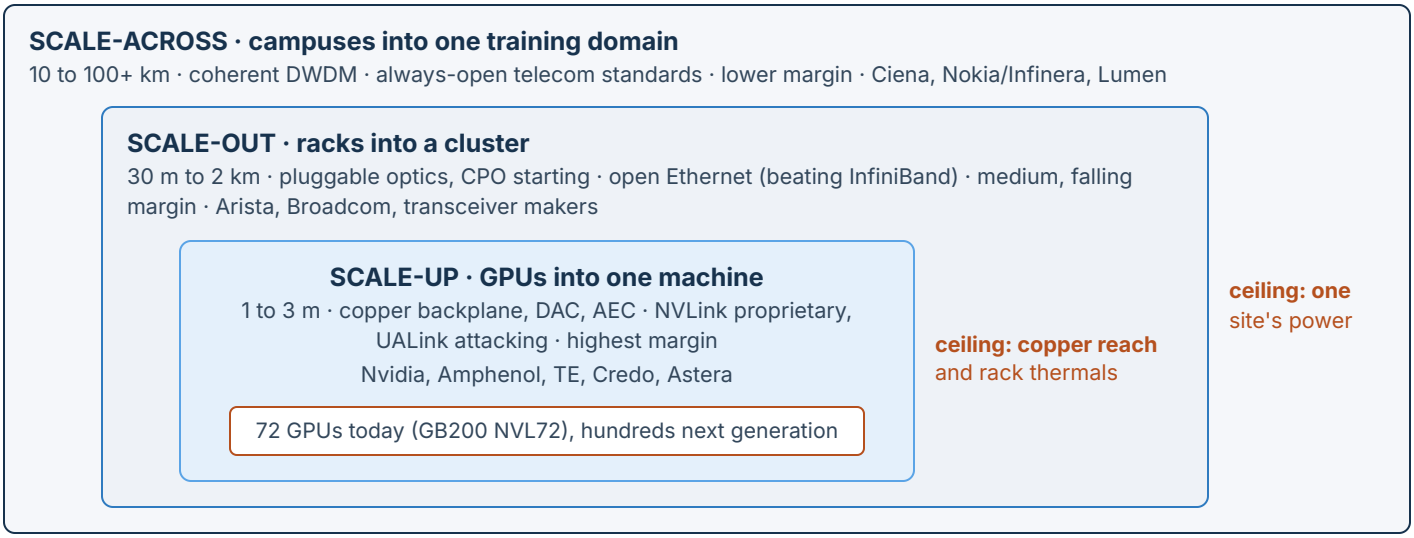
Exhibit 7. Figure 3 detail

	Scale-up inside the rack · tens to hundreds of GPUs into one machine · 1-3 m · copper, with optics entering	**Scale-out** across racks · racks into a cluster · 30 m to 2 km · the transceiver's home market	**Scale-across** across data centers · campuses into one training domain · 10 to 100+ km · coherent optics
Who pays	Hyperscalers *Microsoft · Alphabet · Amazon · Meta · Oracle*; neoclouds *CoreWeave · Nebius*; frontier labs *OpenAI · Anthropic* (private). Cluster architecture (Nvidia full rack vs in-house accelerator + Ethernet) is chosen here, and every cell below is derived from that choice.		
Fabric and protocol	4 NVLink (**Nvidia proprietary**) vs UALink / scale-up Ethernet (**Broadcom · AMD · hyperscaler consortium**). Custom-silicon clusters default to the open route.	2 Ethernet (*Arista · Broadcom · Cisco*) vs InfiniBand (*Nvidia/Mellanox*). The higher the custom-accelerator share, the larger Ethernet's share.	Ethernet + coherent DWDM. Cross-campus training synchronization is new demand; Ciena calls it scale-across and treats it as incremental TAM.
Switching	NVLink switch *Nvidia*; scale-up Ethernet switch *Broadcom Tomahawk Ultra · Arista*	Switch systems *Arista · Cisco · Nvidia Spectrum · Celestica (white box)*; switch ASIC *Broadcom Tomahawk 5/6 · Nvidia Spectrum-X · Marvell Teralynx* 1 pluggable-optics switch vs CPO switch (*Nvidia Quantum-X/Spectrum-X Photonics · Broadcom Bailly*)	Routing and optical transport *Ciena · Nokia/Infinera · Cisco/Acacia*; DCI platforms *Ciena WaveLogic*
Physical link	Copper backplane and DAC *Amphenol · TE* 3 AEC, active electrical cable (*Credo · Marvell*) pushes copper to 5-7 m and squeezes short-reach optics PCIe/CXL connectivity ICs *Astera Labs* Next: CPO inside the accelerator package (*Marvell · TSMC COUPE gen 3*)	Pluggable transceivers 800G / 1.6T *Coherent · Lumentum · Innolight · Eoptolink · Fabrinet (EMS) · AAOI* CPO optical engine *Nvidia · Broadcom · Marvell* LPO / LRO linear pluggable / linear receive optics: in between, drops the DSP but keeps the pluggable form	Coherent pluggables 400ZR / 800ZR *Ciena · Cisco/Acacia · Marvell · Lumentum* Fiber and cable *Corning · Prysmian · YOFC*
Silicon inside the link	SerDes IP and retimers *Broadcom · Marvell · Credo · Astera*	Optical DSP *Marvell · Broadcom*: the chip both CPO and LPO remove Silicon photonics PIC / EIC *Nvidia · Broadcom · Coherent · Innolight (in-house)* 5 Lasers: EML / CW / VCSEL *Lumentum · Coherent (captive) · Broadcom (captive) · AOI · Yuanjie*; CPO's external laser source makes the CW laser a key merchant component; InP capacity expansion raises an oversupply debate	Coherent DSP *Ciena (in-house) · Marvell · Cisco/Acacia* Narrow-linewidth lasers and modulators *Lumentum · Coherent*
Components and foundry	SiPh foundry *TSMC (COUPE) · GlobalFoundries · Tower*; optical engines and components *TFC (Nvidia CPO partner) · Accelink · Fabrinet*; fiber array units, PM fiber, MPO/MT connectors *HYC · Corning · Senko · Sumitomo · Amphenol*; ABF substrate *Ibiden · Unimicron*; OSAT advanced packaging *SPIL · ASE*		
Adjacent constraints	Rack power density and liquid cooling (CPO switches themselves need liquid cooling); grid interconnection queues; TSMC CoWoS and SiPh capacity (the hard ceiling on CPO volume); export controls (geopolitical risk for Chinese transceiver makers shipping to the US)		

Figure 4 · Scale-up, scale-out, scale-across

The three layers are nested, not sequential: a large cluster has all three from day one. Each layer exists because the one below hit a physical limit. The buyer is the same hyperscaler in all three; what changes is the seller, who owns the standard, and the margin.

Exhibit 8. Three nested layers, three physical ceilings



margin falls as you move outward: proprietary inside the rack, open across racks, mature telecom across campuses

KEY TAKEAWAYS

- The three layers are nested, not sequential; each exists because the one below hit a physical ceiling (copper reach, then site power).
- Margin falls as you move outward: proprietary NVLink inside the rack, open Ethernet across racks, mature telecom DWDM across campuses.
- The question that matters for investment is which boundary is moving. Optics per GPU and scale-out bandwidth per GPU are the two things to measure.

Not the IT textbook definition

In traditional IT, scale-up means adding CPUs and memory to one server (vertical), scale-out means adding servers (horizontal). AI networking borrowed the words and changed the meaning: scale-up here means wiring many GPUs into one machine, which itself depends on interconnect; scale-out means wiring those machines into a cluster. The textbook says scale-up is cheap and scale-out expensive; in AI it is reversed. Scale-up interconnect is the most expensive, highest-margin cell.

The useful question is which boundary is moving

Not "which layer happens first" but "which layer's boundary is shifting". When a boundary moves, the business allocation between layers changes. Two observable metrics: **optics per GPU** (sell-side assumptions range from 2 to 4; its direction is the speed at which scale-up eats scale-out) and **scale-out bandwidth per GPU** (can fewer ports be offset by higher price per port). Chinese transceiver makers' capacity plans and next-generation inventory builds are the earliest place to read both.

Effective compute = chips × per-generation performance × cluster efficiency × utilization × algorithmic efficiency × interconnect efficiency. Chip scarcity is not compute scarcity. The first signal of a narrowing supply gap is a fall in compute rental prices, which leads both capex guidance and RPO.

Figure 4 · Scale-up, scale-out, scale-across detail

Exhibit 9. Figure 4 detail

Scale-up · inside the rack

Connect tens to hundreds of GPUs into "one machine": shared memory, every step exchanges intermediate results. Needs bandwidth density and the lowest latency.

Distance 1-3 m · **Medium** copper (backplane, DAC, AEC), optics entering · **Unit** 72 GPUs (GB200 NVL72), hundreds next generation

Ceiling: copper reach and rack thermals → can only span a rack

Scale-out · across racks, inside a cluster

Connect hundreds or thousands of scale-up units into a cluster. Exchanges gradients and parameter updates, an order of magnitude less frequent, tolerates slightly higher latency. Needs scale.

Distance 30 m to 2 km · **Medium** optics (pluggables, CPO starting to replace) · **Scale** tens to hundreds of thousands of GPUs

Ceiling: one building's or one campus's power → can only span a campus

Scale-across · across data centers

Connect several campuses into one training domain when one building's power is not enough. Telecom backbone technology, re-lit by AI.

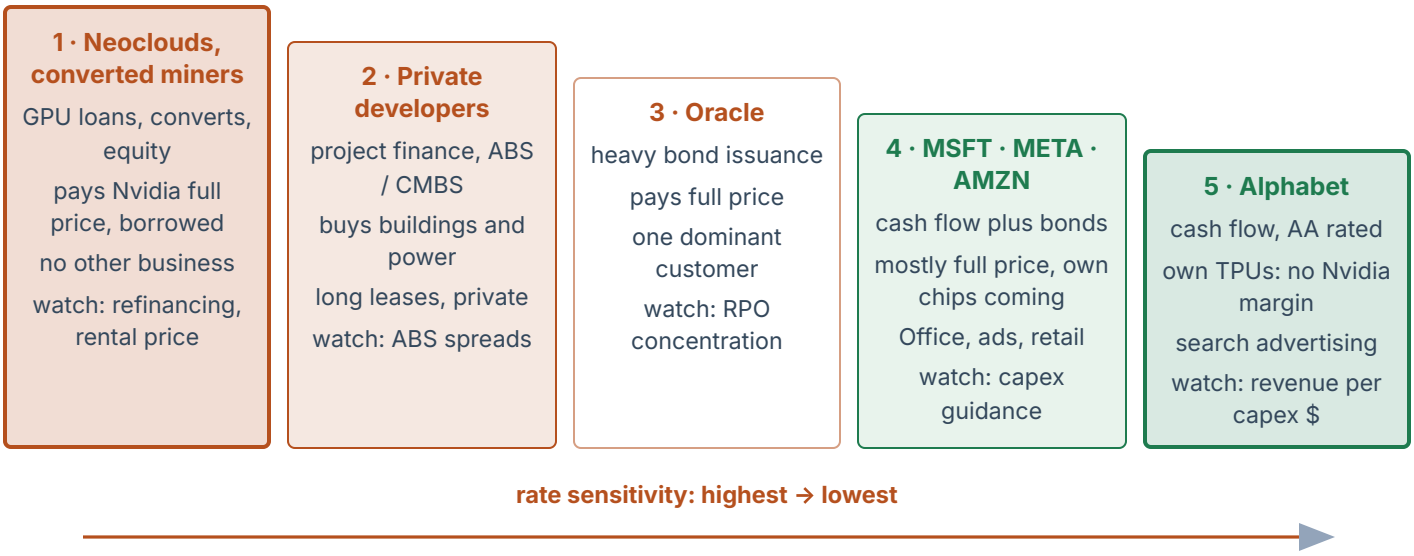
Distance 10 to 100+ km · **Medium** coherent optics (DWDM, 400ZR/800ZR) · **Scale** multiple campuses

	Scale-up	Scale-out	Scale-across
Who owns the standard	**Proprietary**: NVLink. The open camp, UALink (*AMD · Broadcom · hyperscalers*) and scale-up Ethernet, is attacking.	**Open**: Ethernet is winning new deployments over InfiniBand (Arista counts 100+ AI fabric customers), and Nvidia itself now sells Ethernet via Spectrum-X. Installed InfiniBand clusters remain.	**Always open**: DWDM and coherent optics are decades-old telecom standards.
Switching	NVLink switch *(bundled with GB200 racks)*; scale-up Ethernet switch *Broadcom Tomahawk Ultra*	Ethernet / IB switches *Arista · Cisco · Nvidia Spectrum-X / Quantum · Celestica white box on Broadcom silicon*; switch ASIC *Broadcom Tomahawk · Nvidia · Marvell*	Routing and optical transport *Ciena · Nokia/Infinera · Cisco/Acacia · Arista routing*
Physical link	Copper backplane, DAC *Amphenol · TE*; AEC *Credo · Marvell*; PCIe/CXL *Astera*; next, CPO in the accelerator package	Pluggable optics 800G / 1.6T *Coherent · Lumentum · Innolight · Eoptolink · Fabrinet*; CPO switches *Nvidia · Broadcom*; LPO / LRO	Coherent pluggables 400ZR / 800ZR *Ciena · Marvell · Lumentum*; long-haul fiber *Lumen · Corning · YOFC*
Margin and pricing power	**Highest.** Proprietary standard: buy the GPU and you must buy the interconnect. The customer has no option.	**Medium and falling.** Open standards let customers multi-source. Share gains do not equal profit gains.	**Lower.** A mature telecom equipment market; AI adds incremental demand on top.
Debates	4 NVLink vs UALink: does scale-up open? 3 AEC vs short-reach optics: how far does copper go?	2 Ethernet vs InfiniBand: driven by custom-silicon penetration. 1 CPO vs pluggables: pace of substitution. 5 Laser capacity.	Is scale-across a separate market? Arista guides a $15-20B TAM by 2030; Ciena calls it incremental.
Who wins when the boundary moves	Larger scale-up units → Nvidia (more NVLink), copper and AEC vendors; later CPO engines (but defined by Nvidia)	Larger scale-up → fewer scale-out ports per GPU → pressure on transceiver makers and Arista. Offsets: total GPU growth, per-port bandwidth upgrades.	Tighter power constraints → more cross-campus training → Ciena, Lumen, coherent optics benefit

Figure 5 · Downstream payers in five tiers

The same capex has very different outcomes depending on the payer's balance sheet. Tiers are ordered by "does it buy Nvidia at full price with borrowed money, and does it have another business to fall back on". Rate transmission runs from tier 1 down; margin compression appears in the same order.

Exhibit 10. Who is most rate-sensitive, and who is paying Nvidia's margin



rate transmission runs left to right: tier 1 breaks first, its suppliers next, tier 4 adjusts guidance, tier 5 may buy capacity cheaply at the bottom

Source: Q2 2026 reports and public filings as of August 2026.

KEY TAKEAWAYS

- Same capex, different balance sheets: ordered by whether the payer borrows to buy Nvidia at full price and whether it has another business.
- Rate transmission runs tier 1 to tier 5: neoclouds break first, hyperscalers adjust guidance, Alphabet may buy capacity cheaply at the bottom.
- Margin compression is a gradient, not a cloud-wide disease: it shows up below gross margin, in depreciation and interest, in the same order.

Order of rate transmission: tier 1 breaks first → its suppliers (Nvidia's marginal demand, used GPU prices) → tier 4 adjusts capex guidance → tier 5 may buy capacity cheaply at the bottom

Figure 5 · Downstream payers in five tiers continued

A margin gradient, not a "cloud disease"

CoreWeave's 67-74% gross margin is fine. The damage is two lines below: depreciation rising with the GPU base (opex $1.1B → $3.7B → $5.3B) and interest rising with debt. Hyperscalers have the same depreciation ahead of them (Alphabet: infrastructure investment will keep pressing the P&L through higher depreciation and energy cost), but they have no interest line, and Google does not pay Nvidia's margin. So compression arrives in order: borrowed + full price + no in-house silicon first; cash flow + in-house silicon last.

The Nvidia complex, consolidated

The same GPUs: Nvidia books them as revenue at 75% gross margin; tier 1 books them as six years of depreciation plus interest, and rent currently does not cover it. Nvidia's E is real; its quality is set by buyer returns. The weakest 10-20% of buyer balance sheets are the marginal demand that shrinks first, and when it shrinks it takes used GPU prices, and everyone's depreciation assumptions, with it. Three ratios: interest / adj. EBITDA; contracted revenue / debt service; GPU depreciation life / contract length.

Three questions for any RPO

RPO only includes non-cancellable contracts: legally hard, economically not necessarily. The counterparty may not be able to pay (frontier labs are private and raise to spend); contracts have soft spots (minimum commitment plus usage, staged start, renegotiation); the 2001 telecom IRU contracts are the precedent. So ask: who is the customer, where does the customer's money come from, firm commitment or usage-based. RPO growth lagging capex growth is the precursor to payers changing their guidance.

Portfolio implication: the cleanest expression of "AI demand is real, the funding structure is fragile" is long the strongest balance sheets and short the most levered, rather than a blanket view on the theme. Data from Q2 2026 reports and public filings as of August 2026; verify before citing. Structural mapping only.

Figure 5 · Downstream payers in five tiers detail

Exhibit 11. Figure 5 detail

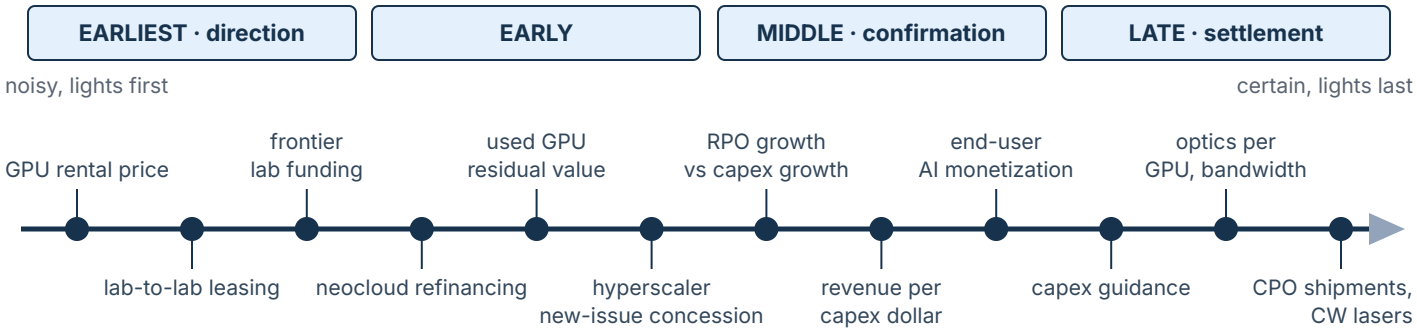
Tier	Funding source	Interest and cash flow	Pays Nvidia's margin?	Other cash cows	Rate sensitivity	Evidence and leading indicator
1 · Neoclouds and converted miners *CoreWeave · Nebius · IREN · Core Scientific · Cipher · TeraWulf · Applied Digital*	GPU-backed loans, convertibles, equity raises; Nvidia equity stakes and backstops (CoreWeave: $6.3B capacity backstop to 2032)	Interest about a quarter of revenue; current ratio 0.46; capex 3.7× revenue; negative FCF is structural, no spend means no revenue	Full price, with borrowed money	None	**Highest**	CoreWeave Q1 interest $536M = 25.8% of revenue, 46.3% of adj. EBITDA; $4.2B principal due in H2. Operating income +$324M (2024) → -$46M (2025) → -$231M TTM: revenue up 2.7×, profit turned negative. **Watch:** refinancing spreads, GPU rental price, used GPU residual value
2 · Private developers *Vantage · QTS · CyrusOne · Aligned · Stack · Stargate (JV)*	Project finance, ABS / CMBS, private equity; tenants are hyperscalers so credit is one notch better	High leverage but long-lease cash flow; private, so stress is invisible	Not directly (they buy the building and the power; chips belong to the tenant)	None, but long leases	**High**	Hidden debt carried by the big five through project-finance structures estimated at $1.65T; only part of AI's real cost sits in public bond markets. **Watch:** data center ABS spreads, pace of new project-finance closings
3 · Oracle *the listed company most like a neocloud*	Heavy bond issuance to fulfil the OpenAI contract; operating cash flow does not cover capex	FCF turned negative; RPO exploded but highly concentrated	Full price	Database and applications, but small relative to capex	**Medium-high**	RPO quality depends on OpenAI's ability to raise, and OpenAI is private. **Watch:** single-customer share of RPO, bond spreads vs the tech sector
4 · Microsoft · Meta · Amazon *hyperscalers buying at full price*	Mostly operating cash flow, gap funded with bonds; also rent part of their capacity from tier 1 (Microsoft ~$60B, Meta up to $62B in neocloud contracts), turning capex into opex	Microsoft FY26 capex guide ~$190B, close to 95% of operating cash flow; FCF squeezed to near zero	Mostly full price; in-house silicon (Maia, Trainium) on the way	Yes: Office, ads, retail	**Medium**	Microsoft Cloud gross margin down to 67%; AWS segment margin expanded 520 bp. Amazon 2026 capex raised to ~$220B; Meta floor lifted to $130-145B. **Watch:** capex guidance, issuance size and new-issue concession
5 · Alphabet *vertically integrated*	Mostly operating cash flow, gap funded with bonds and equity; buybacks and dividends continue	Q2 FCF -$5.9B, but cash-rich and AA rated; negative FCF is a choice, not a structure	Mostly no: TPUs are in-house, the supplier margin stays inside	Yes: search advertising	**Lowest**	Google Cloud Q2 revenue $24.8B, +82% YoY, operating margin 35.6%; backlog $514B vs $106B a year earlier. The market rewards Google and punishes the other three for spending the same money. **Watch:** incremental cloud revenue and backlog per quarter of incremental capex

Source: Q2 2026 reports and public filings as of August 2026.

Figure 6 · The falsifier dashboard

The spine is "funding structure changed → rates touch the chain → tiered transmission by payer → theme shifts to efficiency → lands on the interconnect layer". Every link has an observable. Indicators near the top light up earliest but are noisiest; those near the bottom are most certain but latest. Bear and bull readings are shown side by side.

Exhibit 12. Twelve indicators ordered by lead time



Master falsifier, set downstream:
a hyperscaler cuts capex guidance, or a frontier lab raises less than its annual infrastructure commitment

Reading rule:
one light is not a signal; two in the same group pointing the same way is
2000 ran in this order: bandwidth price, carrier financing, then equipment orders

KEY TAKEAWAYS

- Twelve observables, earliest first. The early ones are noisy, the late ones are certain. One light is not a signal; two in the same group pointing the same way is.
- The master falsifier sits downstream: a hyperscaler cutting capex guidance, or a frontier lab raising less than its annual infrastructure commitment.
- The bull case is written to the same standard, so the dashboard can be read either way.

The master falsifier

Set downstream, not at component makers: any of the four hyperscalers cutting capex guidance, or a frontier lab raising a round below its annual infrastructure commitment. Either one, and transmission begins at tier 1. Neither, and the efficiency theme still holds but "rates touch the whole chain" must be reassessed.

The bull case, written to the same standard

Efficiency gains raise the return on capex (effective compute = chips × performance × cluster efficiency × utilization × algorithmic efficiency × interconnect efficiency); inference share rising to 73% by 2027 means monetization has only started; Google's 36% cloud margin is evidence investment is already paying back; "buildings and power early, chips late" compresses the window where depreciation arrives before revenue; power bottlenecks may extend tier 1's window.

Rules for reading the lights

One light is not a signal; two or more in the same group pointing the same way is. The earliest three are direction, the middle three are confirmation, the last three are settlement. 2000 ran in the same order: bandwidth prices fell first (DWDM pushed effective capacity past demand), carrier financing broke next, equipment orders vanished last, and the fiber itself was used for twenty years.

Figure 6 · The falsifier dashboard detail

Exhibit 13. Figure 6 detail: twelve indicators ordered by lead time

Indicator	Lead	Data source	Threshold	First hit	Current reading (Aug 2026)	Bear read / bull read
GPU rental price	Earliest	H100 / B200 hourly rates on public compute markets and neocloud price lists	Rent falls below the unit cost line of depreciation + interest; tier 1 unit economics turn negative	Tier 1	H100 rents down sharply over two years; CoreWeave operating income +$324M (2024) → -$231M TTM	**Bear:** effective supply has caught up, the stage before glut. **Bull:** only old generations are discounted; new generations still sold out
Lab-to-lab leasing	Earliest	Number and terms of frontier labs renting their own clusters to rivals	Leasing goes from one-off to routine; leases carry short exit clauses	Tier 1, labs	Anthropic rents xAI's Colossus 1 at $1.25B/month to 2029 with 90-day exit both ways; Colossus 1 utilization previously ~11%; SpaceX also rents to Google at $920M/month	**Bear:** capacity built by those who cannot fill it; mismatch precedes glut. **Bull:** buyers still pay extreme prices for capacity available within a month, so the gap is real
Frontier lab funding	Earliest	Size, valuation and terms of OpenAI, Anthropic, xAI rounds; IPO progress	A round smaller than the annual infrastructure commitment; valuation down round; IPO delayed	RPO at every tier	OpenAI infrastructure obligations above $1.15T vs ~$25B annualized revenue (end-February 2026), about 46:1; $122B raised in March at $852B valuation	**Bear:** the whole chain's RPO shares one private denominator. **Bull:** rounds still oversubscribed; ability to pay not yet in question
Neocloud refinancing	Early	Convertible coupons and conversion premia, GPU-loan spreads, whether maturing principal rolls	Interest / adj. EBITDA above 50%; principal cannot roll on original terms	Tier 1	CoreWeave Q1 interest 25.8% of revenue, 46.3% of adj. EBITDA; $4.2B due in H2; Nebius just issued $5.75B of converts	**Bear:** a company that lives on refinancing has a liquidity event the moment terms change. **Bull:** Nvidia backstops and hyperscaler long-term contracts keep ratings investment grade
Used GPU residual value	Early	Bulk GPU disposal prices; miner and neocloud asset sales	Residuals below the depreciation assumption in everyone's model	Tier 1 → Nvidia valuation	No bulk disposal yet; Microsoft stretched data center depreciation life to 25 years	**Bear:** once it appears, every asset-life assumption in the chain is recomputed. **Bull:** old cards still have economic life in inference
Hyperscaler new-issue concession	Early	Concession and book coverage on large AI-related bond deals; hyperscaler spread curve vs tech	Concession keeps widening; coverage falls; financing migrates to private structures faster	Tier 4 capital allocation	Record ~$225B of hyperscaler-related issuance this year; Amazon's $25B deal widened its 30-year by 20 bp; S&P warned of market fatigue	**Bear:** the next dollar costs more; marginal data centers go first. **Bull:** this is supply pricing, not credit pricing; leverage still healthy, AA rated

Figure 6 · The falsifier dashboard detail

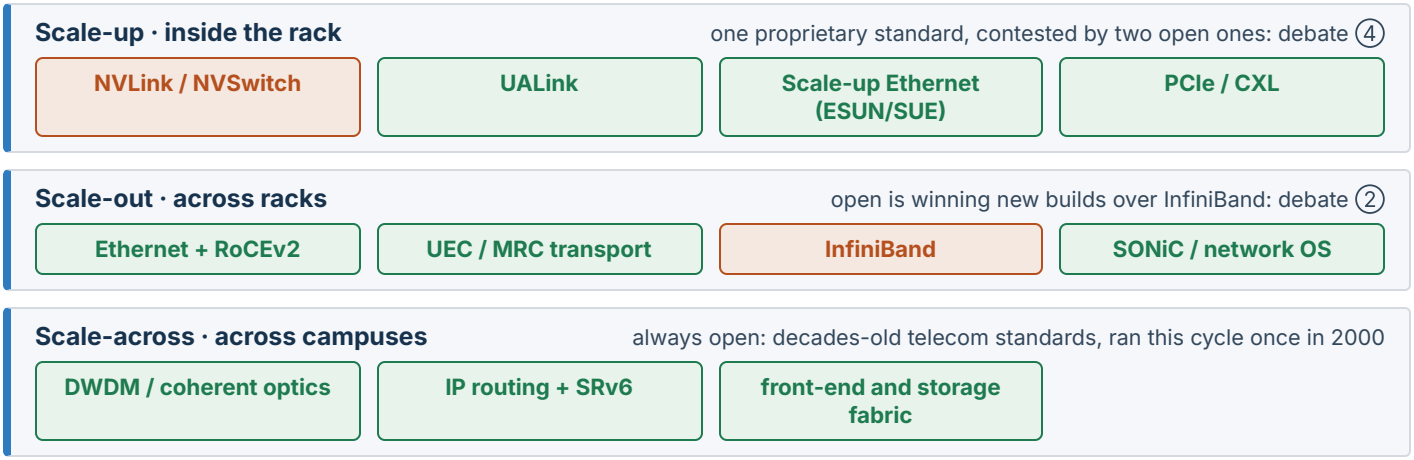
Exhibit 13. Figure 6 detail: twelve indicators ordered by lead time (continued)

Indicator	Lead	Data source	Threshold	First hit	Current reading (Aug 2026)	Bear read / bull read
RPO growth vs capex growth	Middle	Ratio of backlog growth to capex growth each quarter; single-customer share of RPO	RPO growth lags capex growth; one customer above 40%	Tiers 3, 4	Google Cloud backlog $514B (vs $106B); Microsoft backlog $678B, OpenAI ~45%	**Bear:** capacity running ahead of demand is the precursor to payers changing guidance. **Bull:** RPO still far above capex; supply is still the bottleneck
Incremental revenue per capex dollar	Middle	Incremental cloud revenue and backlog per quarter of incremental capex; cloud segment margins	The ratio turns down; margins fall more than 10 points when depreciation arrives	Tiers 4, 5	Google Cloud margin 21% → 36%; AWS +520 bp; Microsoft Cloud gross margin down to 67%; Alphabet Q2 FCF -$5.9B	**Bear:** margin expansion only because depreciation has not arrived. **Bull:** direct evidence that investment is already paying back
End-user AI monetization	Middle	Frontier lab revenue, enterprise AI spend, subscription retention, inference volume growth	Lab revenue growth below their compute commitment growth; enterprise adoption stalls at pilot	Root of the whole chain	OpenAI ~$25B annualized; Anthropic above $30B annualized in April; enterprise adoption moving from pilot to workflow	**Bear:** a 46:1 obligation-to-revenue ratio is unsustainable. **Bull:** revenue growth still triple digits; inference has only begun
Capex guidance	Late	Annual guidance from the four hyperscalers; share of spend "decided late" (chips)	Any one cuts guidance, or defers commitments on short-lived assets	Whole chain	Amazon ~$220B, Google $195-205B, Meta $130-145B, all raised; all four run the same "lock buildings and power early, decide chips late" playbook	**Bear:** this is the master falsifier; a cut reverses the framework. **Bull:** guidance still rising, and risk is already being managed structurally
Optics per GPU, bandwidth per GPU	Late	Cluster architecture announcements; Chinese transceiver makers' capacity plans and inventory; 800G / 1.6T shipment mix	A larger scale-up domain cuts optics per GPU and higher unit price does not offset	Interconnect: transceiver makers, Arista	Sell-side assumptions range 2 to 4; Arista raised guidance three times; 1.6T and liquid-cooled switches ramp in 2027	**Bear:** scale-up eats scale-out and margin accrues to Nvidia's definitions. **Bull:** volume and price growth offset fewer ports
CPO shipments, CW laser orders	Late	Nvidia Quantum-X / Spectrum-X Photonics shipments; TSMC COUPE volume; laser makers' order mix	CPO switches reach volume and customer procurement language shifts	Interconnect: module makers, DSP, lasers	Timelines still on the Q1 2025 framing (Quantum-X 2H25, Spectrum-X 2H26, COUPE gen 2 in 2026); delivery needs verifying	**Bear:** value migrates from module makers to chip makers and foundries. **Bull:** substitution slower than expected; system vendors keep benefiting in the pluggable era

Figure 7 · Interconnect protocols and technologies

Arranged by scaling domain. **Open** means several vendors can build it; **proprietary** means one vendor decides. Every layer's core debate is proprietary vs open, because who owns the standard decides who keeps the margin. Compiled from vendor technical blogs and public standards material, with the product marketing removed.

Exhibit 14. NVLink, UALink, Ethernet, InfiniBand, DWDM: where each sits and why it matters



red = proprietary, one vendor decides · green = open, several can build · who owns the standard decides who keeps the margin

Source: vendor technical blogs and public standards material.

KEY TAKEAWAYS

- Every protocol here is either proprietary (one vendor decides) or open (several can build). That split decides where the margin lands.
- Scale-out already went open (Ethernet over InfiniBand). Scale-up is being contested now (UALink and scale-up Ethernet against NVLink). Scale-across was always open.
- DWDM is the one technology that has run this cycle before, in 2000, and blew up the supply side. Worth remembering when looking at scale-across.

Why Ethernet becomes the common base

Two of three layers are open; the third is contested. The value of open is not better technology but three things: multi-sourcing without lock-in, no management ceiling above hundreds of thousands of GPUs, and a ready engineer pool and toolchain. Ethernet was originally unsuited to AI (lossy, flow-balanced); the RoCEv2 → UEC → MRC line is closing InfiniBand's advantages one by one. What remains is speed: standards change yearly, switch silicon every eighteen months.

Proprietary vs open is who keeps the margin

NVLink is proprietary, so Nvidia keeps the networking margin too; Ethernet is open, so scale-out margin is split among Broadcom, Arista and transceiver makers, and share gains do not equal profit gains. Debates ② and ④ are the same question replicated at two layers: scale-out has answered, scale-up is answering now. Whether UALink or ESUN wins matters less than whether either succeeds; if one does, Nvidia's scale-up pricing power begins to fade.

Why DWDM deserves its own memory

It is the only technology of the three layers that has already run once, in 2000. Carriers laid fiber on "traffic doubles every N months"; DWDM multiplied each fiber's capacity tens of times; effective supply overshot demand; bandwidth prices collapsed; a cohort of fiber builders failed; the fiber itself was used for twenty years. It is back because power constraints push training across campuses. When looking at Ciena, count the scale-across increment, and remember how this technology blew up the supply side last time.

Figure 7 · Interconnect protocols and technologies detail

Exhibit 15. Figure 7 detail: what each is, where it sits, why it matters, who owns it

Protocol	What it is	Primary use	Why it matters	Who owns it	Providers
Scale-up inside the rack · 1-3 m · mostly copper					
NVLink / NVSwitch GPU-to-GPU bus and switch	Nvidia's GPU interconnect: 72 GPUs in a rack (hundreds next generation) read and write each other's memory directly, like one large GPU. NVSwitch is the switch chip that makes NVLink all-to-all.	The highest-frequency data exchange between GPUs in training and inference (tensor and expert parallelism)	Bandwidth an order of magnitude above PCIe, lowest latency. Buy Nvidia GPUs and you must use it: the hardest ring in Nvidia's moat. The larger the scale-up domain, the more NVLink per rack and the fewer scale-out ports per GPU.	**Proprietary: Nvidia**	Nvidia only. Copper backplane and cables *Amphenol · TE*; once optical, CPO engines may enter (still defined by Nvidia)
UALink Ultra Accelerator Link	Open accelerator interconnect standard, consortium formed 2024 by AMD, Broadcom, Intel, Google, Microsoft, Meta, AWS and others, aiming to be the open substitute for NVLink in scale-up.	In-rack interconnect for non-Nvidia accelerators (AMD MI, hyperscaler in-house chips)	Without an open scale-up interconnect, custom-silicon clusters stay small. Whether UALink succeeds decides whether scale-up opens like scale-out did, and whether Broadcom and Arista get into the rack. This is debate ④.	**Open: UALink Consortium**	Switch silicon *Broadcom · Marvell · Astera Labs*; accelerators *AMD · Google TPU · AWS Trainium · Microsoft Maia*; copper and AEC *Credo · Amphenol*
Scale-up Ethernet ESUN / SUE	Using Ethernet for in-rack interconnect: Broadcom's SUE (Scale-Up Ethernet) and OCP's ESUN (Ethernet for Scale-Up Networking). The idea is to reuse Ethernet silicon and ecosystem.	Same as UALink; the other branch of the open scale-up route	Competes with UALink for the "open scale-up" slot. If Ethernet pushes from scale-out into scale-up, all three layers are Ethernet and switch-silicon and system vendors' territory extends into the rack.	**Open: OCP, Broadcom-led**	*Broadcom Tomahawk Ultra · Arista*; hyperscalers (Meta, OpenAI in ESUN)
PCIe / CXL general bus and coherent memory	PCIe is the general bus between CPU, GPU and NIC; CXL adds memory coherence on top so accelerators can share memory pools.	GPU to CPU, NIC and storage; memory expansion and pooling	Not the GPU-to-GPU backbone, but every GPU reaches the NIC and host through it. The PCIe generation (Gen5 → Gen6) caps how fast the scale-out NIC can run. Signal integrity problems created the retimer category.	**Open: PCI-SIG / CXL Consortium**	Retimers and connectivity ICs *Astera Labs · Broadcom · Marvell*; CPUs *Intel · AMD*

Source: vendor technical blogs and public standards material.

Figure 7 · Interconnect protocols and technologies detail

Exhibit 15. Figure 7 detail: what each is, where it sits, why it matters, who owns it (continued)

Protocol	What it is	Primary use	Why it matters	Who owns it	Providers
Scale-out across racks · 30 m to 2 km · optics					
Ethernet + RoCEv2 Ethernet with RDMA over Converged Ethernet	Forty years of general networking, plus RoCEv2 to let NICs read and write remote memory without the CPU (RDMA). Lossless Ethernet via PFC, ECN congestion control.	Connect hundreds or thousands of scale-up units into training and inference clusters; also the front-end and storage network base	Open, cheap, multi-vendor, large engineer pool. Hyperscalers push it to avoid single-vendor lock-in. Proven at hundreds of thousands of GPUs across multiple data centers. Every cell Ethernet wins is another cell of business for Broadcom and Arista. This is debate ②.	**Open: IEEE / IBTA (RoCE)**	Switch silicon *Broadcom Tomahawk · Marvell Teralynx · Nvidia Spectrum*; systems *Arista · Cisco · Nokia · Juniper/HPE · Celestica white box*; NICs *Broadcom · Nvidia ConnectX · AMD/Pensando*; optics *Coherent · Lumentum · Innolight · Eoptolink*
UEC / MRC Ultra Ethernet Transport · Multipath Reliable Connection	Next-generation Ethernet transport layer for AI traffic. UEC is set by the Ultra Ethernet Consortium; MRC is developed under OCP as an evolution of RoCEv2: packet-level rather than flow-level load balancing, multi-plane, deterministic paths, faster congestion feedback.	Solve AI traffic's synchronized, bursty, jitter-sensitive problems; stop one flow blocking thousands of GPUs	This is how Ethernet closes the gap with InfiniBand. Meta data shows network overhead can be up to 60% of iteration time in large training runs, so transport efficiency equals compute utilization. New standards require new switch silicon; an 18-month chip cadence that cannot keep up with the standard becomes a competitive issue.	**Open: UEC / OCP**	Silicon *Broadcom · Nvidia · Marvell · AMD*; systems *Arista (MRC cited in Q2) · Cisco · Nokia*; standards participants *Meta · Microsoft · OpenAI*
InfiniBand HPC-native network	Low-latency lossless network from supercomputing, native RDMA. Became Nvidia's scale-out offering after the 2019 Mellanox acquisition.	Mainstream scale-out for early AI clusters, especially Nvidia-delivered full clusters	Led on performance, but a proprietary ecosystem, single vendor, and management complexity above tens of thousands of GPUs. Share in new deployments is being eroded by Ethernet; Nvidia launched its own Ethernet line (Spectrum-X) as a hedge. Its lifespan sets the margin of Nvidia's networking business.	**Proprietary in practice: Nvidia (IBTA on paper)**	*Nvidia Quantum switches · ConnectX NICs*; optics suppliers as above
SONiC / network OS	The software that runs on the switch. SONiC is Microsoft's open-source network OS, the base of white-box switches; system vendors have their own.	How switches are managed, automated and monitored	Software is the system vendor's moat against white box (Arista's EOS), and also the hyperscaler's tool to escape system vendors (SONiC + Broadcom silicon + white box). Who owns the software decides whether the system layer keeps any margin.	**Open: SONiC (Linux Foundation) / vendor OS**	SONiC *Microsoft-led · Celestica and other white box*; in-house *Arista EOS · Cisco NX-OS · Nokia SR Linux*

Source: vendor technical blogs and public standards material.

Figure 7 · Interconnect protocols and technologies detail

Exhibit 15. Figure 7 detail: what each is, where it sits, why it matters, who owns it (continued)

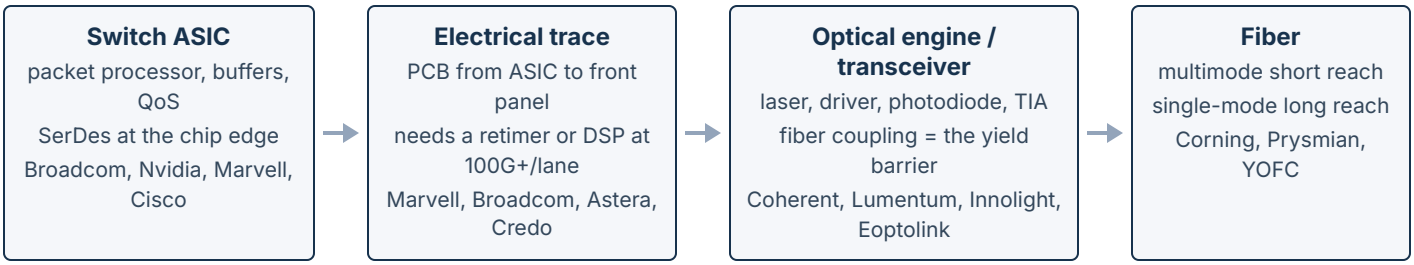
Protocol	What it is	Primary use	Why it matters	Who owns it	Providers
Scale-across across campuses · 10 to 100+ km · coherent optics					
DWDM / coherent optics dense wavelength division multiplexing	Tens to hundreds of wavelengths in one fiber, each coherently modulated (phase and amplitude) to push a single wavelength past 400G / 800G, thousands of km without regeneration.	Data center interconnect (DCI), metro and long-haul backbone; now used to join campuses into one training domain	When one building's power is not enough, training must cross campuses, and cross-campus bandwidth only comes from DWDM. In 2000 DWDM multiplied a fiber's capacity tens of times, effective supply overshot demand, and bandwidth prices collapsed. This cycle it is the physical base of scale-across.	**Open: ITU / OIF (400ZR, 800ZR)**	Optical transport *Ciena · Nokia/Infinera · Cisco/Acacia*; coherent pluggables *Ciena · Marvell · Lumentum · Coherent*; coherent DSP *Ciena in-house · Marvell · Acacia*; fiber *Corning · Prysmian · YOFC*; long haul *Lumen · Zayo*
IP routing + SRv6 segment routing over IPv6	Writes path instructions into each IPv6 packet so the network forwards along a predetermined route: deterministic cross-site routing and traffic engineering.	Coordinate clusters in two campuses as one: path control, failover, cross-site visibility	Across campuses the problem moves from optimization to coordination: same GPU synchronization requirement, much larger failure domain. SRv6 extends scale-out load balancing logic between campuses; MRC builds on it. Arista calls this scale-across routing and guides $15-20B TAM by 2030.	**Open: IETF**	Routers and switches *Arista · Cisco · Nokia · Juniper/HPE*; silicon *Broadcom Jericho · Cisco Silicon One*
Front-end and storage fabric	The cluster's outward-facing network: user requests, object storage, datasets. Conventional Ethernet + IP, no RDMA-grade synchronization needed.	Inference requests in and out, training data reads, checkpoint writes	Not a bottleneck, but as inference share rises its traffic grows faster than the training network, the source of switch vendors' "growth beyond the back-end AI fabric". Vendors often merge it with scale-across to enlarge Ethernet's territory.	**Open**	Same suppliers as scale-out Ethernet

Source: vendor technical blogs and public standards material.

Figure 8 · From switch to transceiver

Top: the signal path from switch ASIC to fiber, what each part is and who makes it. Bottom: the same electro-optic conversion part in three forms (DSP pluggable, LPO/LRO, CPO) side by side. Two common misconceptions are corrected at bottom right.

Exhibit 16. The parts, the three form factors, and who wins



Where does the DSP live? The three form factors are one axis.



KEY TAKEAWAYS

- The signal path is switch ASIC → electrical trace → optical engine → fiber. Every part has a maker, and the DSP is the most expensive and most power-hungry one.
- The three forms are one axis: where the DSP lives. In the module (pluggable today), in the host (LPO), nowhere (CPO).
- Moving along that axis lowers power and latency but deepens the binding to the chip vendor and removes serviceability. Debate ① is about the pace of that move.

The three forms are one axis

The axis is "where does digital signal processing live": in the module (DSP pluggable), in the host (LPO), nowhere (CPO). Moving right lowers power and latency, deepens the binding to the system and the chip, and removes pluggable flexibility. So debate ① (CPO vs pluggables) has LPO in the middle: the pluggable camp's answer to power pressure, buying module makers a few more years before CPO arrives.

Why the DSP is the center of this figure

It is the most expensive and most power-hungry chip in a pluggable module, and the one both LPO and CPO remove. Marvell's optical DSP revenue shows how far the substitution has gone; module makers' gross margins show whether they are holding. These are the first two places debate ① shows up in financial statements.

Two common misconceptions

(1) "CPO is a new packaging of the existing module." Not quite: CPO does not modify the module, it replaces the module form factor by moving the optical engine into the chip package. What it brings is power, latency and bandwidth density, not reach. **(2) "Pluggables are more power-efficient than fixed optics."** That is vendor marketing. The pluggable's advantage is flexibility and serviceability; on power it is the highest of the three. Also, the ~1 km VCSEL figure is multimode short reach; single-mode uses EML or silicon photonics with a CW laser.

Figure 8 · From switch to transceiver detail

Exhibit 17. Figure 8 detail: the signal path and its parts

Switch	The network's logistics hub: a frame enters a port, the destination is looked up, it leaves the matching port. Memory is split into input queues, buffers and output queues, FIFO. Port speed is set by cable and port spec. A hub only repeats; a switch decides. *Systems: Arista · Cisco · Nvidia · Nokia · Celestica (white box). Software: EOS, NX-OS, SR Linux, SONiC*
Switch ASIC	A chip built to decide "where does this packet go" fast, not a general CPU (as a GPU is to graphics). Inside: packet processor, traffic management (priority, QoS), buffer memory, security, management interface. More ports means heavier decisions. **SerDes** is the block at the chip edge that turns parallel data into a high-speed serial signal. *Broadcom Tomahawk / Jericho · Nvidia Spectrum · Marvell Teralynx · Cisco Silicon One*
Electrical trace and retiming	PCB trace from ASIC to the front panel. At 100G / 200G per lane, tens of centimeters already distort the signal badly, so something must reshape it: a retimer on the board or a DSP in the module. The shorter this segment, the less you need them. That is the whole logic of CPO. *DSP: Marvell (Inphi) · Broadcom. Retimer: Astera Labs · Broadcom · Credo*
Optical engine / transceiver	The electro-optic conversion part: laser (VCSEL for multimode short reach within ~1 km; EML or silicon photonics with CW laser for single-mode, longer), laser driver (LDD), photodiode (light to current), TIA/LA (current to voltage and amplify), fiber coupling (a small alignment error loses a lot of light; this is where the yield barrier sits). In pluggable form it is a transceiver; packaged beside the chip it is a CPO optical engine. *Coherent · Lumentum · Innolight · Eoptolink · Fabrinet · AAOI. Lasers: Lumentum · Coherent · Broadcom · AOI · Yuanjie*
Fiber	Multimode (MMF), short reach. Single-mode (SMF), long reach. *Corning · Prysmian · YOFC*

Figure 8 · From switch to transceiver detail

Exhibit 18. Figure 8 detail: DSP pluggable vs LPO vs CPO side by side

	DSP-based pluggable transceiver · today's mainstream	**LPO / LRO linear pluggable optics** · pluggable, DSP removed	**CPO co-packaged optics** · optical engine packaged beside the switch ASIC
Architecture	The module carries its own DSP for retiming, equalization and forward error correction; self-contained, works with almost any host	Only high-linearity analog parts remain in the module (driver, laser, photodiode, TIA); digital signal processing is handed to the host ASIC in the switch or NIC	Optical engine and switch ASIC on one substrate; electrical trace is millimeters; no DSP or retimer; laser is external (ELS)
Power	Highest; the DSP is about half the module's power	Up to 50% below the DSP version	51.2T system power down 25-30% (Cisco); optical interconnect power down 70% (Broadcom Bailly)
Latency	Highest; DSP processing adds a stage	Low; no digital processing in the module	Lowest; shortest path
Reach and tolerance	Handles long, variable-quality links; high tolerance to impairment	Currently short reach (~500 m); low tolerance; must be qualified host by host	Silicon photonics single-mode ~2 km, same as pluggable single-mode; reach is set by the optics, not by CPO
Bandwidth density	Limited by front-panel port count; 3.2T per port not yet demonstrated in this form	Same as pluggable	Bandwidth per mm of chip edge up an order of magnitude; the route vendors show for 3.2T / 6.4T
Serviceability	Hot-swap; replace or upgrade one at a time; multi-vendor	Same as pluggable, but host compatibility is restricted	A failed optical engine cannot be swapped alone; serviceability preserved through the external laser
Standards	QSFP-DD, OSFP form factors; multi-vendor interoperable	LPO MSA, OIF CEI-112G Linear PAM4, CMIS plus VCS	OIF and OCP CPO specs; in practice led by Nvidia, Broadcom and TSMC roadmaps
Who supplies	Modules *Coherent · Lumentum · Innolight · Eoptolink · Fabrinet*; DSP *Marvell · Broadcom*	Same module makers, no DSP vendor needed; higher demands on switch-ASIC SerDes quality (Broadcom, Nvidia benefit)	Definition sits with *Nvidia · Broadcom · Marvell*; SiPh foundry *TSMC COUPE*; engines and parts *Coherent · Lumentum · TFC · Fabrinet · Corning · Senko*
Winners and losers	Status quo beneficiaries: module makers on high gross margin; DSP makers eat the most expensive chip in the module	**Loses:** DSP makers. **Wins:** module makers (pluggable form survives), switch-silicon makers	**Loses:** DSP makers, the pluggable module business. **Wins:** switch-silicon and accelerator makers, SiPh foundries, laser and passive suppliers
Where it fits	All scale-out and scale-across today; long, unpredictable links	Short, high-density links inside AI clusters: rack-to-rack leaf-spine where training sync is latency-sensitive	Next-generation 51.2T / 102.4T switches (scale-out); later inside the accelerator package (scale-up)

DISCLOSURES

This report was prepared by Sonny Xi for informational purposes only and does not constitute investment advice or a recommendation to buy or sell any security. The author may hold positions in securities mentioned in this report.

ABOUT THE ANALYST

Sonny Xi

Independent equity and macro research

Sonny Xi is an independent investment researcher covering equities and macro, with a focus on rates and geopolitics. His work pairs bottom-up company valuation, such as the three-scenario DCF framework used in his equity reports, with top-down macro analysis.

EDUCATION

M.S. in Finance (STEM-designated)
University of Illinois Urbana-Champaign, 2026

B.A. in International Affairs and Finance, cum laude
The George Washington University, Elliott School of International Affairs

CERTIFICATION

CFA Program
Registered for the Level I exam, February 2027

EXPERIENCE

Sell-side equity research
Tianfeng Securities: small- and mid-cap coverage, including institutional reports on nuclear fusion and 3D printing

Private equity due diligence
Lionchase North America and Sunrise Tsinghua Capital: underwriting, technology due diligence, DCF and LBO models

Credit risk modeling
CIBC: Excel dashboards for the U.S. risk model team

Buy-side portfolio research
University of Illinois RCMP: financials coverage and a three-scenario DCF on CME Group

Equity analyst
GW Investment Institute

CONTACT

Email
sonnyxi1205@gmail.com

Research website
sonnyxi-research.com

SELECTED RESEARCH

- Meta Platforms, Inc. (META): equity research report, October 2026
- CME Group Inc. (CME): equity research report, October 2026
- Amazon.com, Inc. (AMZN): equity research report, October 2026
- The Ensign Group (ENSG): sell recommendation
- H2 2026 Macro Outlook
- VAR macro forecasting model
- Constructing a Portfolio for a Five Percent Return Objective

ANALYST CERTIFICATION

The views expressed in this report accurately reflect the personal views of the analyst about the subjects covered in this report. This report is independent research and was not commissioned or paid for by any company or third party.