

THE BOTTOM LINE

Capability stops being about parameters. The constraint stack flipped to power, fiber, and packaging.

W19 is the week the AI compute story stopped being told in parameter counts. No new closed-frontier text flagship shipped — GPT-5.5 and Opus 4.7 still anchor the AA Intelligence Index — yet the most consequential capability move was a 300 MW SpaceX deal that let Anthropic double Claude Code rate limits inside a week. OpenAI shipped a voice trio (GPT-Realtime-2 + Translate + Whisper) with GPT-5-class reasoning baked in, jumping 15 points on Big Bench Audio; xAI moved Grok 4.3 to GA with 1M context at \$1.25/\$2.50 per Mtok; Google promoted Gemini 3.1 Flash-Lite to GA at \$0.25/\$1.50.

Meanwhile NVIDIA wrote two equity checks that buy zero transistors: \$500M in Corning warrants for a 10x expansion of US optical capacity, and up to \$2.1B in IREN warrants alongside a \$3.4B 5-year contract underwriting 5 GW of NVIDIA-aligned AI infrastructure. On the same Monday, OpenAI / AMD / Broadcom / Intel / Microsoft / NVIDIA jointly published Multipath Reliable Connection (MRC) as an OCP-released open RDMA transport — already in production at the Stargate Abilene site, Microsoft Fairwater, and OCI — collapsing 100,000-GPU clusters from three or four switching tiers down to two.

The procurement read: capability is being bought with photonics and power purchase agreements, not parameter counts. Boards should reset AI-vendor diligence to ask 'who's your power partner and which fabric standard do you ship on,' not 'what's your AA Index score.' Investors should treat NVIDIA's equity stakes in CoreWeave and now IREN as a circular-financing pattern that auditors will start framing — supplier financing customers who buy supplier product. Architects should add MRC compliance and AMD-as-substrate (ZAYA1's open frontier-class MoE trained end-to-end on MI300X this week) to procurement bake-offs that previously assumed NVIDIA

+ closed-fabric defaults. Operators in PJM territory should re-bid 5-7 year time-to-power for new Dominion-territory applications and pre-secure interconnect now.

AUTHOR

Brian Letort

BrianLetort.AI

PUBLISHED

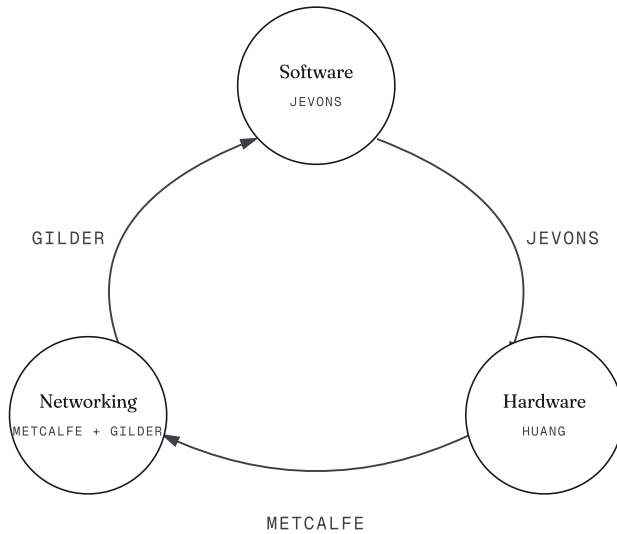
May 8, 2026

Issue 03

HOW WE READ THIS WEEK

Three lenses, one flywheel.

Cheaper inference pulls in more workloads. More workloads need more compute. More compute needs denser fabric. Denser fabric unlocks new architectures, which lower the cost of inference again. Read a week's news across that loop and the noise sorts itself out.



THIS WEEK'S ARC

All three lenses.

Software → Hardware new model capability creates new use cases that consume more compute (Jevons).

Hardware → Networking more compute means more nodes; the value of the connecting fabric scales as the square of the nodes (Metcalf).

Networking → Software denser, higher-bandwidth interconnect makes new model architectures viable (Gilder).

Hardware itself GPU performance doubles faster than Moore's Law (Huang).

THE BAR

The events that actually matter touch at least two of the three lenses. Single-lens reads are noise dressed up as motion. Each section of this brief grades its evidence and ties the implication back to the flywheel.

Software.

Model releases, pricing, capability benchmarks, license posture, and capability-risk disclosures.

MAY 7

EVENT 01

OpenAI ships GPT-Realtime-2 + GPT-Realtime-Translate + GPT-Realtime-Whisper, exiting Realtime API beta — first voice model with GPT-5-class reasoning, 128K context, adjustable effort, priced at \$32/\$64 per Mtok audio (Translate \$0.034/min, Whisper \$0.017/min); Zillow, Priceline, Deutsche Telekom named as early enterprises

OPENAI.COM/INDEX/ADVANCING-VOICE-INTELLIGENCE-WITH-NEW-MODELS-IN-THE-API, CHANNELNEWSASIA, 9TO5MAC, THENEXTWEB, MARKTECHPOST, LATENT.SPACE

MAY 6

EVENT 02

xAI moves Grok 4.3 to GA on the public API with 1M-token context, native video input, and aggressive pricing of \$1.25 / \$2.50 per Mtok (~40-60% below Grok 4.20); eight legacy Grok models scheduled for retirement on May 15

DOCS.X.AI/DEVELOPERS/MODELS, VENTUREBEAT, ARTIFICIAL ANALYSIS, HELP.APIYI.COM

MAY 6

EVENT 03

Anthropic doubles Claude Code 5-hour rate limits across Pro/Max/Team/Enterprise, removes peak-hours throttling for Pro/Max, raises Opus API limits — coordinated with the same-morning announcement of an Anthropic ↔ SpaceX deal: 300+ MW at Colossus 1 (Memphis), ~220K NVIDIA GPUs within the month; Claude Managed Agents ships Multiagent Orchestration (Netflix deployer), Outcomes (+8.4% docx / +10.1% pptx task success), and Dreaming (research preview) at the Code w/ Claude SF event

ANTHROPIC.COM/NEWS/HIGHER-LIMITS-SPACEX, CLAUDE.COM/BLOG/NEW-IN-CLAUDE-MANAGED-AGENTS, BLOOMBERG, CNBC, THE VERGE, ARS TECHNICA, SIMONWILLISON.NET

MAY 5

EVENT 04

US Center for AI Standards and Innovation (CAISI, Department of Commerce) signs pre-deployment evaluation agreements with Google DeepMind, Microsoft, and xAI — completing perimeter coverage of the top-five US labs with OpenAI and Anthropic already renegotiated; labs will provide early access including model versions 'with lowered safety protections' for national-security testing

BBC NEWS, CNBC, THE HILL, NEURA.MARKET

MAY 7
EVENT 05

Google promotes Gemini 3.1 Flash-Lite to GA on Gemini Enterprise at \$0.25 input / \$1.50 output per Mtok with adjustable thinking levels (minimal/low/medium/high) and full multimodal (text, code, image, audio, video, PDF); Gladly, JetBrains, and Astrocade named as enterprise users

CLOUD.GOOGLE.COM/BLOG/PRODUCTS/AI-MACHINE-LEARNING/GEMINI-3-1-FLASH-LITE-IS-NOW-GENERALLY-AVAILABLE, BLOG.GOOGLE, DEEPMIND.GOOGLE/MODELS/GEMINI/FLASH-LITE, SIMONWILLISON.NET

WHAT THIS MEANS

W19 is the week frontier text went into maintenance and capability moved into voice, throughput, and price-band compression. OpenAI's voice trio + Inworld TTS-2 turn 'voice model' into a frontier tier with its own benchmarks (Big Bench Audio, Audio MultiChallenge, Speech Arena Elo) — boards and architects should treat voice/IVR replacement as a 90-day procurement decision, not a 2027 plan. Anthropic's response to a compute deal — doubling paid-seat throughput rather than cutting list price — is the inverse of every other vendor's quarter and signals that the binding constraint on Claude was always GPU supply, not unit economics; operators locked into Claude Code can plan against materially higher concurrent throughput before mid-summer. CAISI completing the top-five-lab pre-deployment perimeter means every US frontier model now ships through a federal capability/risk gate before public release; pair with The Model Pulse for the seven new tree rows this week (gpt-5-5-instant, the Realtime trio, inworld-realtime-tts-2, zaya1-8b, emo-1b14b).

Hardware.

Silicon, density, packaging, memory supply, and the share of incremental compute going to custom silicon.

MAY 5

EVENT 01

AMD Q1 2026 8-K: data-center revenue \$5.8B (+57% YoY, +7% QoQ), Q2 guide \$11.2B (±\$300M); reaffirms Helios + Meta 6 GW Instinct deployment shipping H2 2026; Instinct revenue declined sequentially on China business shift

AMD 8-K (SEC EDGAR), AMD IR PRESS RELEASE, CNBC, REUTERS EARNINGS TRANSCRIPT

MAY 6

EVENT 02

SpaceX files for 'Terafab' Texas logic fab in Grimes County — Phase 1 ~\$55B, full buildout up to ~\$119B, jointly built and operated with Tesla and xAI, with Intel attached for design / fab / packaging on Intel 14A; tax-abatement filing public, hearing scheduled June 3

GRIMES COUNTY PUBLIC FILING, CNBC, TOM'S HARDWARE, TECHCRUNCH, THE VERGE, THE REGISTER

MAY 7

EVENT 03

NVIDIA + IREN strategic partnership: 5-year warrant for up to 30M IREN shares @ \$70 (~\$2.1B) plus separate \$3.4B 5-year AI Cloud contract; up to 5 GW of NVIDIA-aligned AI infrastructure on IREN's pipeline with the 2 GW Sweetwater (TX) campus as flagship — Phase 1 deployment beginning 2027

IREN PRESS RELEASE + Q3 FY26 INVESTOR UPDATE, GLOBENEWSWIRE, CAPACITY MEDIA, THENEXTWEB

MAY 8

EVENT 04

TSMC April 2026 monthly revenue 6-K: NT\$410.73B (~US\$13.05B), +17.5% YoY, -1.1% MoM (first MoM dip of 2026); YTD revenue +29.9% YoY through April against a CoWoS line still booked through 2027 — consistent with packaging being the binding constraint, not wafer fab

TSMC 6-K (SEC), TSMC IR, STOCKTITAN, ETMANUFACTURING, MARKETSCREENER

WHAT THIS MEANS

The binding constraint has migrated off the die. Two of NVIDIA's biggest W19 checks — Corning warrants (see Networking lens) and IREN warrants — buy no transistors; they buy optical-fiber capacity and 5 GW of deployment runway. AMD's \$5.8B data-center quarter (+57% YoY) is the first merchant-GPU read of the new cycle and confirms a credible second supply path with Meta's 6 GW Helios commitment shipping in H2; SpaceX's \$55-119B Terafab filing brings a fourth US foundry option onto the radar in a market that had three. TSMC's first MoM dip against +29.9% YoY YTD is consistent with packaging — not wafers — being the rate-limiter on Huang's doubling cadence. For boards, the procurement framework reset is: power, packaging, and photonics now price like the substrate; for investors, watch NVIDIA's equity stakes in customers (CoreWeave, IREN) as a circular-flow risk auditors will start framing. Architects sizing 2027 capacity should pre-bid AMD-substrate options against NVIDIA defaults — ZAYA1 this week is the first arxiv-grade frontier-class MoE trained end-to-end on AMD MI300X.

Networking.

Interconnect, fabric standards, optical capacity, sovereign and operator-level networking products.

MAY 5

EVENT 01

OpenAI / AMD / Broadcom / Intel / Microsoft / NVIDIA jointly publish Multipath Reliable Connection (MRC) as an OCP-released open RDMA transport — running in production at the Stargate Abilene site, Microsoft Fairwater, and Oracle Cloud Infrastructure; collapses 100,000+ GPU clusters from three or four switching tiers down to two

OPENAI PRESS POST, AMD BLOG, NVIDIA BLOG, MICROSOFT AZURE HPC BLOG, BROADCOM BLOG, ORACLE CLOUD INFRASTRUCTURE BLOG, OCP MRC V1.0 SPEC, ARXIV 2605.04333 (JOINT OPENAI / MICROSOFT / NVIDIA PAPER)

MAY 5

EVENT 02

NVIDIA + Corning announce a multi-year supply and technology partnership with up to ~\$3.2B in equity exposure (\$500M pre-funded share purchase + warrants on 18M shares @ \$180), funding a 10x expansion of US optical-connectivity manufacturing, +50% US fiber capacity, three new North Carolina and Texas plants, and 3,000 jobs

CORNING 8-K + EX 99.1 (SEC EDGAR), NVIDIA NEWSROOM, CNBC, BLOOMBERG, TOM'S HARDWARE, RCR WIRELESS

MAY 5

EVENT 03

Astera Labs ships the Scorpio X-Series 320-lane Smart Fabric Switch — the industry's largest open memory-semantic scale-up switch, with hardware-accelerated Hypercast and In-Network Compute claiming up to 2x collective-operation speedup across CXL, Ethernet, NVLink Fusion, PCIe, and UALink

ASTERA LABS IR RELEASE, GLOBENEWSWIRE

MAY 4

EVENT 04

GlobalFoundries unveils SCALE — the first OCI MSA-capable co-packaged-optics platform — with 8λ and 16λ bidirectional DWDM demonstrated on its silicon-photonics PDK, plus 50G and 100G micro-ring modulators and TSV-ready 2.5D / 3D stacking down to sub-45μm copper pitch

GLOBALFOUNDRIES PRESS RELEASE, MICROWAVE JOURNAL, NEW ELECTRONICS

MAY 5
EVENT 05

EllaLink lights a new 670 km submarine-cable extension into Nouadhibou, Mauritania — the country's second direct path to European cloud and AI services, EU-co-funded via the Connecting Europe Facility, two fiber pairs scalable to multi-terabit, with a neutral landing station

DATACENTER DYNAMICS, CONNECTING AFRICA, TECHAFRICA NEWS, TOTAL TELECOM

WHAT THIS MEANS

This is the most consequential AI-fabric week of the year. MRC as an OCP-blessed multi-vendor RDMA transport (already running at Stargate Abilene, Microsoft Fairwater, OCI) commoditizes the inside-the-cluster fabric at the protocol layer the same week NVIDIA pre-allocates the underlying optical supply chain (\$3.2B equity exposure to Corning, 10x US optical capacity). Durable value migrates from cage fabric to cross-cage and cross-DC interconnect — the operator category that controls dense metro fiber and meet-me ecosystems wins the next two upgrade cycles. For investors and boards, this is Metcalfe at the standards layer: every additional MRC-conformant NIC, switch, and OCI MSA optical engine multiplies the value of every other connected node, while closed alternatives lose option value by the week. Architects should re-test single-vendor scale-up roadmaps against an open MRC + UEC + ESUN + OCI MSA blueprint stack now bookable today; operators should treat the Mauritania landing as a template for how sovereign-AI capacity now follows submarine fiber rather than the other way around.

CAPITAL FLOW

Money in, revenue out.

Capital deployed (forward) vs revenue out (quarterly or run-rate). Burn-to-revenue = revenue / capital — lower means more out than in. Bars normalized to 58.0 \$B; On-Prem revenue is indirect.

Frontier Labs

OPENAI, ANTHROPIC, GOOGLE DEEPMIND, XAI

BURN / REV

0.5



~\$52B



~\$45B

Anthropic round still pending close at \$850-900B target with \$30B+ disclosed run rate (some sources cite ~\$45B internal); a May 8 FT report floats potential repricing toward \$1T but no primary close. Anthropic's 300 MW SpaceX compute deal (May 6) was the W19 capability event, not the raise.

Hyperscaler-Hosted

MICROSOFT AZURE, GOOGLE CLOUD, AWS, META

BURN / REV

0.22



~\$58B



~\$13B

Mid-cycle quiet between Q1 prints (W18) and NVIDIA Q1 FY27 (May 20). Forbes coverage May 6 extends the W18 OpenAI-on-Bedrock procurement story into federal/enterprise narrative; Oracle GA's OCI Compute with NVIDIA RTX PRO 6000 Blackwell at \$4.50/GPU-hr (May 7). No fresh capex print. Carry W18 priors flat.

Neoclouds

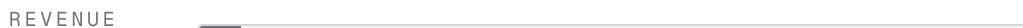
COREWEAVE, CRUSOE, NEBIUS, APPLIED DIGITAL, IREN, LAMBDA

BURN / REV

0.18



~\$28B



~\$3.0B

The loudest W19 across categories. CoreWeave Q1 print (May 7) discloses \$2.08B revenue (+112% YoY) and a record \$99.4B revenue backlog; IREN-NVIDIA \$3.4B 5-year contract + \$2.1B warrant + 5 GW partnership (May 7); Lambda \$1B credit facility (4x prior); plus Cipher Mining + Digi Power X + DeepInfra + RadixArk smaller tickets. Capital-in stepped up materially via debt + customer-contract value; revenue out is approaching 1.5x the W18 prior driven by CoreWeave alone.

On-Prem / Hybrid

ENTERPRISE GPU CLUSTERS, SOVEREIGN AND NATIONAL PROGRAMS

BURN / REV

n/a



~\$42B



Indirect

Two new UAE platform-tier sovereign initiatives (Core42 + Solutions+, e& 'Agents Factory') in window; no disclosed dollar values. TSMC monthly print (May 8) confirms packaging is binding — squeeze on every on-prem server OEM that ships discrete AI accelerator boards or HBM-heavy SKUs. Carry W18 priors flat.

SIGNAL VS NOISE

What's real, what's noise.

5 claims that drove headlines this week, scored 1–5 on source quality and triangulation. 2 flagged as noise. The bar is at least one explicit noise call per issue.

5 / 5

CLAIM 01

OpenAI / AMD / Broadcom / Intel / Microsoft / NVIDIA jointly publish Multipath Reliable Connection (MRC) as an OCP-released open RDMA transport — running in production at the Stargate Abilene site, Microsoft Fairwater, and Oracle Cloud Infrastructure; collapses 100,000+ GPU clusters from three or four switching tiers down to two.

SOURCES: OPENAI PRESS POST, AMD BLOG, NVIDIA BLOG, MICROSOFT AZURE HPC BLOG, BROADCOM BLOG, ORACLE CLOUD INFRASTRUCTURE BLOG, OCP MRC V1.0 SPEC PDF, ARXIV 2605.04333 (JOINT OPENAI / MICROSOFT / NVIDIA PAPER)

Real and structural. Six independent primary vendor blogs plus an OCP-published spec plus a peer-reviewed paper plus production deployment make this the most-confirmed networking event of 2026 to date. The procurement implication is durable: closed alternatives lose option value by the week. Read changes only if a major hyperscaler explicitly defects from MRC for its next-gen scale-out fabric — none have signaled intent.

5 / 5

CLAIM 02

CoreWeave Q1 2026 revenue backlog reaches \$99.4B as of March 31, 2026 — up from \$66.8B at YE 2025 (+49%); revenue \$2.08B (+112% YoY); \$31-35B 2026 capex guide tied entirely to signed customer contracts.

SOURCES: COREWEAVE Q1 2026 PRESS RELEASE (MAY 7); MORNINGSTAR MIRROR; SEC EDGAR 10-Q FILING WINDOW MAY 11-21

Real and operationally consequential. The disclosure validates the anchor-tenant contract structure that underpins neocloud capex commitments. Read changes if the May 11-21 10-Q reveals contract-amendment language or if Q2 print shows backlog churn or recognition restatement on the 36% expected to convert in 24 months.

5 / 5

CLAIM 03

NVIDIA + Corning announce a multi-year supply and technology partnership with up to ~\$3.2B in equity exposure (\$500M pre-funded share purchase + warrants on 18M shares @ \$180), funding a 10x expansion of US optical-connectivity manufacturing, +50% US fiber capacity, and three new NC / TX plants.

SOURCES: CORNING 8-K + EX 99.1 (SEC EDGAR), NVIDIA NEWSROOM, CNBC, BLOOMBERG, TOM'S HARDWARE, RCR WIRELESS

Real. Corning shares closed +12% and NVIDIA +6% on the announcement. The structural read: NVIDIA is now an equity holder in two of its largest neocloud customers (CoreWeave Class A; IREN warrant) AND its largest US optics supplier — auditors and short-side analysts will start framing this as a circular-financing pattern. Read changes if NVIDIA discloses warrant terms and exercise pricing logic on May 20 (Q1 FY27 print) and if similar arrangements with three+ counterparties surface.

2 / 5 -
NOISE

Anthropic's \$50B funding round will reprice toward a \$1 trillion valuation as investor demand intensifies (FT report May 8).

CLAIM 04

SOURCES: FINANCIAL TIMES VIA TECH STARTUPS SUMMARY (MAY 8) — SINGLE SECONDARY SOURCE LAYERED ON THE UNDERLYING \$850-900B REPORTING

Likely modest hype layered on a real underlying deal. The \$1T number is one FT report being amplified; multi-source reporting still has the round at \$850-900B and the deal has been 'within two weeks' of closing for two weeks already. Watch the May 12-22 window for primary close confirmation. Read changes only on Anthropic primary disclosure: \$1T = signal, \$900B = the FT marker was wishful pricing.

1 / 5 -
NOISE

Panthalassa's \$140M Series B (Thiel-led, May 4) for ocean-powered AI compute validates wave-energy AI inference at sea as a viable architecture.

CLAIM 05

SOURCES: BUSINESSWIRE PRESS RELEASE MAY 4, 2026; NO CUSTOMER DISCLOSURE, NO UNIT-ECONOMICS DISCLOSURE, COMMERCIAL DEPLOYMENT TARGETED 2027

Mostly narrative at this stage. The \$140M is real and the investor list is real, but the technical premise — densely-packed AI accelerators surviving in marine atmosphere with wave-energy power — has well-known thermal, corrosion, vibration, latency, and serviceability problems. No anchor customer, no production node uptime data, no cost/MWh disclosure. Read changes only on a deployed pilot node with measurable inference uptime and per-token economics.

EARLY WARNING PANEL

The levers we monitor.

10 metrics, current vs prior period. **5 rising**, **0 falling**, 5 steady. Each metric carries a threshold value where the read materially changes.

METRIC	CURRENT	PRIOR	DIR	THRESHOLD / NOTE
Frontier lab cash position (avg months runway, top 3)	~30 mo	~30 mo	→	<18 mo triggers re-rating risk
Hyperscaler capex / AI revenue ratio (top 4 weighted)	~5.0-5.2	~5.0-5.2	→	>6.0 invites investor pushback at next earnings
CoreWeave revenue backlog	\$99.4B	~\$87B+ (est)	↑	Conversion velocity matters more than gross figure
NVIDIA Q-over-Q data center revenue	\$62.3B (Q4 FY26)	\$62.3B (Q4 FY26)	→	Q1 FY27 print May 20 — Goldman Sachs forecasts ~\$70B+ DC implied
Open vs closed gap on SWE-Bench Pro (coding)	Closed +6 to +20pp	Closed +6 to +19pp	→	Sustained open lead reshapes enterprise procurement
Sovereign AI commitments (count / aggregate \$)	10 / ~\$80B+	8 / ~\$80B+	↑	—
PJM 2026/27 capacity auction price (\$/MW-day)	\$329.17	\$329.17	→	11x in 24 months — power is the new binding constraint
Time-to-power, busiest US markets (months)	60-84 (new PJM); 36-48 (existing PJM queue)	36-48 (PJM large >100 MW); ~53 nat'l avg	↑	—
Cost-per-task, frontier reasoning model	~\$0.10-\$0.15 (effective, with hidden reasoning tokens)	~\$0.05	↑	Methodology correction: published rates understate effective cost

METRIC	CURRENT	PRIOR	DIR	THRESHOLD / NOTE
Custom silicon share of incremental AI compute	~ 33-36% (estimated)	~31%	↑	>35% materially compresses merchant GPU pricing

THRESHOLD VALUES ARE THE POINTS WHERE THE READ FLIPS. CROSSINGS ARE FLAGGED IN THE ISSUE BODY WHEN THEY HAPPEN.

PREDICTIONS

What we expect next.

5 falsifiable, time-bounded predictions. Each carries a confidence (1–99%, never 50%), a deadline, and a specific signal we'll watch. Future issues score them hit / miss / partial.

PREDICTION 01 SOFTWARE

65%

At least one Fortune 500 enterprise discloses replacement of an existing voice / IVR platform with a frontier voice model (GPT-Realtime-2, Inworld TTS-2, or competitor) on a contract greater than \$25M annualized.

BY JULY 31, 2026

TRIGGER: Enterprise customer-disclosure announcements; Q2 vendor commentary from OpenAI / Anthropic / Inworld; named Zillow / Priceline / Deutsche Telekom contract values.

PREDICTION 02 NETWORKING

70%

At least one major hyperscaler other than Microsoft publicly discloses a production AI fabric running MRC at greater than 50,000-GPU scale.

BY AUGUST 31, 2026

TRIGGER: Google I/O 2026 (May 19–20) fabric session; AWS re:Invent prep coverage; OCP Future Technologies Symposium filings; vendor blog disclosures.

PREDICTION 03 CAPITAL

75%

Anthropic's pending raise closes at a final post-money valuation between \$850B and \$1T, with the lead investor and check size publicly disclosed.

BY JUNE 15, 2026

TRIGGER: Anthropic primary disclosure; SEC Form D if applicable; secondary market activity; lead investor naming.

PREDICTION 04 HARDWARE

55%

A second public open-weights frontier-class model trained end-to-end on AMD silicon (MI300X or MI400) is announced, confirming AMD-as-training-substrate is structural rather than a one-off proof.

BY SEPTEMBER 30, 2026

TRIGGER: Vendor announcements; arxiv preprints; AMD customer-list updates at Advancing AI 2026 (July 22–23); Rackspace + AMD MOU production milestones.

PREDICTION 05 CAPITAL

70%

At least one major equity research firm or audit-grade publication explicitly frames NVIDIA's 2026 customer-equity stakes (CoreWeave + IREN, plus any subsequent counterparty) as a circular-financing risk in a written report or audit qualifier.

BY JULY 31, 2026

TRIGGER: NVIDIA Q1 FY27 earnings call (May 20); Q2 audit cycle commentary; sell-side and short-seller report releases.

WATCHLIST

On the radar.

5 catalysts in the next 7–14 days that would change the read materially. Watching these tells us whether the thesis is strengthening or weakening.

MAY 19 - 20

WATCH 01

Google I/O 2026

Likely venue for Gemini 3.2 Flash confirmation (the May 5 leak surfaced in iOS app + AI Studio metadata at \$0.25 / \$2.00 per Mtok) and any Gemini 3.2 Pro announcement that would force a frontier scorecard refresh. Watch for fabric-side disclosures: Google has been silent through W18-W19 while OpenAI / Microsoft / NVIDIA shipped MRC publicly.

MAY 20

WATCH 02

NVIDIA Q1 FY27 earnings

First disclosure on Vera Rubin ramp velocity, Samsung HBM4 second-source effect, custom-silicon competitive pressure (AMD Helios + Meta 6 GW, Trainium \$225B book), and warrant-stake terms for the Corning, IREN, and CoreWeave equity investments. Goldman preview (May 6) raised the bar to \$80.05B total / ~\$70B+ DC implied. A beat resets the H2 hyperscaler capex tail; a miss flattens the doubling slope; circular-flow framing gets stress-tested on the call.

MAY 11 - 21

WATCH 03

CoreWeave 10-Q filing

Audited detail behind the W19 Q1 print: \$99.4B backlog conversion schedule, \$31-35B 2026 capex composition, customer concentration disclosure, NVIDIA \$2B Class A equity terms. The ~36% expected to recognize within 24 months is the line auditors and short-side analysts will read first. Material differences between management commentary and 10-Q footnotes would re-rate the entire neocloud category.

MAY 12 - 22

WATCH 04

Anthropic round close

Talks have been 'within two weeks' of close for two weeks. The price (\$850B floor / \$900B mid / \$1T per FT) re-rates the entire frontier-lab valuation curve. A walk-away from the price would be the loudest re-rating signal in a year; a clean close at \$1T validates the FT signal as primary.

MAY 19 - 23

WATCH 05

Microsoft Build 2026 + Computex 2026 prep

Microsoft Build (May 19-22) is the venue for Azure AI roadmap, Maia 200 numbers, and any production MRC fabric disclosure. Computex 2026 has now slipped to June 2-5 (with Jensen Huang's GTC Taipei keynote June 1) but vendor prep coverage will start in W21 — a bridge week for hardware-and-fabric news cycles.

METHODOLOGY AND AUTHORSHIP

How this brief is built.

Compiled from public announcements, SEC filings, earnings transcripts, and official lab and vendor publications. Every quantitative claim is graded on source quality. Every prediction is falsifiable, time-bounded, and scored hit / miss / partial / pending in future issues.

SIGNAL SCORE RUBRIC

- 5 / 5** SEC filing or audited disclosure. Multi-source independent confirmation. Operational, not aspirational.
- 4 / 5** Earnings-call disclosure or primary lab/vendor announcement. Two or more independent sources.
- 3 / 5** Single primary source. Reasonably consistent with sector data.
- 2 / 5** Analyst report or secondary press only. Or single primary source with credibility caveats.
- 1 / 5** Rumor, social media, single non-primary source. Or contradicted by alternate primary sources.

ANYTHING GRADED 2 OR BELOW IS FLAGGED AS NOISE.

PREDICTION OUTCOMES

- HIT** The specific observable outcome occurred by the deadline.
- MISS** The deadline passed and the outcome did not occur.
- PARTIAL** A meaningfully similar outcome occurred but not the literal wording.
- PENDING** Deadline not yet reached.

HIT RATE OF 65-75% IS THE TARGET. ABOVE 80% MEANS TOO CONSERVATIVE; BELOW 50% MEANS THE FRAMEWORK IS WRONG.

AUTHORSHIP**Brian Letort**

BrianLetort.AI · Independent analysis. All sources public. Not investment guidance.

THIS ISSUE**Issue 03**

Week 19 of 2026 · May 8, 2026

WEB

brianletort.ai/industry

Past issues, the working framework, and the LLM Evolutionary Tree companion.