SENTIENT REVIEW

Forthcoming · Not Published

The Economics of AI Infrastructure: Capacity, Capital and Compute

The Economics of AI · Part I · Capital & Infrastructure Economics

Author: Fritzie M. Alex

Editor: Fritzie M. Alex

Publisher: Sentient Review

Reserved Publication № 10

https://sentient-review.com/perspectives/the-economics-of-ai-infrastructure

AI infrastructure investment becomes economically productive through the alignment of capital commitments, physical delivery, workload performance and financial returns. Across fourteen chapters, this publication examines financing, power, GPU architecture and depreciation, utilization, fully loaded compute costs, contractual revenue, cash flow, risk and capital allocation.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Publication Editorial Note

This publication examines AI infrastructure from capital commitment through productive operation and investment judgment. Its five parts distinguish financing and physical delivery, workload and asset economics, financial performance and capital allocation.

Company figures retain their stated reporting periods and accounting boundaries. Benchmark results describe performance under documented conditions. Modeled estimates and hypothetical financial examples are distinguished from reported outcomes.

Investment interpretations are analytical rather than company-specific valuation conclusions. Project-level decisions require verified contractual and asset-level evidence, dated cash flows, explicitly defined assumptions and appropriately supported discount rates.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Contents

Part I · The Capital Commitment

01 · The Architecture of AI Capital

02 · Financing the AI Infrastructure Cycle

Part II · The Physical Infrastructure

03 · From Announced Capacity to Operational Compute

04 · Power as an Economic Constraint

05 · The Economics of GPU Architecture

Part III · The Productive Asset

06 · The GPU as a Depreciating Financial Asset

07 · Capacity Utilization and Commercial Productivity

08 · Training, Inference and Workload Economics

Part IV · Infrastructure Financial Economics

09 · The Fully Loaded Cost of AI Compute

10 · Revenue Quality and Infrastructure Contracts

11 · Cash Flow and Return on Invested Capital

12 · Infrastructure Risk and Economic Sensitivity

Part V · Capital Productivity and Investment Judgment

13 · Where AI Infrastructure Value Accumulates

14 · The Economics of Productive AI Infrastructure

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 01 · The Architecture of AI Capital

AI infrastructure investment connects semiconductor suppliers, infrastructure operators and cloud platforms through distinct capital and revenue cycles. This chapter examines how capital moves through that system, becomes productive capacity and generates financial returns.

Artificial intelligence is becoming an industrial capital system. Its expansion connects semiconductor production with physical infrastructure investment and commercial computing services.

The financial architecture deserves examination because capital enters this system at different points and generates returns on different schedules.

Alphabet reported $91.4 billion in capital expenditure during 2025, compared with $52.5 billion in 2024. These company-wide investments primarily support technical infrastructure and other property assets.

The scale establishes the financial significance of infrastructure expansion. The underlying economics depend on how investment becomes productive capacity and how that capacity generates returns.

Capital Across the AI Supply Chain

The AI infrastructure economy connects businesses with distinct financial roles.

Semiconductor companies develop accelerated computing technology. Equipment manufacturers integrate that technology into systems. Infrastructure developers establish the facilities required for operation. Cloud providers convert computing capacity into commercial services.

Each participant earns revenue through a different transaction.

For an accelerator supplier, an equipment sale represents commercial revenue. For the purchaser, the same transaction creates an asset investment.

This distinction is central to understanding capital flows.

Revenue recognized by one participant can represent capital expenditure for another. Aggregating both as independent economic value would misrepresent the system.

The relevant question is where capital ultimately produces earnings and cash generation after the costs of delivering each product or service.

The distribution of value depends on contractual position and the capital required to sustain commercial activity.

The Conversion of Capital Into Productive Assets

Infrastructure investment develops over time.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Alphabet's 2025 Form 10-K describes data-center construction as a multiyear process. Technical assets can remain under construction or assembly before becoming ready for their intended use.

The company begins depreciation when qualifying property and equipment enter service.

These disclosures establish a distinction between capital deployment and operating readiness.

Expenditure can occur during development. Asset commissioning determines when infrastructure becomes available for productive operation. Depreciation then reflects the accounting allocation of asset cost over its estimated useful life.

Commercial revenue follows the business model supported by the infrastructure.

For an internal cloud platform, additional capacity may support existing services or new products. A specialized infrastructure provider may monetize capacity through customer contracts.

The economic return depends on the contribution generated over the investment horizon.

The Financial Position of Infrastructure Owners

Infrastructure ownership introduces a different financial exposure from equipment supply.

The owner must finance the assets required to deliver computing services. Revenue depends on commercial demand and contractual arrangements.

CoreWeave provides a documented example.

Its FY2025 Form 10-K describes asset-level debt financing supported by take-or-pay customer contracts. Corporate financing supplements these arrangements.

At December 31, 2025, CoreWeave reported $60.7 billion in remaining performance obligations.

The disclosure establishes the scale of contracted future revenue under the company's accounting definitions.

It also illustrates how customer commitments can support infrastructure financing.

The financial relationship connects assets being deployed with expected future service revenue. Capital recovery depends on the performance of those arrangements over time.

The financing structures and their associated risks receive detailed treatment in Chapter 02.

The Timing of Economic Value

The investment cycle creates several distinct financial events.

Capital commitment establishes an obligation or investment intention according to its contractual terms.

Cash expenditure records the deployment of financial resources.

Operational readiness establishes the availability of productive assets.

Commercial activity generates revenue according to the applicable business model.

Operating contribution reflects revenue after the relevant costs of delivering services.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Capital productivity evaluates those returns against the resources invested.

These measures describe different dimensions of the same economic system.

A financial assessment must preserve their timing and measurement boundaries.

Infrastructure construction may precede commercial activity by several years. Customer commitments may arise before capacity becomes operational. Revenue recognition can follow contractual schedules that differ from cash collection.

Understanding these differences is essential to interpreting the financial performance of AI infrastructure investment.

Where Economic Value Accumulates

The AI supply chain creates distinct opportunities for financial returns.

Technology suppliers monetize intellectual property and equipment. Infrastructure operators commercialize deployed computing assets. Integrated cloud platforms combine infrastructure with software and customer distribution.

The capital requirements of these businesses differ.

Equipment suppliers invest in product development and manufacturing relationships. Infrastructure operators carry substantial physical asset exposure. Cloud platforms distribute capital investment across a broader service portfolio.

Their financial performance therefore requires different measures.

Revenue growth establishes commercial scale. Operating margins describe earnings performance within a defined accounting boundary. Cash flow reveals the timing of financial receipts and expenditure.

Return on invested capital connects operating earnings with the capital supporting them.

These measures help establish where financial value is generated and retained.

A complete comparison requires consistent definitions and an understanding of the business model underlying each result.

The Capital Allocation Judgment

AI infrastructure investment connects technology architecture with financial performance.

The physical system determines the productive capacity available for commercial use. Customer demand establishes the opportunity to monetize that capacity. Operating economics determine the contribution generated during its useful life.

Capital allocation must consider all three.

The financial attractiveness of an investment depends on the returns produced relative to the resources committed and the risks assumed.

This relationship extends beyond the scale of infrastructure expenditure.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

It concerns the productivity of the capital deployed.

The economics of AI begin with capital allocation. Their financial significance emerges through the conversion of that capital into productive assets and sustainable returns.

Sources

2025 Form 10-K

Alphabet Inc.

Company Filing

February 5, 2026

https://www.sec.gov/Archives/edgar/data/1652044/000165204426000018/goog-20251231.htm

Capital expenditure, infrastructure development and property depreciation.

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Infrastructure financing, customer commitments and remaining performance obligations.

Editorial Note

Financial figures reflect the companies' FY2025 disclosures. Capital-flow and value-capture interpretations represent analysis of the reported business models.

The discussion distinguishes reported financial facts from economic interpretation. Subsequent chapters examine the financing and operating conditions governing infrastructure returns.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 02 · Financing the AI Infrastructure Cycle

AI infrastructure financing connects equipment-backed borrowing with long-term compute contracts. This chapter examines how asset productivity and customer commitments shape capital recovery and financial exposure.

The financing of AI infrastructure is becoming a defining feature of the industry's capital architecture.

Computing capacity requires substantial investment before commercial operation. Equipment procurement creates immediate financial commitments. Physical infrastructure follows construction schedules. Customer revenue develops as contracted services become available.

The financing structure connects these different timelines.

CoreWeave provides a useful case. Its 2025 financial disclosures reveal an infrastructure business supported by equipment-backed borrowing and long-term customer commitments. The arrangements demonstrate how future compute demand can finance present investment.

The central economic consideration is the alignment between asset productivity and the capital obligations supporting it.

Capital Ahead of Commercial Operation

An AI data center requires coordinated investment across physical infrastructure and computing systems.

Site development establishes the facility. Electrical infrastructure determines available power. Accelerator procurement commits capital to equipment whose commercial value depends on deployment and productive use.

Each investment follows its own schedule.

Construction expenditure can precede commissioning by several years. Equipment financing may begin during installation. Customer commitments can establish future revenue expectations while the associated capacity remains under development.

The resulting interval has financial consequences.

Borrowing costs accrue according to contractual terms. Development expenditure consumes liquidity. Commissioning determines when assets become available to serve commercial workloads.

A shorter delivery period can accelerate capital recovery. Extended development schedules increase financing requirements and expose investors to changes in market conditions.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Commissioning is therefore a financial milestone as well as an engineering milestone.

Financing Against Computing Assets

CoreWeave finances infrastructure through several capital sources.

Its 2025 Form 10-K identifies asset-level borrowing supported by take-or-pay customer contracts. Corporate debt and equity provide additional financing.

At December 31, 2025, the company reported approximately $9.7 billion outstanding under delayed-draw term loan facilities.

These facilities are secured by assets associated with the underlying customer contracts and pledged contractual cash flows. Borrowing occurs as infrastructure is developed to meet customer requirements.

Repayment follows contractual schedules supported by expected customer receipts.

The arrangement links three economic interests.

Lenders provide capital against secured assets and contractual payment streams. Infrastructure operators undertake delivery and operating obligations. Customers establish commercial demand through capacity commitments.

The relationship between these parties shapes financial exposure.

Asset-backed borrowing can support infrastructure expansion by connecting financing to identifiable equipment and revenue contracts. Its effectiveness depends on the quality of the collateral and the durability of the expected cash flows.

Customer Contracts as Financial Support

Long-term customer commitments have become important financing instruments within the AI infrastructure economy.

CoreWeave primarily serves customers through multi-year take-or-pay arrangements. Customers purchase specified capacity over agreed contractual periods.

The company also provides consumption-based services.

These models establish different relationships between utilization and revenue.

Committed capacity creates contractual payment obligations. Consumption-based services generate revenue through actual usage.

At December 31, 2025, CoreWeave reported $60.7 billion in remaining performance obligations. Its committed contracts had a weighted-average duration of approximately five years.

Management expected 43% of the outstanding obligations to be recognized during the following 24 months.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

That proportion represents approximately $26.1 billion in expected future revenue recognition, subject to the company's disclosed contractual and accounting assumptions.

The distinction between contracted revenue and realized cash flow is central to financing analysis.

Remaining performance obligations provide visibility into future service revenue. Operating cash generation depends on delivery and payment timing.

Contract quality therefore becomes a determinant of financing capacity.

Customer Concentration and Financial Exposure

Anchor customers can support infrastructure development through substantial commercial commitments.

CoreWeave's relationship with Microsoft illustrates this structure.

Microsoft accounted for approximately 67% of CoreWeave's recognized revenue during 2025.

Such concentration can strengthen demand visibility while increasing dependence on a major counterparty.

The financial significance extends beyond the customer's creditworthiness.

Contract duration influences the period of expected revenue. Payment provisions determine cash timing. Renewal economics shape the commercial life of the infrastructure beyond its initial commitments.

CoreWeave describes credit monitoring and customer prepayments as mechanisms used to manage counterparty exposure.

For capital providers, contractual cash-flow quality is fundamental to evaluating repayment capacity.

For infrastructure operators, customer concentration influences future capacity planning and refinancing resilience.

SENTIENT REVIEW

The Financial Life of Computing Assets

AI accelerators introduce a distinctive relationship between technology evolution and capital recovery.

New hardware generations can improve computational efficiency. Existing equipment may remain commercially productive through alternative workload assignments.

The economic life of an accelerator depends on its ability to generate useful computing services at competitive cost.

Financing arrangements follow separate schedules.

Debt matures according to contractual terms. Accounting depreciation allocates asset cost over an estimated useful life. Commercial productivity changes with workload demand and technology performance.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

These timelines influence investment returns.

CoreWeave's disclosures identify equipment upgrades and useful-life estimates as material considerations. The company continually evaluates the deployment and replacement of infrastructure components.

The financing assessment must therefore consider future equipment investment alongside existing repayment obligations.

Chapter 06 examines the relationship between hardware economics and asset depreciation in greater depth.

The Cost of Infrastructure Capital

Borrowing costs influence the returns available to infrastructure investors.

CoreWeave reported $1.229 billion in net interest expense for 2025.

Its year-end market-risk disclosure estimated approximately $107 million in annual interest-expense sensitivity for a 100-basis-point movement in interest rates, based on the outstanding debt exposure used in that calculation.

The company also employs interest-rate swaps to manage certain floating-rate obligations.

These disclosures demonstrate the financial significance of borrowing terms.

Operating earnings reflect the performance of the infrastructure business before financing costs. Interest expense affects earnings available after debt financing. Principal repayment determines additional demands on cash.

The distinction becomes important during periods of rapid infrastructure expansion.

A business can generate operating cash while continuing to require substantial capital for new facilities and computing equipment.

The investment assessment must therefore consider operating performance alongside financing obligations and future capital requirements.

The Capital Allocation Judgment

AI infrastructure financing connects future compute demand with present capital commitments.

Equipment-backed borrowing supports asset deployment. Customer contracts provide commercial visibility. Prepayments can strengthen liquidity during expansion.

Financial resilience emerges when these arrangements remain aligned with the productive life of the infrastructure.

For investors, the relevant consideration is the durability of cash flows supporting capital recovery.

For cloud operators, financing terms influence expansion capacity and the economics of equipment renewal.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Enterprise technology buyers also have an interest in supplier financial resilience because long-term compute services depend on continued infrastructure operation.

The strength of AI infrastructure financing rests on the alignment of asset productivity, contractual cash flows and capital obligations.

Capital enables infrastructure development. Commercial performance sustains its financial architecture.

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Item 1, Business. Infrastructure financing model and customer contracts.

Item 1A, Risk Factors. Customer concentration and equipment replacement exposure.

Item 7, Management's Discussion and Analysis. Liquidity and capital resources.

Item 7A, Market Risk. Interest-rate sensitivity.

Notes to Consolidated Financial Statements. Revenue recognition and remaining performance obligations.

Note 10, Debt. Delayed-draw facilities and contractual financing terms.

Editorial Note

Company-specific financial figures reflect CoreWeave's FY2025 disclosures. The discussion of financing arrangements distinguishes reported contractual obligations from analytical interpretation.

Infrastructure financing structures vary according to asset ownership and contractual terms. Investment returns depend on commercial performance throughout the relevant financing and asset horizons.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 03 · From Announced Capacity to Operational Compute

AI infrastructure capacity progresses through contractual commitments, physical development and commercial operation. This chapter examines how grid availability and commissioning schedules influence capital productivity.

The economics of AI infrastructure begin long before the first GPU enters service.

Capital commitments establish the investment. Commercial agreements establish future demand. Physical infrastructure determines when computing capacity becomes available. Financial returns emerge through productive operation.

These activities follow different schedules.

A data-center developer may secure customer commitments years ahead of commissioning. Equipment procurement can begin while electrical infrastructure remains under development. Financing obligations may accumulate throughout construction.

The interval between capital commitment and operational readiness is therefore a material component of infrastructure economics.

SENTIENT REVIEW

The Economic States of Infrastructure Capacity

AI infrastructure advances through distinct commercial and physical conditions.

Announced capacity represents a proposed development. The announcement establishes intended scale and investment ambition.

Contracted capacity reflects rights or obligations established through commercial agreements. These may concern electricity supply or computing services.

Capacity under construction represents physical development supported by committed expenditure.

Energized capacity has access to electrical supply under defined operating conditions.

Commissioned capacity has completed the technical acceptance requirements for its intended operation.

Revenue-producing capacity supports commercial services under customer agreements.

Economically productive capacity generates sufficient financial contribution to support operating requirements and capital recovery.

These categories establish an analytical progression. Contractual and physical development can advance concurrently.

Their financial significance lies in the evidence each provides.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A power agreement establishes a development position. Commissioning demonstrates operational capability. Commercial service delivery establishes a basis for revenue recognition.

Investment returns depend on the contribution generated over the asset's economic life.

The Conversion of Contracted Power

CoreWeave's 2025 financial disclosures illustrate the scale of infrastructure development.

At December 31, 2025, the company operated 43 data centers with more than 850 megawatts of active power.

Its total contracted power capacity reached approximately 3.1 gigawatts.

The figures represent distinct operating conditions.

Active power describes the company's existing infrastructure footprint. Contracted power includes capacity expected to enter service during future periods.

The contracted figure is equivalent to approximately 3,100 megawatts.

This establishes a significant development pipeline relative to the reported active footprint. Its conversion into operational capacity requires additional infrastructure investment and successful deployment.

The financial implications extend beyond construction expenditure.

New facilities require commissioning. Installed equipment must become operational. Commercial commitments must translate into service delivery.

CoreWeave's annual filing identifies substantial future capital requirements associated with its expansion.

The investment assessment must therefore consider the timing of capacity deployment alongside the financing required to support it.

A complete project-level conversion analysis would require comparable capacity measurements across development stages and reporting dates.

Power Availability as a Commercial Constraint

The proposed Nscale development in Loughton, Essex, provides a documented example of the relationship between electricity infrastructure and commercial commitments.

The facility was intended to support Microsoft's Azure services through approximately 50 megawatts of computing capacity.

Plans included 23,040 NVIDIA GB300 GPUs, with operations originally anticipated in early 2027.

In October 2026, Reuters reported that Microsoft withdrew from the agreement to anchor the facility after grid connection delays pushed expected electricity availability into the 2030s.

Microsoft described the decision as specific to the Loughton site and reaffirmed its broader relationship with Nscale.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Nscale continued exploring alternative power arrangements and locations.

The development illustrates a critical feature of infrastructure investment.

Commercial demand can be established ahead of the physical resources required to satisfy it.

A delay in grid availability changes the expected delivery schedule. That change can affect capital deployment and future revenue timing.

The financial consequences depend on expenditure already incurred and the contractual allocation of development risk.

The publicly reported case establishes the disruption to planned delivery. Quantification of project-specific financial exposure requires further contractual evidence.

The Economics of Grid Access

Power availability influences both infrastructure location and capital productivity.

The International Energy Agency's Energy and AI research examined grid congestion and connection timelines across planned data-center developments.

Its analysis estimated that grid constraints could delay approximately 20% of global data-center capacity planned for construction by 2030.

This is a modeled exposure based on infrastructure conditions and development assumptions.

The estimate highlights the economic importance of electrical infrastructure readiness.

Connection schedules influence commissioning. Grid capacity determines the scale of equipment that can operate at a location. Electrical infrastructure investment affects project expenditure.

Location decisions therefore influence the relationship between capital commitments and productive capacity.

Alternative power arrangements may improve development flexibility. Their economics depend on reliability requirements and the investment needed to support continuous computing operations.

The relevant assessment connects power availability with commercial delivery and capital recovery.

The Financial Value of Time

Infrastructure development delays affect the timing of expected financial contribution.

Consider a hypothetical facility expected to generate $300 million in annual operating contribution after commissioning.

Under a uniform monthly contribution assumption, a six-month delay shifts approximately $150 million of expected contribution beyond the original operating schedule.

The calculation illustrates deferred contribution.

Its effect on investment value depends on the revised cash-flow schedule.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Construction costs may continue during the delay. Financing obligations can create additional carrying expenses. Commercial agreements determine the treatment of delivery changes.

A project may recover deferred operating contribution during an extended commercial horizon. A fixed contractual period can produce a different financial outcome.

Discounted cash-flow analysis captures the value of timing through the present value of expected future cash flows.

The relevant financial measure is the change in project value across the complete investment horizon.

From Commissioning to Economic Productivity

Commissioning establishes operational readiness.

Commercial productivity follows from the services delivered through that capacity.

Revenue recognition depends on customer agreements. Actual compute utilization follows workload demand and deployment characteristics.

Reserved-capacity arrangements can establish payment obligations across contractual periods. Consumption-based services create a closer relationship between metered activity and revenue.

Financial productivity introduces the cost of delivering those services and the capital supporting the infrastructure.

An operational facility can generate commercial revenue while continuing to require substantial investment for expansion or equipment replacement.

The financial assessment therefore extends across the productive life of the assets.

Investment Judgment

AI infrastructure investment requires coordination between commercial commitments and physical delivery.

Power availability establishes a critical development condition. Commissioning creates operational capacity. Customer demand supports commercial utilization.

The financial outcome depends on the timing and contribution of that capacity relative to invested capital.

The economic value of AI infrastructure capacity emerges through physical deliverability, commercial productivity and disciplined capital recovery.

The transition from planned megawatts to productive computing assets is ultimately a question of investment economics.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Infrastructure footprint, active power and contracted capacity. Future capital requirements.

Microsoft pulls out of UK data centre deal with Nscale

Reuters

Independent Reporting

October 2026

https://www.reuters.com/business/media-telecom/microsoft-pulls-out-uk-data-center-deal-with-nscale-telegraph-reports-2026-10-18/

Loughton development, planned computing capacity and electricity-grid delays.

Energy and AI

International Energy Agency

Institutional Research

2025

https://www.iea.org/reports/energy-and-ai/ai-and-energy-security

Section: AI and Energy Security.

Grid connection constraints and modeled capacity-delay exposure.

SENTIENT REVIEW

Editorial Note

CoreWeave figures reflect its infrastructure position at December 31, 2025. The Microsoft–Nscale case reflects publicly reported developments in October 2026.

The IEA estimate represents modeled exposure. The financial sensitivity calculation is hypothetical and illustrates the effect of commissioning delays on contribution timing.

Project-specific investment returns require verified financial and contractual evidence.

Contents

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 04 · Power as an Economic Constraint

Electricity availability governs the physical and financial development of AI infrastructure. This chapter examines grid constraints, power density, facility efficiency and their effects on capital productivity.

Electricity has become a governing variable in the economics of artificial intelligence infrastructure.

The availability of computing equipment determines potential processing capacity. Electrical infrastructure determines where that capacity can operate and when it becomes commercially productive.

Power availability influences site selection. Connection schedules shape commissioning dates. Electricity prices affect operating expenditure throughout the life of an asset.

The financial consequences extend from initial capital allocation to long-term infrastructure returns.

The Electricity Requirements of AI Infrastructure

AI data centers operate as integrated electrical and computing systems.

Accelerators perform computational work. Networking equipment supports communication between systems. Cooling infrastructure manages the heat generated during operation.

Electrical distribution systems sustain these functions under demanding reliability requirements.

The International Energy Agency estimated global data-center electricity consumption at approximately 415 terawatt-hours in 2024. This represented around 1.5% of global electricity consumption.

Its April 2025 Energy and AI report projected consumption reaching approximately 945 terawatt-hours by 2030 under its Base Case.

The IEA's updated 2026 assessment places estimated 2025 consumption at 485 terawatt-hours and projects approximately 950 terawatt-hours by 2030.

These estimates describe the broader data-center industry, including AI and conventional computing workloads.

The projections establish the scale of anticipated electricity demand. Their economic significance emerges through the infrastructure required to supply that demand.

Grid Access and Capital Deployment

Data-center development and electricity-grid expansion follow different investment schedules.

SENTIENT PREVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A computing facility can be constructed within a few years. Transmission infrastructure may require longer periods for planning and delivery.

Connection availability therefore influences when completed facilities can enter commercial service.

The IEA's 2025 analysis estimated that grid constraints could delay approximately 20% of global data-center capacity planned for construction by 2030.

This estimate represents modeled exposure based on regional infrastructure conditions.

For investors, the critical variable is the relationship between capital deployment and the date of operational readiness.

Construction expenditure can accumulate while a facility awaits electrical connection. Equipment procurement may establish additional financial commitments. Financing costs can continue throughout the development period.

A reliable commissioning schedule therefore contributes directly to capital efficiency.

Grid access is both an engineering requirement and an investment consideration.

Power Density and Infrastructure Architecture

Accelerated computing changes the electrical characteristics of data-center facilities.

High-performance computing systems concentrate substantial processing capability within individual racks. Greater density increases the importance of electrical distribution and thermal management.

The IEA identifies accelerated server deployment as a principal driver of future data-center electricity demand.

Infrastructure architecture must accommodate the power requirements of the intended computing configuration.

Electrical capacity determines the equipment a facility can support. Cooling capability influences operating conditions and equipment reliability.

The resulting investment extends beyond accelerator acquisition.

Power distribution and cooling systems form part of the productive infrastructure asset. Their design affects capital expenditure and operating efficiency.

Architecture decisions therefore influence both the cost of establishing capacity and the cost of sustaining it.

Measuring Energy Efficiency

Power Usage Effectiveness, or PUE, provides a standard measure of data-center infrastructure energy efficiency.

The US Department of Energy defines PUE as the ratio of total facility energy consumption to IT equipment energy consumption over the same measurement period.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A PUE of 1.20 means the facility consumes 1.20 units of energy for every unit consumed by its IT equipment.

The additional energy supports the infrastructure required to operate the computing environment.

PUE helps quantify the energy associated with facility operation.

Its economic significance depends on electricity pricing and the scale of IT consumption.

The metric measures facility energy efficiency. Workload productivity requires separate measures of useful computational output.

The Financial Effect of Energy Efficiency

Consider a hypothetical data center operating at an average IT load of 10 megawatts throughout a full year.

At a PUE of 1.40, annual facility electricity consumption reaches 122,640 megawatt-hours.

At a PUE of 1.20, consumption falls to 105,120 megawatt-hours.

The difference is 17,520 megawatt-hours annually.

At an illustrative electricity price of $100 per megawatt-hour, the annual expenditure difference is $1.752 million.

The calculation assumes constant average IT load and a uniform electricity price.

Actual financial outcomes depend on measured energy consumption and contractual tariffs. Infrastructure investments required to improve efficiency also affect the economic return.

The example establishes the relationship between facility design and recurring operating expenditure.

Energy efficiency becomes financially meaningful when the value of reduced consumption is evaluated against the investment required to achieve it.

SENTIENT REVIEW

Reliability and Commercial Performance

Electricity availability also determines the continuity of computing services.

Infrastructure operators must support customer workloads under defined service requirements. Electrical interruptions can affect availability and commercial performance.

Backup systems and resilient distribution architectures support operational continuity.

These capabilities require capital investment and ongoing maintenance.

Their economic value depends on the reliability requirements of the workloads being served and the financial consequences of service disruption.

Training systems and latency-sensitive inference services can carry different operating requirements.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The appropriate infrastructure design therefore follows the commercial obligations and workload characteristics of the facility.

Reliability investment becomes part of the economics of service delivery.

Geography and Infrastructure Concentration

Electricity infrastructure also influences the geographic distribution of AI computing capacity.

Regions with available grid capacity can offer development advantages. Electricity pricing shapes operating economics. Local infrastructure conditions influence delivery schedules.

The IEA identifies concentrated data-center development as a source of pressure on regional electricity systems.

Geographic concentration can support access to established infrastructure and technical expertise.

It can also create exposure to local power constraints.

Investment decisions must consider the relationship between regional capacity availability and the commercial demand expected from the facility.

The most attractive location depends on the complete economic requirements of the project.

Investment Judgment

Power economics connect physical infrastructure with financial productivity.

Electricity availability determines the feasibility and timing of deployment. Facility architecture influences capital requirements. Energy efficiency affects recurring expenditure.

Reliability supports commercial service delivery.

The investment assessment must consider these relationships over the operating life of the infrastructure.

SENTIENT REVIEW

Electricity is a determinant of AI infrastructure capacity, operating economics and capital productivity.

The financial quality of an AI infrastructure investment depends partly on securing dependable power at a cost and schedule consistent with its commercial objectives.

Sources

Energy and AI

International Energy Agency

Institutional Research

April 10, 2025

https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai

Global data-center electricity consumption, infrastructure development and demand projections.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Energy and AI

International Energy Agency

Institutional Research

2025

https://www.iea.org/reports/energy-and-ai/ai-and-energy-security

Grid connection constraints and modeled capacity-delay exposure.

Section: AI and Energy Security.

Key Questions on Energy and AI

International Energy Agency

Institutional Research

2026

https://www.iea.org/reports/key-questions-on-energy-and-ai/executive-summary

Executive Summary.

Updated data-center electricity demand estimates and projections.

Best Practices Guide for Energy-Efficient Data Center Design

US Department of Energy

Technical Publication

July 2024

https://www.energy.gov/sites/default/files/2024-07/best-practice-guide-data-center-design.pdf

Section 8: Benchmarking and Power Usage Effectiveness.

SENTIENT REVIEW

Editorial Note

IEA consumption figures distinguish historical estimates from projected demand. The 2025 and 2026 publications represent separate analytical vintages.

The PUE calculation uses hypothetical operating assumptions. Financial outcomes depend on facility characteristics and electricity tariffs.

Economic interpretations connect the cited technical evidence with infrastructure investment considerations.

Contents

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 05 · The Economics of GPU Architecture

GPU architecture determines the computing capabilities supported by infrastructure investment. This chapter examines memory systems, accelerator interconnects and workload performance as determinants of capital productivity.

The economics of accelerated computing begin with architecture.

A GPU represents more than computational throughput. Its memory system determines the data available during execution. Interconnect architecture governs communication across accelerators. System design determines how these capabilities operate together.

The financial significance emerges when architectural performance is translated into productive computing capacity.

For infrastructure investors, accelerator selection is a capital allocation decision. For cloud operators, it influences the cost and commercial capacity of the computing estate.

The relevant economic measure is the productive work delivered by the complete system relative to its capital and operating requirements.

SENTIENT REVIEW

The Architecture of Computational Capacity

AI workloads impose different demands on computing systems.

Model training requires substantial mathematical processing and communication across distributed accelerators. Inference introduces workload-specific requirements for memory access and response latency.

The architecture of the computing system influences performance in both environments.

Computational throughput measures the processing capability available for mathematical operations. Memory capacity determines how much data can reside close to the processor. Memory bandwidth governs the rate at which data moves between memory and computation.

These resources interact.

A workload constrained by memory bandwidth may gain limited benefit from additional arithmetic throughput. A model requiring more memory than an accelerator provides may need distribution across devices.

The resulting system configuration influences infrastructure cost and execution efficiency.

Architectural evaluation therefore begins with the workload and the resources required to execute it.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The Economics of High-Bandwidth Memory

High-bandwidth memory has become a defining component of AI accelerator design.

NVIDIA's HGX reference architecture documents the following specifications.

Accelerator	GPU memory	Memory bandwidth
H100 SXM	80 GB	3.35 TB/s
H200 SXM	141 GB	4.8 TB/s
B200 SXM	180 GB	Up to 8 TB/s

These figures describe manufacturer-specified hardware capabilities.

The B200 provides 2.25 times the memory capacity of the H100 SXM. Its specified maximum memory bandwidth is approximately 2.39 times that of the H100.

These ratios establish architectural differences.

Their economic significance depends on the model and execution configuration.

Greater memory capacity can support larger model states or additional inference requests within the available memory budget. Higher bandwidth can improve throughput for workloads constrained by memory access.

The resulting commercial benefit depends on realized performance under production conditions.

A financial comparison must therefore connect hardware specifications with useful workload output.

SENTIENT REVIEW

Interconnect and Distributed Computing

Modern AI systems depend on communication between accelerators.

Large training workloads distribute computation across multiple devices. Certain inference configurations also require model execution across several GPUs.

Communication introduces an architectural dependency.

NVIDIA's HGX H100 and H200 configurations provide up to 900 GB/s of GPU-to-GPU NVLink bandwidth. The HGX B200 configuration increases this specification to 1.8 TB/s.

The figures describe NVIDIA's stated interconnect capabilities.

Higher communication bandwidth can improve performance when data exchange constrains distributed execution.

The benefit depends on the communication requirements of the workload and the efficiency of its implementation.

System-level performance also depends on the network connecting computing nodes.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A facility's productive capacity therefore reflects the interaction between accelerators and the communication infrastructure supporting them.

The System as the Investment Unit

Infrastructure procurement increasingly involves integrated computing systems.

An eight-GPU HGX platform combines accelerator resources through a shared high-speed interconnect architecture.

The complete installation also requires host computing resources and system memory. Networking equipment connects the node to the wider computing environment.

These components establish the system's productive capability.

Their costs belong within the economic assessment of the installation.

A comparison based solely on accelerator specifications can overlook differences in the supporting system required to deliver a workload.

The appropriate investment boundary depends on the decision being evaluated.

An accelerator purchase may be compared at the component level. A cloud infrastructure investment requires a broader view of the computing system and its operating environment.

The financial denominator must reflect the assets required to produce the measured computational output.

Performance Under Commercial Conditions

Benchmarking provides evidence of computing performance under specified conditions.

MLCommons maintains the MLPerf Inference benchmark suite. Its methodology evaluates systems using defined models and workloads under documented performance requirements.

The September 2025 MLPerf Inference v5.1 release included results across multiple inference workloads.

The benchmark framework establishes a structured basis for evaluating performance.

Commercial interpretation requires attention to the tested configuration.

A result achieved under one model and request pattern represents performance within that measurement environment. Different workload characteristics can change the effective throughput of the same infrastructure.

Latency requirements are particularly important for interactive inference.

Higher throughput can improve the productive capacity of an installation when the system continues to satisfy its service requirements.

For an infrastructure operator, the economic benefit emerges through the relationship between workload performance and the resources consumed.

The relevant comparison connects measured output with a consistent cost boundary.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Architecture and Workload Placement

Computing infrastructure becomes more economically useful when workloads are assigned to systems suited to their requirements.

Large distributed workloads may benefit from substantial interconnect capacity. Memory-intensive inference can benefit from accelerators with greater high-bandwidth memory.

Other workloads may achieve attractive economics on earlier hardware generations.

This creates an opportunity for workload placement to influence asset productivity.

The decision depends on the relationship between performance requirements and effective execution cost.

A newer accelerator can provide superior technical performance while carrying a different acquisition cost. Existing equipment can retain economic usefulness where its performance satisfies commercial requirements.

The resulting infrastructure strategy may involve several accelerator generations serving different workloads.

This is an architectural decision with direct financial consequences.

From Hardware Performance to Capital Productivity

Accelerator performance establishes the productive potential of a computing asset.

Economic productivity requires a further calculation.

The relevant cost includes the computing system required to execute the workload. Its operating expenses must be measured across the appropriate period.

Useful output must be defined consistently.

For inference, the measurement may involve completed requests or generated tokens under specified service requirements.

For training, it may involve completed training work under defined quality and performance conditions.

These measures support comparisons between computing configurations.

A defensible investment decision evaluates the cost of delivering equivalent useful work.

Hardware specifications establish technical capability. Workload benchmarks provide performance evidence. Financial analysis determines the economic attractiveness of the resulting system.

Investment Judgment

GPU architecture influences the capital required to establish computing capacity and the productivity of that investment.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Memory architecture shapes the workloads an accelerator can accommodate. Interconnect performance influences distributed execution. System integration determines the resources available for commercial operation.

The financial assessment must connect these capabilities to actual workload requirements.

The economic value of GPU architecture emerges through the productive computation delivered by the complete system.

Accelerator selection becomes a capital allocation decision when performance is evaluated against the investment required to sustain it.

Sources

NVIDIA HGX AI Factory Reference Architecture

NVIDIA

https://docs.nvidia.com/enterprise-reference-architectures/hgx-ai-factory-h100-h200-b200/latest/components.html

Hardware Components: H100, H200 and B200.

Memory specifications and GPU interconnect architecture.

NVIDIA HGX Platform

NVIDIA

https://www.nvidia.com/en-us/data-center/hgx/

Product specifications.

System configurations and NVLink capabilities.

MLPerf Inference v5.1 Benchmark Results

MLCommons

September 9, 2025

https://mlcommons.org/2025/09/mlperf-inference-v5-1-results/

Benchmark methodology and workload-specific performance evaluation.

MLPerf Inference: Datacenter

MLCommons

https://mlcommons.org/benchmarks/inference-datacenter/

Benchmark documentation.

Performance measurement conditions and benchmark definitions.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Editorial Note

Hardware specifications reflect manufacturer documentation for the identified accelerator configurations. Actual workload performance depends on the system configuration and execution environment.

The memory-capacity and bandwidth ratios are calculated from NVIDIA's published specifications. They represent architectural comparisons rather than measured application-performance improvements.

Economic interpretations connect the documented technical characteristics with infrastructure investment considerations. Detailed workload pricing and inference optimization receive dedicated treatment in Part II of The Economics of AI series.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 06 · The GPU as a Depreciating Financial Asset

GPU infrastructure combines technological capability with depreciating capital assets. This chapter examines useful-life assumptions, asset consumption and hardware replacement economics as determinants of investment returns.

The financial life of an AI accelerator begins with capital acquisition. Its economic life is determined by the work it can perform profitably.

Between these two events lies one of the central investment questions in AI infrastructure: how long can computing equipment generate returns sufficient to recover its acquisition cost and support its eventual replacement?

GPU performance advances through successive architectures. Commercial workloads evolve alongside it. Infrastructure owners must reconcile these changes with depreciation policies and financing obligations.

The resulting asset economics influence earnings and cash generation throughout the investment horizon.

SENTIENT REVIEW

The GPU as a Capital Asset

An AI accelerator enters an infrastructure business as part of a broader computing investment.

Its acquisition creates capital expenditure or an associated financing obligation. Installation places the equipment within an operating system. Commissioning establishes its readiness for productive use.

Accounting treatment follows the asset's intended service life.

Under straight-line depreciation, the depreciable cost is allocated across the estimated useful life of the equipment.

The estimate reflects management's expectations about productive use and asset longevity.

Commercial performance follows another measure: the contribution generated by the computing services delivered through that equipment.

These perspectives serve different financial purposes.

Depreciation allocates historical asset cost to accounting periods. Economic productivity depends on future operating contribution and capital recovery.

What Infrastructure Operators Disclose

CoreWeave's 2025 Form 10-K establishes a six-year estimated useful life for technology equipment.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The company adopted this estimate in January 2023, extending the previous five-year life for computing equipment used in its data centers.

Management attributed the revision to developments in hardware performance and software optimization. Improvements in data-center design also informed the decision.

The accounting effect was measurable.

CoreWeave reported that the revised estimate reduced total expenses by approximately $20 million during 2023.

The adjustment changed the timing of depreciation expense. The historical acquisition cost of the equipment remained unchanged.

Alphabet provides a comparable accounting reference.

Its 2025 Form 10-K reports that servers and network equipment are generally depreciated over six years. The company reviews useful-life estimates using expected technology developments and planned infrastructure utilization.

These disclosures establish how major infrastructure owners account for computing assets.

They also reveal the judgment embedded in useful-life assumptions.

The Financial Sensitivity of Useful Life

Consider a hypothetical computing installation with a depreciable cost of $600 million and zero residual value.

Under straight-line depreciation:

Estimated useful life	Annual depreciation
Four years	$150 million
Five years	$120 million
Six years	$100 million

A six-year estimate produces $50 million less annual depreciation than a four-year estimate during their overlapping depreciation periods.

The initial investment remains $600 million.

This distinction matters when interpreting operating earnings.

Depreciation affects reported profitability. Capital expenditure affects cash flow according to the timing of payment. The economic value of the asset depends on the cash it generates throughout its productive life.

Useful-life assumptions therefore influence accounting comparisons between infrastructure operators.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A longer accounting life can improve reported earnings during earlier periods. Its economic justification depends on the expected productive use of the underlying assets.

The Scale of Asset Consumption

CoreWeave's FY2025 financial statements illustrate the significance of depreciation in a rapidly expanding infrastructure business.

The company reported approximately $2.4 billion in FY2025 property and equipment depreciation and amortization.

Its FY2025 technology and infrastructure expense included approximately $2.3 billion in depreciation and amortization, compared with $843 million in the corresponding FY2024 technology and infrastructure expense category.

These figures use different reporting boundaries and should retain their respective definitions.

CoreWeave also reported $20.9 billion in gross technology equipment and $9.4 billion in construction in progress at December 31, 2025.

The figures reveal the scale of assets supporting the company's commercial operations and future deployment.

They also demonstrate why asset accounting is material to infrastructure profitability.

The annual cost of capital consumption becomes a significant component of the operating model as equipment enters service.

Economic Obsolescence and Productive Life

Accelerator technology advances through improvements in computational throughput and memory architecture. Interconnect design also influences the performance of distributed computing systems.

Each generation changes the available technical capabilities.

The economic consequences depend on workload requirements.

A newer accelerator may execute a workload with greater efficiency. Earlier equipment may retain commercial value through applications suited to its performance characteristics.

Software improvements can extend the useful life of installed hardware.

Workload placement therefore influences asset longevity.

An infrastructure operator may continue generating contribution from an earlier GPU generation while deploying newer systems for workloads with more demanding requirements.

The investment decision depends on the contribution available from each asset relative to its remaining operating costs and future capital requirements.

Commercial useful life emerges from this relationship between performance and demand.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Replacement Economics

Hardware replacement introduces a new capital allocation decision.

An operator can continue using an existing installation or invest in a newer computing system.

The financial comparison should examine the incremental economics of each option.

Continued operation preserves the opportunity to earn contribution from existing assets. Replacement introduces new capital expenditure alongside potential gains in performance and operating efficiency.

The relevant analysis includes the timing of replacement and expected residual value.

Financing obligations associated with the existing equipment remain part of the broader financial position.

A replacement decision may therefore be attractive operationally while creating additional demands on capital.

The investment judgment must account for the productive contribution expected from both alternatives.

Accounting Life and Investment Returns

The economic return on computing infrastructure depends on cash generation over the relevant investment horizon.

Depreciation contributes to accounting earnings and may affect taxable income. It represents the allocation of previously capitalized expenditure.

Cash-flow analysis captures acquisition expenditure and subsequent operating receipts. Replacement investments enter the analysis when expected to occur.

Return on invested capital introduces the relationship between after-tax operating earnings and the capital employed.

These measures answer related financial questions through different accounting boundaries.

Their interpretation requires consistent definitions and measurement periods.

A six-year depreciation policy establishes an accounting estimate. The investment assessment must establish whether the equipment can generate sufficient economic contribution over the relevant horizon.

Investment Judgment

The GPU is both a computing resource and a depreciating financial asset.

Its architecture determines productive capability. Workload demand influences commercial usefulness. Depreciation policy governs the recognition of historical asset cost.

Replacement decisions shape future capital requirements.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The financial quality of the investment depends on the contribution generated throughout the asset's economic life.

The value of AI computing equipment rests on its ability to convert technical performance into durable financial contribution.

The most productive infrastructure investment is determined by the relationship between asset cost and the useful work delivered over time.

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Relevant disclosures:

Note 1. Property and Equipment, Net. Estimated useful lives and depreciation policy.

Note 5. Property and Equipment, Net. Asset balances and depreciation.

Management's Discussion and Analysis. Technology and infrastructure expenses.

Impairment of Long-Lived Assets. Recoverability and asset valuation.

Primary source:

2025 Form 10-K

Alphabet Inc.

Company Filing

February 5, 2026

https://www.sec.gov/Archives/edgar/data/1652044/000165204426000018/goog-20251231.htm

Relevant disclosures:

Property and Equipment. Depreciation policy and useful-life estimates.

Critical Accounting Estimates. Technical infrastructure useful-life assessment.

Primary source:

Editorial Note

Company-specific figures reflect the reported accounting periods and classifications. The depreciation example uses hypothetical asset values and a zero-residual-value assumption.

Accounting useful life and commercial productive life are evaluated as distinct concepts. The analysis of replacement economics represents financial interpretation of documented asset characteristics and accounting practices.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 07 · Capacity Utilization and Commercial Productivity

AI infrastructure productivity depends on the relationship between available computing capacity, commercial utilization and financial contribution. This chapter examines how contract structures and workload demand shape infrastructure economics.

Installed computing capacity represents an investment in productive potential. Its financial performance depends on the commercial services delivered through that capacity and the contribution those services generate.

The distinction matters because AI infrastructure operates through several forms of utilization.

An accelerator can be installed and available for service. It can execute computational workloads. Its capacity can be commercially committed to a customer. Each condition describes a different aspect of productivity.

Infrastructure economics emerge from the relationship between these measures.

For investors, the relevant consideration is the financial contribution generated by computing assets throughout their productive lives.

The Economic Meaning of Utilization

Capacity utilization begins with a defined resource.

Installed capacity measures the computing equipment deployed within an infrastructure environment.

Available capacity represents resources ready to serve workloads under the relevant operating conditions.

Computational utilization measures processing activity against an explicitly defined hardware or system capacity.

Commercial utilization concerns the capacity committed or consumed under customer agreements.

Financial productivity measures the resulting economic contribution relative to the resources required to deliver it.

Each measure serves a different purpose.

Computational utilization helps operators understand resource consumption. Commercial utilization supports revenue planning. Financial productivity connects infrastructure operation with investment returns.

Their interpretation depends on consistent measurement periods and clearly defined denominators.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A percentage describing GPU processing activity has a different economic meaning from the percentage of capacity committed under customer contracts.

Commercial Capacity and Contract Structure

CoreWeave's FY2025 Form 10-K provides a useful example of the relationship between capacity and revenue.

The company serves customers through committed contracts and on-demand computing services.

Committed arrangements typically operate on a take-or-pay basis. Customers purchase access to specified computing capacity over an agreed period.

Revenue is generally recognized over the contract term as the associated services are provided.

On-demand arrangements follow a consumption model. Customers are billed according to actual usage of computing services.

These structures produce different relationships between workload activity and recognized revenue.

Under a committed-capacity agreement, contractual payments can continue as the customer's computational demand varies.

Under consumption-based pricing, commercial revenue follows metered usage more directly.

The contract determines how infrastructure availability becomes a revenue-generating service.

Approximately 98% of FY2025 revenue on a rounded reported basis originated from committed customer arrangements.

This concentration of revenue within committed arrangements illustrates the importance of contractual economics to infrastructure productivity.

The disclosure establishes the company's revenue composition. Workload-level utilization requires separate operating evidence.

Physical Scale and Commercial Deployment

CoreWeave reported more than 850 megawatts of active power across 43 data centers at December 31, 2025.

Its contracted power capacity reached approximately 3.1 gigawatts.

These figures establish the scale of the company's operating footprint and development pipeline.

Power capacity measures electrical infrastructure. GPU utilization measures computational activity.

The relationship between them depends on the equipment installed and the facility's operating characteristics.

An infrastructure investment therefore requires several layers of measurement.

Physical capacity establishes the resources available for deployment. Installed computing equipment determines technical capability. Commercial agreements establish expected demand.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Actual operating performance determines the resources consumed in delivering services.

Financial performance follows the revenue and cost structure of those services.

The distinction becomes particularly important during periods of rapid infrastructure expansion.

Revenue per Billable GPU-Hour

A consumption-based infrastructure model can be examined through billable computing hours.

Consider a hypothetical installation containing 1,000 GPUs.

Assume each GPU has 8,760 potential service hours annually. The installation therefore provides 8.76 million theoretical GPU-hours per year.

At a realized rate of $3 per billable GPU-hour, annual revenue depends on the proportion of theoretical hours that become billable.

Billable utilization	Annual billable hours	Annual revenue
40%	3,504,000	$10.512M
60%	5,256,000	$15.768M
80%	7,008,000	$21.024M

These calculations assume constant pricing and uniform service availability throughout the year.

The example represents a hypothetical consumption-based commercial model.

Reserved-capacity agreements require analysis of the contractual payment schedule and the services delivered.

The revenue calculation establishes the commercial sensitivity to billable utilization. Profitability requires the corresponding cost evidence.

The Cost of Maintaining Capacity

Computing infrastructure incurs costs throughout its operating life.

Facility expenses support the physical environment. Electricity powers the computing equipment and its supporting systems. Technical operations sustain availability and service delivery.

Depreciation allocates the historical cost of qualifying assets across their estimated useful lives.

The economic significance of these costs depends on their behavior as workload activity changes.

Some costs remain relatively stable over the relevant operating range. Others vary with electricity consumption or service demand.

An increase in billable utilization can improve financial contribution when incremental revenue exceeds the additional cost of serving the workload.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The magnitude of that improvement depends on the realized commercial price and the operating cost structure.

Infrastructure profitability therefore requires an analysis of contribution rather than revenue alone.

Workload Productivity and Service Quality

The productive capacity of a GPU installation depends on the workloads it executes.

Training systems often emphasize sustained computational processing across distributed accelerators.

Inference systems can experience variable request patterns and latency requirements.

The same hardware may exhibit different operating characteristics under these workloads.

Computational throughput provides one measure of technical performance. Service quality establishes whether the delivered computation satisfies the workload's requirements.

A system serving interactive inference must maintain performance within the applicable response-time conditions.

Batch workloads may permit different scheduling and throughput arrangements.

These characteristics influence the amount of useful computing work delivered by installed capacity.

Workload placement becomes an economic decision when alternative configurations offer different performance and cost profiles.

The following chapter examines these differences between training and inference.

From Utilization to Financial Contribution

Infrastructure utilization becomes financially meaningful when connected to the economics of service delivery.

Three measures are particularly useful.

Realized revenue establishes the commercial receipts attributable to the service under the applicable accounting treatment.

Operating contribution measures revenue after the costs assigned to delivering that service.

Capital productivity evaluates operating returns against the invested resources supporting them.

These measures establish progressively broader views of infrastructure performance.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

A highly active computing installation may support demanding workloads with substantial operating costs.

A capacity reservation may generate predictable revenue while the customer's actual computational activity varies.

The financial significance of either arrangement depends on the commercial contract and the resources required to fulfill it.

The investment assessment must preserve these distinctions.

Investment Judgment

AI infrastructure productivity emerges through the interaction between computing capacity and commercial demand.

Installed assets establish technical potential. Workload execution consumes computing resources. Customer contracts determine the revenue associated with service delivery.

Operating costs and capital requirements determine the financial contribution available to investors.

The relevant economic assessment connects these conditions over a consistent measurement period.

The financial productivity of AI infrastructure is established through commercially valuable computation and the contribution generated by the capital employed.

Utilization provides operating evidence. Contract economics and financial performance establish its investment significance.

SENTIENT REVIEW

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Relevant disclosures:

Item 1. Business. Infrastructure capacity and commercial services.

Item 1A. Risk Factors. Committed contracts and on-demand revenue exposure.

Revenue Recognition. Committed and on-demand service arrangements.

Note 2. Revenue. Contractual revenue composition and recognition policies.

Primary source:

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Editorial Note

CoreWeave's financial and capacity figures reflect its FY2025 disclosures. Commercial contract descriptions follow the company's reported revenue recognition policies.

The GPU-hour analysis is hypothetical. Its utilization assumptions and pricing are illustrative rather than observed company performance.

Financial interpretations distinguish physical capacity from computational activity and commercial revenue. Workload profitability requires additional cost evidence.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 08 · Training, Inference and Workload Economics

Training and inference create distinct demands for AI infrastructure. This chapter examines workload performance, service economics and computing-resource allocation as determinants of capital productivity.

The economics of AI computing are determined by the work performed on the infrastructure.

Training creates model capability. Inference converts that capability into computational outputs. Each activity places different demands on the underlying systems and creates distinct conditions for capital productivity.

Infrastructure owners must understand these differences when allocating computing capacity. Model developers face decisions about training investment and deployment efficiency. Enterprise buyers encounter the resulting economics through the price and performance of AI services.

The governing relationship is between useful computational work and the resources required to deliver it.

Two Distinct Computing Economies

Training develops or adapts model parameters through repeated computational operations.

Large training runs can occupy accelerator clusters for extended periods. Their performance depends on sustained processing efficiency and communication between distributed computing resources.

Inference applies trained models to new inputs.

Its operating profile depends on the application. Interactive services require timely responses. Batch processing permits different scheduling arrangements. Reasoning workloads may consume additional computation as models generate intermediate steps.

These differences influence infrastructure requirements.

Training economics concern the cost of producing or improving a model capability. Inference economics concern the recurring cost of using that capability.

For infrastructure operators, both activities create commercial demand. Their contribution depends on realized pricing and the cost of supplying the required computing resources.

Training as a Capital Allocation Decision

Training expenditure represents an investment in model capability.

The computing requirement depends on model architecture and training methodology. Dataset characteristics influence the work performed. Distributed execution introduces communication and synchronization requirements.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The economic analysis begins with the total resources consumed during training.

Accelerator hours establish one measure of computing expenditure. Networking and storage contribute additional costs. Engineering effort and experimental iterations influence the broader investment.

The resulting model capability becomes economically valuable through subsequent use or commercialization.

For a model developer, training expenditure may support future product revenue or internal operating benefits. For a cloud provider, the same workload can generate infrastructure-service revenue during execution.

The distinction reflects their different commercial positions.

Training efficiency matters because it influences the resources required to achieve a defined model result.

A meaningful comparison must preserve the intended capability and evaluation conditions.

Inference as a Recurring Operating Economy

Inference creates a continuing relationship between computational demand and service delivery.

Each request consumes computing resources according to its characteristics.

Input length influences the processing required before generation. Output length affects the work performed during response generation. Model size and architecture determine memory and computational requirements.

Application design introduces further variability.

An interactive assistant may prioritize response latency. A document-processing service may prioritize aggregate throughput. An analytical application may require extended reasoning and larger context windows.

These workloads create different economic demands on the same infrastructure.

Inference efficiency is therefore measured against the service being delivered.

The appropriate denominator may be a completed request or a defined quantity of generated output. Quality and latency requirements establish the conditions under which that output is useful.

The commercial significance emerges when workload performance is connected to realized revenue and delivery cost.

Throughput and Latency

Throughput measures computational output over time. Latency measures the time required to deliver a response or complete an operation.

The relationship between them influences infrastructure economics.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Batching can improve the efficiency of accelerator use by processing requests together. Larger batches may increase aggregate throughput while affecting response times.

Interactive workloads impose limits on this trade-off.

A service designed for rapid responses may require additional available capacity to satisfy demand during peak periods.

That capacity carries financial cost even when demand varies.

NVIDIA's inference documentation describes batching and serving optimizations as methods for balancing throughput with latency requirements.

The economic decision concerns the cost of meeting a defined service standard.

An efficient infrastructure configuration delivers the required workload performance with an appropriate allocation of computing resources.

What Independent Benchmarks Establish

MLCommons provides a structured basis for evaluating inference performance.

Its September 2025 MLPerf Inference v5.1 release introduced benchmarks for DeepSeek-R1 and Llama 3.1 8B. Whisper Large V3 expanded coverage of speech recognition.

The release also extended interactive testing under tighter latency conditions.

These additions reflect the diversity of production inference requirements.

MLPerf Inference v6.1, released in September 2026, introduced End-to-End Retrieval-Augmented Generation and Edge Agentic Inference benchmarks. The release expanded evaluation beyond individual model execution to include multi-component inference pipelines and agentic workloads.

Benchmark results provide technical evidence under documented conditions.

They can inform hardware procurement and workload placement decisions when the tested configurations are relevant to the intended application.

Financial interpretation requires additional evidence about equipment costs and operating expenditure.

The benchmark establishes measured performance. Investment analysis connects that performance to the resources required to produce it.

The Economics of Workload Placement

Infrastructure estates increasingly contain several generations of computing equipment.

Each generation offers a different relationship between performance and cost.

A workload requiring substantial memory or distributed computation may benefit from newer accelerator architectures.

Other workloads can remain economically attractive on established hardware.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

This creates an allocation problem.

Infrastructure operators must determine which computing resources can deliver each workload at an appropriate cost while satisfying its service requirements.

The decision affects asset utilization and equipment replacement economics.

Assigning suitable workloads to existing equipment can extend its productive life. Deploying newer systems for demanding workloads can improve capacity where architectural advantages are material.

The financial result depends on the effective cost of useful computation.

Workload placement therefore becomes part of infrastructure capital management.

Measuring Workload Economics

A useful economic measure requires a consistent definition of output and cost.

Consider a hypothetical inference service processing two million accepted requests during a month.

Assume the fully allocated monthly delivery cost is $80,000.

The resulting cost per accepted request is four cents.

This calculation is illustrative. It assumes all two million requests meet the defined acceptance requirements and that the cost allocation covers the relevant service-delivery resources.

The measure becomes commercially useful when compared with realized revenue per request.

For an infrastructure provider, the analysis may instead focus on revenue and contribution per contracted computing unit.

For an enterprise customer, the economic assessment extends to the business activity supported by the inference service.

These perspectives belong to different positions within the value chain.

The full economics of inference pricing and model routing receive dedicated treatment in Part II of The Economics of AI.

Investment Judgment

Training and inference create different forms of demand for AI infrastructure.

Training supports the development of model capability. Inference converts that capability into recurring computational services.

Their economics depend on workload characteristics and service requirements.

Infrastructure architecture determines available computing performance. Scheduling and workload placement influence productive capacity. Commercial arrangements establish revenue opportunities.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The investment assessment connects these elements through the cost and contribution of useful computation.

The financial productivity of AI compute emerges from the relationship between workload performance, service economics and the capital required to deliver them.

The quality of infrastructure investment ultimately depends on how effectively computing capacity serves commercially valuable workloads.

Sources

MLPerf Inference v5.1 Benchmark Results

MLCommons

September 9, 2025

https://mlcommons.org/2025/09/mlperf-inference-v5-1-results/

Benchmark workloads, interactive inference and performance measurement.

MLPerf Inference v6.1 Benchmark Results

MLCommons

September 2026

https://mlcommons.org/2026/09/mlperf-inference-v6-1-results/

End-to-End Retrieval-Augmented Generation and Edge Agentic Inference benchmarks, multi-component inference pipelines and agentic workloads.

MLPerf Inference Reference Implementations and Rules

MLCommons

https://github.com/mlcommons/inference

What Is AI Inference?

NVIDIA

https://www.nvidia.com/en-us/glossary/ai-inference/

Inference serving architecture and throughput-latency considerations.

Inference Performance for Data Center Deep Learning

NVIDIA Developer

https://developer.nvidia.com/deep-learning-performance-training-inference/ai-inference

Workload-specific performance measurements and configuration disclosures.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Editorial Note

MLCommons benchmark results reflect defined test conditions and published submission methodologies. NVIDIA technical material provides manufacturer documentation and performance evidence.

The cost-per-request calculation is hypothetical. Its assumptions are explicit and its results represent an illustrative unit-economic relationship.

The analysis distinguishes measured technical performance from commercial workload profitability. Detailed model-serving optimization and inference pricing belong to Part II of the series.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 09 · The Fully Loaded Cost of AI Compute

AI infrastructure profitability depends on the complete cost of delivering computing services. This chapter examines facility expenditure, technology depreciation, operating margins and workload-level cost allocation.

AI infrastructure profitability begins with the cost of delivering computational capacity.

The purchase of accelerators establishes an investment. Facility operations create recurring expenditure. Technology assets contribute depreciation throughout their accounting lives. Engineering and software sustain the computing environment.

Each cost enters the financial statements according to its economic character and accounting treatment.

For investors, the relevant question concerns the resources required to generate a dollar of compute revenue and the financial contribution remaining after those resources are accounted for.

The Economic Boundary of Compute

A computing service draws upon an integrated operating system.

Accelerators execute workloads. Networking equipment supports communication. Electrical infrastructure supplies power. Cooling systems maintain operating conditions.

Software and engineering sustain the platform through which customers consume these resources.

The financial boundary must reflect the service being evaluated.

At the facility level, analysis may focus on energy consumption and operating expenditure.

At the infrastructure-platform level, the cost boundary expands to include equipment depreciation and technical operations.

At the enterprise level, the assessment incorporates broader operating expenses and capital requirements.

The selected boundary determines the meaning of the resulting margin.

What CoreWeave's Financial Statements Reveal

CoreWeave's fiscal 2025 results provide a documented illustration of infrastructure cost classification.

The company reported $5.131 billion in revenue and $1.453 billion in cost of revenue.

Its technology and infrastructure expense reached $2.929 billion.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Operating expenses also included $144 million in sales and marketing and $651 million in general and administrative expenditure.

The resulting operating loss was $46 million.

These amounts establish the company's reported operating economics for the year ended December 31, 2025.

FY2025 financial measure	USD millions
Revenue	5,131
Cost of revenue	1,453
Technology and infrastructure	2,929
Sales and marketing	144
General and administrative	651
Operating loss	(46)

The figures demonstrate how the financial interpretation changes as the cost boundary expands.

Facility Costs and Infrastructure Operations

CoreWeave classifies data-center facility expenses within cost of revenue.

These include rent and utilities. Depreciation associated with power installation and distribution systems also enters this category.

Personnel supporting data-center operations and customer success contribute further expense.

Technology and infrastructure costs follow a separate classification.

This category includes depreciation associated with computing equipment and networking systems. Technical personnel and software expenditure also contribute.

The separation reflects the company's reported accounting policies.

During 2025, cost of revenue increased by $960 million compared with 2024.

Management attributed part of that increase to approximately $566 million in additional rent expense and $203 million in additional utilities expenditure.

Depreciation associated with power installation and distribution systems increased by approximately $84 million.

These are reported year-over-year expense increases. They illustrate the operating consequences of infrastructure expansion.

The Margin Boundary

CoreWeave's reported revenue and cost of revenue produce gross profit of $3.678 billion.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The corresponding gross margin is approximately 71.7%.

This measure reflects the company's reported cost-of-revenue classification.

A broader analytical calculation incorporates technology and infrastructure expense.

Revenue of $5.131 billion less cost of revenue of $1.453 billion and technology and infrastructure expense of $2.929 billion leaves $749 million.

That represents approximately 14.6% of revenue.

Analytical measure	FY2025
Gross profit	$3.678B
Gross margin	71.7%
Revenue after cost of revenue and technology infrastructure expense	$749M
Corresponding analytical margin	14.6%
Operating margin	−0.9%

The 14.6% figure is an analytical subtotal. Its purpose is to illustrate the effect of expanding the expense boundary.

Sales and marketing expenditure and general administrative costs remain outside that subtotal.

The company's full operating statement produces the reported $46 million operating loss.

Margin interpretation depends on the expenses included in its calculation.

SENTIENT REVIEW

Depreciation and Capital Consumption

Computing equipment represents a significant capital asset.

CoreWeave reported approximately $2.3 billion in depreciation and amortization within technology and infrastructure expense during 2025, compared with $843 million in 2024.

The increase reflects the expansion of the company's computing infrastructure.

Depreciation allocates qualifying asset costs across their estimated useful lives.

The corresponding capital expenditure follows the timing of acquisition and payment.

These measures serve different financial purposes.

Depreciation contributes to accounting profitability. Capital expenditure affects investment cash flows. Asset replacement introduces future capital requirements.

A comprehensive infrastructure assessment must consider all three.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The resulting financial analysis should connect current operating performance with the investment required to sustain future computing capacity.

From Infrastructure Cost to Workload Cost

Company-level financial statements establish aggregate operating economics.

Workload economics require a more specific allocation.

A GPU-hour represents a defined unit of accelerator availability or consumption. Its cost depends on the resources assigned to that service.

For a consumption-based model, the calculation may allocate relevant annual costs across billable GPU-hours.

A reserved-capacity contract requires treatment consistent with the contracted service and associated revenue.

The denominator must reflect the commercial arrangement.

Consider a hypothetical computing service generating six million billable GPU-hours annually.

Assume the relevant annual cost pool totals $18 million.

The allocated cost is $3 per billable GPU-hour.

At a realized commercial rate of $4 per hour, the contribution before expenses outside the selected cost boundary is $1 per billable hour.

This calculation assumes a consistent service definition and complete allocation of the stated cost pool.

The example illustrates how cost boundaries influence unit economics.

SENTIENT REVIEW

The Financial Cost of Capacity

Infrastructure providers invest in computing resources ahead of commercial consumption.

Capacity availability carries costs through the operating life of the assets.

Electricity expenditure varies with operating conditions. Equipment depreciation follows accounting schedules. Facility obligations depend on contractual arrangements.

Commercial utilization influences how these costs are absorbed across revenue-producing services.

A higher volume of billable computing activity can improve cost absorption when additional revenue exceeds incremental service costs.

The resulting financial contribution depends on realized pricing and the resources required to fulfill demand.

The economic assessment must therefore connect cost structure with contract terms and workload characteristics.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Investment Judgment

AI infrastructure profitability is determined by the relationship between commercial revenue and the full economic resources required to deliver computing services.

Facility expenses establish one operating boundary. Technology and infrastructure costs establish another. Corporate operating expenses complete the reported operating picture.

Capital expenditure and financing introduce further investment considerations.

The relevant measure depends on the decision being made.

The economics of compute become meaningful when the cost boundary reflects the assets and operating resources required to deliver the service.

Financial productivity emerges through sustainable contribution and disciplined capital recovery.

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Relevant disclosures:

Consolidated Statements of Operations and Comprehensive Loss.

Management's Discussion and Analysis: Components of Results of Operations.

Cost of Revenue.

Technology and Infrastructure.

Property and Equipment.

Primary source:

Editorial Note

CoreWeave financial figures reflect the company's FY2025 reporting classifications.

The 14.6% analytical margin uses an expanded expense boundary and represents an independently calculated subtotal. The GPU-hour example uses hypothetical inputs.

Company-level financial statements provide evidence of aggregate operating performance. Workload-specific profitability requires measured resource consumption and attributable cost data.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 10 · Revenue Quality and Infrastructure Contracts

AI infrastructure contracts establish future commercial demand and influence financing capacity. This chapter examines revenue recognition, customer concentration, payment timing and contractual economics as determinants of financial value.

AI infrastructure contracts have become financial instruments as well as commercial agreements.

Long-term compute commitments establish future demand. They can support equipment financing and provide visibility into revenue. Their economic value depends on delivery obligations and the financial resources required to fulfill them.

For infrastructure investors, the quality of contracted revenue is as important as its scale.

The central consideration is the relationship between commercial commitments and sustainable financial contribution.

SENTIENT REVIEW

The Commercial Architecture of Compute

Infrastructure providers generate revenue through different contractual arrangements.

Under committed-capacity agreements, customers secure access to computing resources for an agreed period. Payments generally follow the contracted capacity and service terms.

Consumption-based arrangements establish a closer relationship between actual usage and billing.

CoreWeave operates both models.

Its FY2025 Form 10-K describes committed contracts with initial terms generally ranging from one to six years. These arrangements typically use take-or-pay provisions requiring payment regardless of actual utilization.

On-demand services are billed according to consumption.

The distinction determines how operating capacity becomes recognized revenue.

A committed contract can support predictable revenue recognition while computational demand varies. Consumption-based revenue responds more directly to workload activity.

Both models create economic opportunities. Their risk characteristics differ.

The Scale of Contracted Demand

At December 31, 2025, CoreWeave reported $60.7 billion in remaining performance obligations, or RPO.

Its committed contracts had a weighted-average duration of approximately five years.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

RPO represents the transaction price allocated to outstanding contractual performance obligations, after accounting for estimated variable consideration.

The balance includes billed and unbilled amounts associated with committed contracts.

CoreWeave expects 43% of its year-end RPO to be recognized during the 24 months ending December 31, 2027.

A further 38% is expected during months 25 through 48. The remaining balance is expected during months 49 through 84.

These disclosures establish a substantial future revenue schedule.

They also create an important financial distinction between contracted demand and recognized operating performance.

The economic value of the commitments depends on the services delivered and the associated cost structure.

Revenue Recognition and Cash Timing

CoreWeave accounts for its cloud computing services under ASC 606.

Committed-contract revenue is generally recognized ratably across the service period.

On-demand revenue is recognized as customers consume services.

Customer prepayments introduce a separate cash-flow dimension.

Deferred revenue represents consideration received or billed ahead of the associated revenue recognition. Revenue is subsequently recognized as the contractual performance obligations are satisfied.

CoreWeave reported $8.2 billion in deferred revenue at December 31, 2025, compared with $4.1 billion a year earlier.

Remaining performance obligations represent the transaction price allocated to outstanding contractual performance obligations. Deferred revenue represents consideration received or billed ahead of the associated revenue recognition. These measures have different accounting boundaries.

The increase reflected growth in committed contracts that typically include prepayments.

This creates a financing advantage through earlier cash receipts.

It also establishes future service obligations.

The financial assessment must therefore examine cash timing alongside revenue recognition and the cost of fulfilling contracted capacity.

Revenue Concentration

Customer concentration influences the financial quality of infrastructure contracts.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

CoreWeave reported that Microsoft represented approximately 67% of its recognized revenue during 2025.

Its other individual customers each accounted for less than 10%.

The concentration reflects the importance of large customers within specialized AI infrastructure.

An anchor customer can provide substantial demand visibility. Long-term commitments may support equipment financing and capacity development.

The same commercial structure creates dependence on the durability of the customer relationship.

Contract renewal terms influence future revenue. Customer credit quality affects payment expectations. The customer's infrastructure strategy can shape demand beyond the original agreement.

CoreWeave disclosed an additional agreement under which OpenAI committed to payments of up to approximately $6.5 billion through May 2031.

That commitment provides evidence of future commercial demand. Its realized contribution will depend on service delivery and the applicable contractual terms.

Revenue concentration therefore requires assessment alongside contract duration and counterparty quality.

Contract Quality and Infrastructure Financing

Long-term customer agreements can strengthen the financing capacity of an infrastructure operator.

CoreWeave describes asset-level financing supported by computing equipment and contractual cash flows.

These arrangements connect future customer payments with present infrastructure investment.

The quality of that relationship depends on several conditions.

The customer must retain the financial capacity to meet its obligations. The operator must deliver the contracted service. The financing structure must accommodate the timing of receipts and required expenditure.

Contractual provisions can allocate certain risks between counterparties.

CoreWeave's filing also identifies potential customer credits and delivery delays within its treatment of variable consideration.

These adjustments influence the calculation of remaining performance obligations.

The economic analysis must therefore examine contractual terms alongside reported backlog.

From Revenue Visibility to Financial Contribution

Contracted revenue establishes an expectation of future commercial activity.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Financial contribution emerges after the resources required to deliver the service are accounted for.

Consider a hypothetical infrastructure contract providing $120 million in annual recognized revenue.

Assume annual attributable service-delivery costs of $85 million.

The resulting contribution is $35 million before expenses outside the selected cost boundary.

The contribution margin is approximately 29.2%.

This calculation illustrates the relationship between revenue quality and cost-to-serve.

A longer contract may improve revenue visibility. Its economic attractiveness depends on the costs incurred throughout the service period.

Changes in electricity prices can affect contribution. Hardware replacement may introduce additional capital requirements. Financing obligations influence cash available to investors.

The contract's financial value therefore extends beyond its stated revenue amount.

Commercial Durability

AI infrastructure contracts operate within a changing technology market.

New accelerator generations influence the performance available to customers. Workload demand evolves with model architecture and application adoption.

These developments can affect future pricing and renewal economics.

CoreWeave identifies customer retention and infrastructure upgrades as material considerations in its annual filing.

An operator may benefit from committed demand during the original contract period while facing different commercial conditions at renewal.

The investment assessment should therefore distinguish contracted revenue visibility from the expected economics of subsequent operating periods.

Long-term value depends on the contribution generated across the productive life of the assets.

Investment Judgment

Infrastructure contracts establish the commercial foundation for capacity investment.

Their duration influences revenue visibility. Payment terms affect liquidity. Customer concentration shapes counterparty exposure.

Financial contribution depends on the operating resources and capital required to deliver the contracted services.

The investment judgment must connect contractual commitments with their economic fulfillment.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The quality of AI infrastructure revenue rests on the durability of customer commitments and the financial contribution generated through their delivery.

Contracted demand supports investment. Sustainable operating economics establish its value.

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Relevant disclosures:

Item 1. Business. Committed capacity and take-or-pay contracts.

Item 1A. Risk Factors. Customer concentration and contractual exposure.

Revenue Recognition. ASC 606 policies and customer prepayments.

Note 2. Revenue. Deferred revenue and remaining performance obligations.

Management's Discussion and Analysis. Commercial commitments and financing.

Primary source:

Editorial Note

CoreWeave figures reflect its FY2025 disclosures. Remaining performance obligations represent management's reported contractual revenue schedule under applicable accounting requirements.

The contribution-margin example is hypothetical. Its cost boundary excludes corporate expenses and financing obligations.

The financial analysis distinguishes contracted demand from recognized revenue and realized cash generation.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 11 · Cash Flow and Return on Invested Capital

AI infrastructure cash flows reveal the relationship between operating performance and capital investment. This chapter examines free cash flow, financing obligations and return on invested capital as measures of financial productivity.

AI infrastructure investment creates a financial relationship between operating cash generation and capital consumption.

Computing assets require substantial expenditure before reaching productive service. Customer contracts generate receipts according to commercial terms. Equipment financing introduces repayment obligations across the asset's operating life.

The investment return emerges from the timing and scale of these cash flows.

CoreWeave's 2025 financial statements illustrate the relationship. The company generated positive operating cash flow while deploying substantial capital into infrastructure expansion. Its financing activities supplied additional liquidity.

These results establish a useful distinction between operating cash generation and the capital required to sustain growth.

SENTIENT REVIEW

The Cash Flow Architecture of Infrastructure

Infrastructure financial statements separate cash flows into operating activities, investing activities and financing activities.

Operating cash flow reflects cash generated or consumed through business operations. Customer payments influence this measure alongside working-capital movements.

Investing cash flow captures expenditures on productive assets and other investing transactions.

Financing cash flow records capital raised from investors and lenders, together with repayments and related financing transactions.

Each category describes a different financial activity.

Their relationship becomes especially important during periods of infrastructure expansion.

An operator may collect substantial customer prepayments while simultaneously investing in computing equipment. Debt proceeds can finance the resulting expenditure.

The consolidated cash-flow statement reveals how these activities interact.

CoreWeave's 2025 Cash Flow

CoreWeave reported the following cash flows for the year ended December 31, 2025.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Financial measure	FY2025
Operating cash flow	$3.058B
Net investing cash flow	−$10.271B
Net financing cash flow	$9.308B
Cash purchases of property and equipment	$10.309B

The figures reveal the financial scale of infrastructure expansion.

Operating activities generated $3.058 billion. Investing activities consumed $10.271 billion on a net basis. Financing activities supplied $9.308 billion.

Management attributed the increase in operating cash flow partly to upfront payments associated with new and expanded customer contracts.

Customer prepayments provide liquidity before the associated services are fully delivered. The corresponding revenue recognition follows the company's contractual performance obligations.

The timing of customer receipts therefore influences reported operating cash generation.

For investors, this distinction is material to understanding the sustainability of future cash flows.

Free Cash Flow and Capital Intensity

Operating cash flow less cash purchases of property and equipment.

Using CoreWeave's 2025 disclosures:

$3.058B − $10.309B = −$7.251B

The calculation produces operating cash flow less cash purchases of property and equipment of −$7.251 billion: $3.058 billion in operating cash flow less $10.309 billion in cash purchases of property and equipment.

This measure reflects cash capital expenditure exceeding operating cash generation during the reporting year.

CoreWeave's expansion also involves equipment financing and other transactions that affect the timing of cash payments.

The company disclosed $11.151 billion in liabilities associated with property and equipment additions, including OEM-financed additions, within its supplemental financing and investing disclosures.

This figure describes a separate accounting category from cash purchases of property and equipment. Financed equipment additions require separate analysis and are not included in this cash-capital-expenditure calculation.

A comprehensive capital-investment analysis must reconcile cash expenditure with financed asset additions and related obligations.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The economic interpretation depends on the investment boundary being measured.

Capital Investment and Operating Earnings

Capital expenditure and depreciation affect financial performance through different mechanisms.

Capital expenditure records the acquisition or development of assets. Depreciation allocates qualifying asset costs across their estimated accounting lives.

Operating earnings reflect depreciation according to the company's expense classifications.

Cash-flow analysis records capital expenditure according to payment timing.

These distinctions become material when an infrastructure operator is expanding its computing estate.

A new facility may require expenditure during construction. Its depreciation begins when qualifying assets become ready for their intended use.

Commercial contribution develops through service delivery.

The timing of these events shapes both accounting earnings and investment cash flows.

Financial analysis must preserve their respective measurement periods.

The Cost of Infrastructure Financing

Debt financing supports infrastructure deployment while introducing contractual repayment obligations.

CoreWeave reported $21.615 billion in debt principal outstanding at December 31, 2025.

Its scheduled principal payments included $6.708 billion during 2026 and $4.298 billion during 2027.

The company also reported $1.229 billion in net interest expense for 2025.

These figures demonstrate the financial significance of debt obligations within the infrastructure operating model.

Interest expense influences earnings after financing costs. Principal repayment affects cash available to equity investors.

Refinancing arrangements can alter the timing of these obligations.

An investment assessment therefore requires a distinction between operating performance and financing cash flows.

Return on Invested Capital

Return on invested capital connects operating profitability with the resources employed to generate it.

A conventional formulation uses after-tax operating profit divided by average invested capital.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

The calculation requires consistent definitions.

Operating profit must reflect the relevant business activities. Tax treatment must correspond to the chosen operating-profit measure.

Invested capital requires an explicit treatment of operating assets and associated liabilities.

For infrastructure businesses, construction in progress and leased assets can materially affect the denominator.

A defensible calculation must also preserve the distinction between capital committed and capital already deployed.

The resulting measure supports an assessment of operating returns relative to the capital supporting the business.

Cash Returns and Investment Valuation

A project-level investment assessment uses dated cash flows.

Construction expenditure establishes the initial investment profile. Operating receipts and expenditure determine subsequent contribution.

Hardware replacement introduces future capital requirements. Residual asset value influences the final investment period.

Net present value discounts expected future cash flows using a rate appropriate to the investment and its risk.

An unlevered project analysis evaluates operating investment economics before debt financing cash flows.

An equity analysis incorporates the financing structure and cash flows attributable to equity holders.

These perspectives serve different investment decisions.

Scenario analysis can examine commissioning timing, utilization and financing sensitivity through explicitly defined assumptions. Project-specific valuation requires dated cash flows, verified inputs and an appropriately supported discount rate.

Company-specific investment conclusions require evidence supporting the relevant project cash flows and discount rates.

Investment Judgment

AI infrastructure can generate positive operating cash flow while consuming substantial capital during expansion.

The financial outcome depends on the contribution generated by deployed assets and the investment required to sustain them.

Customer prepayments influence liquidity. Financing arrangements determine repayment obligations. Depreciation affects reported earnings.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Cash-flow analysis connects these elements across the investment horizon.

Capital productivity is established through sustainable operating returns relative to the full resources committed to productive infrastructure.

The investment judgment ultimately concerns the financial returns generated throughout the economic life of the assets.

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Relevant disclosures:

Consolidated Statements of Cash Flows.

Management's Discussion and Analysis: Liquidity and Capital Resources.

Capital Investments and Financing Activities.

Note 10: Debt.

Supplemental Noncash Investing and Financing Disclosures.

Primary source:

SENTIENT REVIEW

Editorial Note

CoreWeave figures reflect its FY2025 financial disclosures. The free-cash-flow calculation uses operating cash flow less cash purchases of property and equipment.

Financed equipment additions represent a separate disclosure category. Investment-return methodology is presented as analytical interpretation.

Project-level valuation requires verified cash-flow assumptions and a justified discount rate.

Contents

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 12 · Infrastructure Risk and Economic Sensitivity

AI infrastructure investment carries interacting delivery, financing and commercial exposures. This chapter examines how commissioning delays, customer concentration and hardware replacement influence capital recovery and investment returns.

AI infrastructure investment concentrates substantial capital in assets whose financial performance depends on physical delivery and commercial demand.

The investment case develops across several timelines. Power availability determines commissioning. Financing obligations establish capital requirements. Customer contracts influence revenue visibility. Hardware evolution shapes the productive life of computing equipment.

These variables interact throughout the investment horizon.

The economic significance of infrastructure risk emerges through its effect on cash generation and capital recovery.

The Architecture of Infrastructure Risk

Infrastructure exposure begins during development and continues throughout commercial operation.

Physical delivery determines when computing assets can enter service. Financing conditions influence the cost of capital. Commercial contracts establish expected revenue. Equipment performance affects the cost of delivering computing services.

Each variable has a distinct financial mechanism.

A commissioning delay shifts the timing of expected contribution. Higher borrowing costs affect financing cash flows. Weaker utilization can reduce revenue under consumption-based contracts. Hardware replacement creates additional capital requirements.

The interaction between these exposures determines the resilience of the investment.

An infrastructure assessment therefore requires a consistent financial horizon and explicit assumptions about how operating conditions change.

Power Availability and Delivery Exposure

Electrical infrastructure has become a critical determinant of data-center development schedules.

The International Energy Agency's 2025 Energy and AI research estimated that grid constraints could delay approximately 20% of global data-center capacity planned for construction by 2030.

The estimate represents modeled exposure based on regional grid conditions and development pipelines.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

For infrastructure investors, the economic concern is the interval between capital deployment and productive operation.

Construction expenditure may continue while electrical connection remains pending. Equipment commitments can create further financial obligations. Financing arrangements influence the carrying cost of deployed capital.

The timing of commissioning therefore affects the present value of expected returns.

Delivery risk becomes financially material when it changes the investment's cash-flow schedule.

Financing and Refinancing Exposure

CoreWeave's FY2025 Form 10-K illustrates the financing requirements of specialized AI infrastructure.

At December 31, 2025, the company reported approximately $21.615 billion in outstanding debt principal.

Its scheduled principal repayments included $6.708 billion during 2026 and $4.298 billion during 2027.

These amounts represent contractual maturities as reported at year-end. Subsequent refinancing or repayment activity can change the schedule.

CoreWeave also disclosed approximately $107 million in annual interest-expense sensitivity for a 100-basis-point movement in interest rates.

The calculation reflects the company's stated debt exposure and assumptions. Interest-rate swaps mitigate exposure associated with certain floating-rate obligations.

These disclosures establish two financial considerations.

Debt maturities create requirements for repayment or refinancing. Interest-rate exposure influences the cost of outstanding borrowing.

Infrastructure cash generation must be evaluated alongside both.

Customer Concentration and Contract Quality

Commercial commitments can support the financing of computing infrastructure.

Their economic value depends on contractual performance and the financial capacity of the counterparties.

CoreWeave reported that Microsoft accounted for approximately 67% of its recognized revenue during 2025.

The concentration illustrates the importance of anchor customers within specialized compute infrastructure.

Long-term agreements can strengthen revenue visibility. Their payment provisions influence liquidity. Renewal terms shape future commercial performance.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Concentration also increases the sensitivity of the business to decisions made by major customers.

The financial assessment should examine contract duration and payment security. Capacity-delivery obligations require equal attention.

The investment return ultimately depends on the contribution generated through those commitments.

Hardware Longevity and Replacement Risk

Accelerator technology evolves through successive improvements in computing performance.

New architectures can improve throughput or reduce the resources required for particular workloads. Existing equipment may remain commercially productive through alternative workload assignments.

CoreWeave reports a six-year estimated useful life for technology equipment.

That accounting estimate establishes the period over which qualifying asset costs are depreciated.

Commercial competitiveness follows workload demand and realized service pricing.

The relationship between accounting life and economic productivity becomes important when evaluating replacement investment.

Earlier replacement increases capital requirements. Extended productive use can support additional contribution from installed assets.

The investment decision depends on the incremental cash flows associated with continued operation and replacement.

Measuring Financial Sensitivity

A useful sensitivity analysis connects changes in operating assumptions to their financial consequences.

Consider a hypothetical infrastructure project with $5 billion in floating-rate borrowing.

A 100-basis-point increase in the applicable interest rate would create approximately $50 million in additional annual interest expense.

The calculation assumes the entire principal remains outstanding and exposed for the year.

Now consider a separate hypothetical facility expected to generate $300 million in annual operating contribution after commissioning.

Under uniform monthly contribution, a six-month delay would shift approximately $150 million of expected contribution beyond the original operating schedule.

The final financial effect depends on subsequent operations and the revised cash-flow horizon.

These examples illustrate different mechanisms.

Interest-rate sensitivity affects financing expenditure. Commissioning delays affect the timing of operating contribution.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

An integrated model must preserve those distinctions.

Combined Economic Scenarios

Infrastructure risks can reinforce one another.

A delayed facility may require additional construction-period funding. Higher borrowing costs can increase financing expenditure during the extended development period.

Commercial utilization may develop more slowly than expected after commissioning. Hardware replacement can introduce further capital requirements before the original investment has achieved its anticipated recovery.

The combined effect belongs within a dated cash-flow analysis.

A central scenario represents the selected operating assumptions. A downside scenario tests adverse changes in delivery and commercial performance. An upside scenario examines stronger execution and financial productivity.

Each scenario requires explicit definitions.

Financial conclusions should reflect the modeled cash flows and their associated discount rates.

The resulting analysis provides a basis for evaluating capital recovery under alternative operating conditions.

Conditions Supporting Attractive Returns

Infrastructure risk also creates opportunities for disciplined investment.

Projects with secured power availability can achieve greater delivery certainty. Strong customer commitments may support financing on favorable terms.

Efficient infrastructure architecture can improve operating contribution. Workload placement may extend the productive life of existing computing equipment.

These advantages become financially meaningful when they improve expected cash generation relative to invested capital.

Their value depends on the actual investment and operating conditions.

The objective of sensitivity analysis is to establish how financial returns respond to those conditions.

Investment Judgment

AI infrastructure risk emerges from the interaction between physical delivery and financial obligations.

Power availability influences commissioning. Customer commitments establish commercial demand. Hardware economics shape asset longevity. Financing determines the cost and timing of capital obligations.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Investment returns depend on the combined effect of these variables throughout the operating horizon.

The financial resilience of AI infrastructure is established through deliverable capacity and durable commercial contribution relative to capital exposure.

Sensitivity analysis translates these operating conditions into investment judgment.

Sources

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Relevant disclosures: Item 1A, Risk Factors. Item 7A, Market Risk. Note 10, Debt. Property and Equipment accounting policies.

Energy and AI

International Energy Agency

April 2025

https://www.iea.org/reports/energy-and-ai/ai-and-energy-security

Section: AI and Energy Security.

Grid connection delays and modeled infrastructure delivery exposure.

SENTIENT REVIEW

Editorial Note

CoreWeave figures reflect its financial position at December 31, 2025. Debt maturity schedules and interest-rate sensitivity follow the company's disclosed assumptions.

The IEA capacity-delay estimate represents modeled exposure. The financing and commissioning examples are hypothetical.

The scenario analysis illustrates economic sensitivity. Project-specific investment returns require verified cash-flow inputs and appropriately supported discount rates.

Contents

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 13 · Where AI Infrastructure Value Accumulates

AI infrastructure generates financial returns across technology suppliers, infrastructure operators and cloud platforms. This chapter examines how commercial position and capital requirements influence profitability and value capture.

The AI infrastructure economy distributes capital and financial returns across businesses occupying different positions in the technology supply chain.

Semiconductor suppliers monetize computing technology. Infrastructure operators convert physical assets into commercial capacity. Cloud platforms combine computing services with software and customer distribution.

Each business captures a different portion of expenditure. Each carries a distinct financial structure.

The distribution of economic value depends on the relationship between commercial pricing and the capital required to sustain each position.

Three companies illustrate these differences: NVIDIA, CoreWeave and Microsoft.

Their financial disclosures provide evidence of how revenue and profitability emerge across the AI infrastructure economy.

The Economics of Technology Supply

NVIDIA occupies a central position in accelerated computing.

Its products combine processor architecture with memory systems and interconnect technology. Software platforms support the deployment and operation of accelerated workloads.

The company earns revenue through the commercialization of these technologies and associated computing systems.

For fiscal 2026, NVIDIA reported $215.9 billion in revenue, representing 65% annual growth.

GAAP gross margin reached 71.1%. Operating income totaled $130.4 billion.

Its Compute & Networking segment generated approximately $193.5 billion in revenue.

These results demonstrate the financial scale of NVIDIA's technology supply business.

The economic structure is important.

Equipment purchases create revenue for NVIDIA while establishing capital investment for infrastructure customers.

The same transaction occupies different positions in the financial statements of the two parties.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

NVIDIA's profitability reflects its commercial position and operating model. Infrastructure customers must subsequently generate returns from the equipment acquired.

This distinction establishes the first major point of value capture in the AI investment cycle.

The Economics of Infrastructure Ownership

CoreWeave represents a different financial model.

Its business converts computing infrastructure into commercial cloud services. Capital is committed to equipment and supporting facilities before the full commercial contribution develops.

The company reported $5.131 billion in revenue during 2025.

Its operating expenses totaled $5.177 billion, producing an operating loss of $46 million.

Net interest expense reached approximately $1.229 billion.

These figures reflect the capital and operating requirements of specialized infrastructure services.

CoreWeave also reported $3.058 billion in operating cash flow while cash purchases of property and equipment reached $10.309 billion.

The relationship between these measures reveals the demands of infrastructure expansion.

Commercial revenue grows through the delivery of computing services. Physical assets require substantial investment. Financing arrangements support deployment and introduce repayment obligations.

The operator's economic position therefore depends on the productivity of installed assets throughout their operating lives.

The Economics of Integrated Cloud Platforms

Microsoft combines infrastructure investment with a broader commercial platform.

Its Intelligent Cloud segment includes Azure and other enterprise infrastructure services. The business generates revenue through computing capacity and associated platform offerings.

For fiscal 2026, Intelligent Cloud reported $137.791 billion in revenue.

Segment operating income reached $56.972 billion.

The resulting operating margin was approximately 41.3%.

Microsoft's annual report identifies AI infrastructure investment as a contributor to cloud operating costs and margin pressure.

The segment's financial results reflect a diversified portfolio of cloud services.

Its reported profitability therefore provides evidence of integrated cloud-platform economics.

The distinction from a specialized infrastructure operator lies partly in the commercial activities supported by the underlying assets.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Computing infrastructure can serve multiple products within an integrated platform. Established distribution channels and software services create additional opportunities for monetization.

The financial contribution of AI infrastructure must be evaluated within that broader business structure.

Comparing Financial Positions

The three companies demonstrate different economic roles.

Financial measure	NVIDIA	CoreWeave	Microsoft Intelligent Cloud
Reporting period	FY2026	FY2025	FY2026
Revenue	$215.9B	$5.131B	$137.791B
Operating income	$130.4B	−$46M	$56.972B
Operating margin	60.4%	−0.9%	41.3%

The comparison uses reported GAAP figures and segment results.

NVIDIA's figures represent consolidated company performance. CoreWeave's results describe its consolidated infrastructure business. Microsoft's figures represent the Intelligent Cloud segment.

The reporting periods also differ.

These boundaries matter because each business generates earnings through a different combination of assets and commercial activities.

Operating margins reveal the profitability of the reported businesses. Capital productivity requires further analysis of invested capital and asset requirements.

The comparison illustrates financial differences across the supply chain while preserving the distinct economic structures underlying them.

The Movement of Capital and Revenue

AI infrastructure investment creates financial transactions between participants.

A cloud operator purchases computing systems. That expenditure generates revenue for equipment suppliers.

The operator subsequently commercializes computing services. Customer payments generate infrastructure revenue.

Enterprise customers consume those services to support applications and business processes.

Economic value can emerge at each stage.

The financial interpretation requires attention to transactions occurring between companies.

Supplier revenue and customer capital expenditure represent opposite sides of the same transaction.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Aggregating financial activity across the supply chain requires appropriate treatment of these relationships.

The relevant analytical objective is to establish where profits and cash generation arise after the resources required to produce them are accounted for.

The Distribution of Capital Risk

Financial exposure also differs across the supply chain.

Technology suppliers invest in product development and commercial delivery. Infrastructure operators commit capital to depreciating physical assets.

Integrated cloud platforms distribute investment across infrastructure and related services.

These positions create different financial sensitivities.

Hardware demand influences supplier revenue. Capacity utilization and contract terms influence infrastructure operators. Cloud-platform profitability depends on the economics of the services supported by its infrastructure.

Financing introduces additional exposure for asset-intensive operators.

Capital requirements and commercial performance therefore determine the distribution of investment risk alongside financial returns.

The economic position of each participant depends on the assets it owns and the obligations it carries.

Where Value Capture Becomes Durable

Financial performance provides evidence of value capture during a reporting period.

Durability requires examination across a longer investment horizon.

For technology suppliers, future performance depends on product demand and competitive position.

Infrastructure operators must sustain commercial contribution through equipment utilization and asset replacement.

Cloud platforms must generate returns across their service portfolios while continuing to invest in infrastructure.

The investment assessment connects operating performance with future capital requirements.

A business can demonstrate strong revenue growth while expanding its asset base. Financial value becomes durable when operating returns support the capital required to sustain the business.

The distinction is particularly important in an industry undergoing rapid technological change.

Investment Judgment

AI infrastructure creates economic opportunities across several commercial positions.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

NVIDIA demonstrates substantial profitability in technology supply. CoreWeave illustrates the capital requirements of specialized computing infrastructure. Microsoft demonstrates the earnings capacity of an integrated cloud platform.

Their financial results reveal how business models influence the distribution of returns.

The economic value of AI infrastructure accumulates according to commercial position and the capital required to sustain it.

Investment judgment requires an understanding of both.

The next chapter brings these relationships together through the economics of productive infrastructure and disciplined capital allocation.

Sources

Fiscal 2026 Annual Review

NVIDIA Corporation

https://s201.q4cdn.com/141608511/files/doc_financials/2026/ar/2026-annual-report-web-Hyperlinks.pdf

Fiscal year ended January 25, 2026.

Revenue, gross margin, operating income and Compute & Networking segment disclosures.

2025 Form 10-K

CoreWeave Inc.

Company Filing

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Fiscal year ended December 31, 2025.

Operating results, financing expenses and cash-flow disclosures.

2026 Annual Report

Microsoft Corporation

https://www.microsoft.com/investor/reports/ar26/

Fiscal year ended June 30, 2026.

Intelligent Cloud segment financial performance and infrastructure investment disclosures.

Editorial Note

Company financial figures reflect their respective fiscal periods and accounting boundaries.

The comparative operating margins are calculated from reported operating income and revenue. NVIDIA and CoreWeave are represented through consolidated results. Microsoft is represented through its Intelligent Cloud segment.

The analysis examines differences in business-model economics. Company-specific return on invested capital requires additional financial reconciliation.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Chapter 14 · The Economics of Productive AI Infrastructure

AI infrastructure investment creates economic value through the conversion of capital into productive computing capacity. This concluding chapter integrates infrastructure delivery, workload economics and financial returns into a unified investment perspective.

AI infrastructure investment is an exercise in converting capital into productive assets and those assets into sustainable financial returns.

The conversion begins with semiconductor technology and physical development. It progresses through electrical infrastructure and commissioning. Commercial workloads establish revenue opportunities. Operating performance determines contribution. Capital recovery establishes the financial result.

Each stage influences the economics of the investment.

The evidence examined throughout this publication reveals an infrastructure market shaped by extraordinary demand and substantial capital requirements. Its financial performance depends on the relationship between technical capability and productive commercial use.

The central investment judgment concerns the economic return generated across the complete life of the infrastructure.

SENTIENT REVIEW

Capital and the Productive Asset

Infrastructure investment begins with capital commitments.

Equipment procurement creates assets and payment obligations. Facility development establishes the physical environment. Electrical infrastructure determines the power available for computing operations.

These investments follow different development schedules.

Alphabet's 2025 financial disclosures demonstrate the scale of technical infrastructure investment. Its reported capital expenditure reached $91.4 billion, compared with $52.5 billion during 2024.

The figures represent company-wide capital expenditure. Their significance lies in the investment required to support future operating capacity.

For an infrastructure owner, capital productivity develops when installed assets generate commercial contribution.

The financial analysis must therefore connect expenditure with asset readiness and the operating activity supported by that investment.

Physical Delivery and Commercial Readiness

Computing capacity progresses through contractual commitments and physical development.

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Power agreements establish access to electrical resources. Construction creates the facility. Commissioning confirms operational capability.

Commercial service delivery introduces revenue under the applicable customer arrangements.

CoreWeave's December 2025 disclosures illustrate the distinction between development commitments and operating capacity.

The company reported approximately 3.1 gigawatts of contracted power and more than 850 megawatts of active power.

These measures describe different stages of infrastructure availability.

The contracted pipeline establishes potential future expansion. Active capacity reflects the company's reported operational footprint.

Investment analysis must examine the resources and time required to convert future capacity into commercial operation.

The economic importance of this conversion extends beyond construction schedules. Financing obligations and customer commitments create financial exposure throughout development.

The Economics of Compute Productivity

Operational infrastructure generates economic value through the services it delivers.

Accelerator architecture determines the computing resources available to workloads. Memory systems influence execution efficiency. Interconnect design shapes distributed processing performance.

Training and inference introduce different requirements.

Training emphasizes sustained computation and the development of model capability. Inference creates recurring service demand with workload-specific throughput and latency requirements.

The resulting infrastructure economics depend on the cost of delivering useful computational work.

Commercial contracts influence revenue realization. Consumption-based services connect revenue more directly to metered usage. Reserved-capacity agreements establish payment obligations according to contracted service availability.

These arrangements create different relationships between workload activity and financial contribution.

Productivity must therefore be evaluated within the relevant commercial model.

Asset Longevity and Capital Recovery

Computing equipment represents a depreciating capital asset.

Accounting useful life determines the allocation of historical asset cost. Commercial productive life depends on the equipment's ability to serve workloads economically.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

CoreWeave's six-year estimated useful life for technology equipment illustrates the accounting treatment.

Hardware evolution introduces a separate investment consideration.

New accelerator generations may improve performance for demanding workloads. Existing equipment can remain productive where its capabilities satisfy commercial requirements.

The economic assessment must examine continued operation and replacement through their incremental financial consequences.

Asset longevity influences capital recovery. Replacement investment affects future cash requirements.

The financial return emerges across the full investment horizon.

Revenue Quality and Operating Contribution

Infrastructure contracts establish the commercial foundation for capital deployment.

CoreWeave reported $60.7 billion in remaining performance obligations at December 31, 2025.

The figure represents contracted future revenue under the company's accounting definitions.

Its financial significance depends on the delivery of computing services and the costs associated with those obligations.

Customer concentration introduces another dimension.

Large customers can support infrastructure financing through substantial commitments. Their contractual terms influence future cash-flow visibility.

Operating contribution provides a further measure of commercial performance.

Revenue must be evaluated alongside the resources required to deliver the contracted service.

Facility costs support physical operations. Technology expenses include depreciation and technical resources. Corporate operating expenditure completes the broader financial picture.

The selected cost boundary determines the meaning of the resulting margin.

Cash Generation and Financial Productivity

Infrastructure investment requires a distinction between operating performance and capital expenditure.

CoreWeave's 2025 financial statements illustrate the relationship.

The company generated approximately $3.058 billion in operating cash flow while cash purchases of property and equipment reached $10.309 billion.

Operating cash flow less cash purchases of property and equipment therefore equaled −$7.251 billion: $3.058 billion less $10.309 billion.

The calculation reflects a period of substantial infrastructure expansion.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Customer prepayments contributed to operating liquidity. Financing arrangements supported capital deployment. Financed equipment additions require separate analysis and are not included in this cash-capital-expenditure calculation.

The resulting financial position must be interpreted across the investment cycle.

Operating cash generation establishes one measure of performance. Accounting depreciation allocates asset cost over its estimated useful life. Cash capital expenditure records cash purchases of property and equipment. Financed equipment additions and related obligations require separate analysis when evaluating total infrastructure investment.

Return on invested capital connects operating profitability with the capital employed.

Project valuation extends the analysis through dated cash flows and appropriately supported discount rates.

These measures establish different perspectives on infrastructure financial productivity.

Conditions Supporting Attractive Returns

AI infrastructure can generate attractive economic returns when commercial demand aligns with productive capacity.

Power availability supports timely commissioning. Efficient architecture improves the resources required for useful computation. Customer contracts establish revenue visibility.

Operating discipline influences contribution. Asset longevity affects replacement investment and capital recovery.

The financial result depends on the interaction between these conditions.

An infrastructure project with dependable power and durable demand may achieve a favorable operating profile.

A computing estate with suitable workload placement can extend the commercial productivity of installed equipment.

Long-term customer commitments can support financing when contractual receipts align with repayment obligations.

The economic attractiveness of each arrangement must be established through its own investment and operating evidence.

Capital Allocation Across the AI Economy

The AI infrastructure economy distributes financial returns across distinct business models.

Technology suppliers earn revenue through computing systems and related intellectual property.

Infrastructure operators monetize deployed capacity.

Integrated cloud platforms combine computing resources with broader commercial services.

Their capital requirements differ. Their reported financial results reflect those differences.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

NVIDIA's fiscal 2026 results demonstrate substantial profitability within accelerated computing technology supply.

CoreWeave's fiscal 2025 results illustrate the investment and financing requirements of specialized infrastructure services.

Microsoft's fiscal 2026 Intelligent Cloud results demonstrate the earnings capacity of a diversified cloud platform.

These comparisons establish the importance of commercial position.

Financial returns must be evaluated against the capital supporting each business model.

The distribution of value across the AI economy depends on where profitable commercial activity develops and how much investment is required to sustain it.

The Investment Judgment

AI infrastructure has become a central component of technology capital allocation.

Its economic evaluation requires a connected understanding of physical infrastructure and computing architecture.

Financial analysis must then establish the contribution generated by those assets throughout their productive lives.

The governing considerations are clear.

Capital investment establishes productive potential.

Physical delivery determines operational readiness.

Workload performance influences computing efficiency.

Commercial agreements establish revenue opportunities.

Operating contribution and capital recovery determine financial returns.

These relationships provide the foundation for disciplined infrastructure investment.

The economics of productive AI infrastructure are established through the conversion of capital into useful computation and sustainable financial contribution.

The quality of that conversion determines the returns available to infrastructure owners and the broader participants in the AI economy.

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

Sources

2025 Form 10-K

Alphabet Inc.

February 5, 2026

https://www.sec.gov/Archives/edgar/data/1652044/000165204426000018/goog-20251231.htm

Capital expenditure and technical infrastructure disclosures.

2025 Form 10-K

CoreWeave Inc.

March 2, 2026

https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

Infrastructure capacity, revenue obligations and cash-flow disclosures.

Fiscal 2026 Annual Report

NVIDIA Corporation

https://investor.nvidia.com/financial-info/annual-reports-and-proxies/default.aspx

Fiscal year ended January 25, 2026.

Accelerated computing business and financial performance.

2026 Annual Report

Microsoft Corporation

https://www.microsoft.com/investor/reports/ar26/

Fiscal year ended June 30, 2026.

Intelligent Cloud financial results.

Energy and AI

International Energy Agency

April 2025

https://www.iea.org/reports/energy-and-ai

Infrastructure development and electricity supply constraints.

Editorial Note

This chapter synthesizes the technical and financial evidence developed throughout the publication.

Company disclosures retain their respective accounting periods and reporting boundaries. The financial interpretation distinguishes capital expenditure from operating contribution and investment returns.

The analysis establishes principles for evaluating infrastructure economics. Company-specific investment decisions require verified asset-level evidence and appropriately supported financial assumptions.

Contents

SENTIENT REVIEW

© 2026 Sentient Review. All rights reserved. · Original Intellectual Property · sentient-review.com

SENTIENT REVIEW

Intellectual Property Notice

This publication and its original frameworks, models, structures, diagrams, terminology, and analytical materials are intellectual property of Sentient Review.

© 2026 Sentient Review. All rights reserved.

No part of this artifact may be reproduced, republished, distributed, adapted, incorporated into commercial materials, or represented as another party’s work without prior written permission, except as permitted by applicable law.

Published by Sentient Review

sentient-review.com

SENTIENT REVIEW