Graphics Processing Units (GPUs)

Information Technology

April 1, 2024

Industry Weighting — **NEUTRAL**

Investment Thesis

Artificial Intelligence (AI) is poised to usher in unprecedented innovation and productivity across the global economy. GPU sales and accompanying software are expected to sustain high growth over the next decade as the backbone of this technological revolution. Nevertheless, valuations matter, and we see limited upside for investors at current prices. For this reason, we recommend a NEUTRAL market rating finding the industry to be fairly price in public markets.

Investment Drivers

- **Rapid Adoption of AI Technologies:** Increasing adoption of AI technologies is driving significant demand for GPUs. These units are essential for AI training and inference, with a projected market of $450.9 billion by 2030, representing a 33.5% CAGR.[1] NVIDIA, AMD, and Intel are poised to capture much of this growth.

- **Tailwind from AI Market Expansion:** The forecasted surge in the global AI market, expected to exceed $1.9 trillion by 2030 with a CAGR of 36.6%, acts as a significant tailwind for GPU sales.[2]

- **Expansion in Cloud Gaming and Extended Reality:** The growing popularity of cloud gaming and extended reality (XR) is expected to boost GPU demand. The XR market is predicted to grow to $100.8 billion by 2026.[3]

Investment Risks

- **Market Valuation and Growth Expectations:** The significant rise in the market values of major GPU competitors, coupled with the concern that future growth is already priced into their stocks, presents a risk of limited upside potential.

- **Geopolitical and Regulatory Challenges:** The industry faces considerable risk from geopolitical tensions, particularly in regions critical to the GPU supply chain. Strict export controls and the potential for conflict could severely limit access to essential manufacturing resources and markets.

- **Supply Chain Concentration Risks:** The heavy reliance on a limited number of suppliers and manufacturers places the GPU industry in a vulnerable position. Any disruptions, from geopolitical strife, regulatory changes, or logistical challenges, could lead to significant supply chain bottlenecks, affecting product availability and profitability. In 2020, 87% of semiconductors were produced in Asia.[4]

Industry Statistics

Market Capitalization	($ in billions)
NVIDIA	$2,293
AMD	$290
Intel	$181

Price to Earnings	
NVIDIA	51.1
AMD	280.5
Intel	125.3

Price To Sales	
NVIDIA	25.0
AMD	10.6
Intel	3.9

Revenue	($ in millions)
NVIDIA	$60,922
AMD	$22,680
Intel	$54,228

EBITDA	($ in millions)
NVIDIA	$34,480
AMD	$4,176
Intel	$9,663

Margins (Latest Fiscal Year)

Margins bar chart — NVIDIA, AMD, Intel:
Gross Margin: NVIDIA 72.7, AMD 37.9, Intel 39.7
Operating Margin: NVIDIA 54.1, AMD 2.8, Intel 0.1
Profit Margin: NVIDIA 48.9, AMD 3.8, Intel 3.1
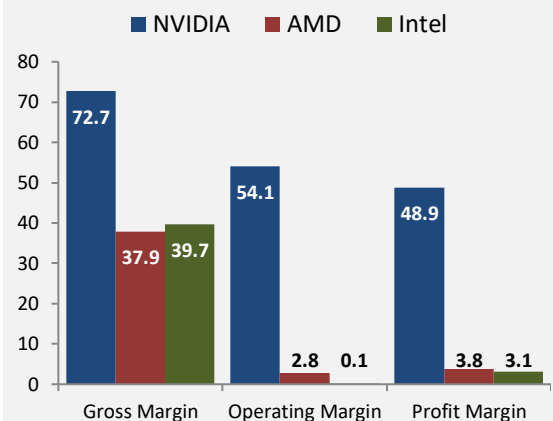
12-Month Relative Performance

Cumulative Return (%) chart from 03/23 to 03/24. Series: S&P 500, SOXX, NVIDIA, AMD, Intel.

Source: FactSet
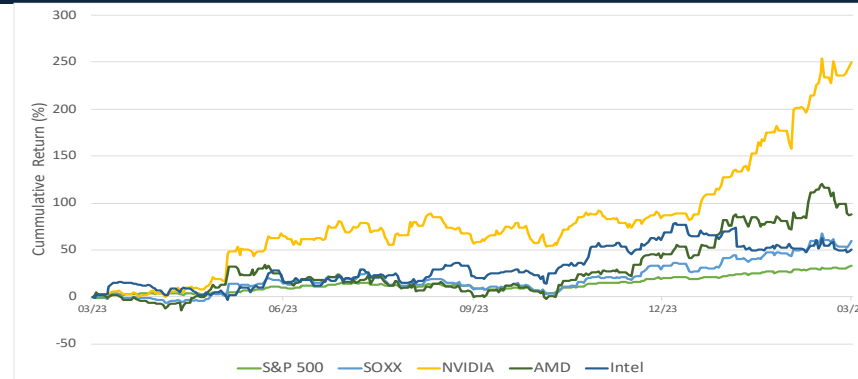
Industry Description

GPUs are optimized for parallel computing, enabling them to process multiple calculations simultaneously. This makes them ideally suited for computationally intensive tasks such as machine learning, computer graphics, image processing, and cryptocurrency mining. The GPU industry is experiencing a surge in demand, driven by catalytic breakthroughs in large language models (LLMs) and other forms of generative AI. Advancements in AI will continue to serve as a sales catalyst for the GPU industry.

Important disclosures appear on the last page of this report.

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

INDUSTRY OVERVIEW

GPUs Explained

GPUs use thousands of small, efficient cores to solve multiple calculations simultaneously. This type of problem-solving is called parallel computation. CPUs execute linear computation. In the image below, a GPU performs the operations concurrently, whereas the CPU multiplies 1x7, then 2x9, then 3x11 before summing up the results. Each additional problem adds a proportional amount of time to solve the equation using linear computation. The matrix math below is how generative AI works, except at scale—speed matters when training models on massive datasets. Startups and enterprises generate strong demand for GPUs as they experiment with AI technology.

Matrix Multiplication
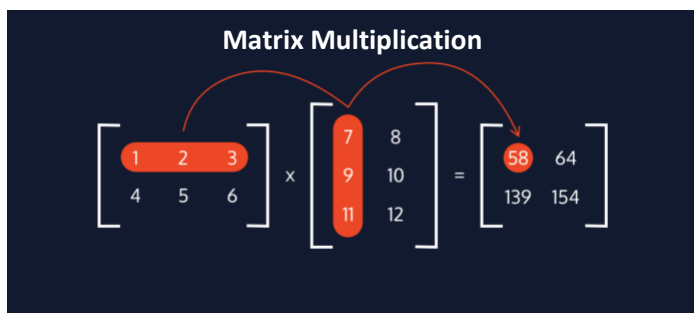
Source: Codeacademy

Competition

The GPU industry, a subset of the technology sector, is distinguished by its oligopolistic nature and dominated by three major firms: NVIDIA, AMD, and Intel.

|	Market Cap ($B)	Sales ($M)	Net Income ($M)
NVIDIA	$2,293	$60,922	$29,760
AMD	$290	$22,680	$854
Intel	$181	$54,228	$1,689

Source: FactSet

Within this industry, firms adopt one of two business models: fab or fabless. Companies following the fab model take ownership of the entire production process by investing in semiconductor manufacturing facilities. This vertical integration insulates firms from supply chain uncertainties, a critical advantage given the geopolitical sensitivities associated with crucial manufacturing hubs in regions like Taiwan and South Korea. However, the capital-intensive nature of this model demands substantial ongoing investment, potentially impacting financial liquidity and agility.

Conversely, firms adhering to the fabless model concentrate on the design and innovation aspect of GPUs, outsourcing manufacturing to specialized foundries such as Taiwan Semiconductor Manufacturing Company (TSMC) and Samsung. This approach enables an emphasis on R&D, driving advancements in GPU architecture and efficiency without the financial burdens of manufacturing infrastructure. The fabless model offers scalability and adaptability to market shifts, though it introduces risks related to supplier dependencies and production chain vulnerabilities.

NVIDIA: NVIDIA invented the GPU and specializes in designing GPU architecture and software. The company categorizes GPU sales based on specialized markets defined by the end-user. These markets include Data Center, Gaming, Professional Visualization, Automotive, and OEM & Other. GPUs used for AI training and inference fall into the Data Center category, representing 78% of NVIDIA's total sales.


NVIDIA 2024 Sales by Segment

- Data Center
- Gaming
- Professional Visualization
- Automotive
- OEM & Other
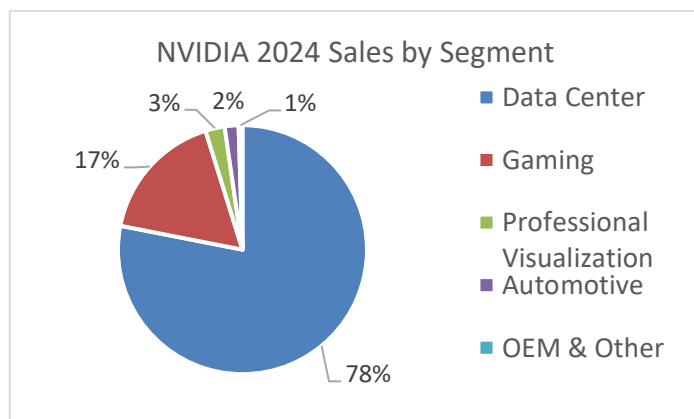
$ in millions   Source: FactSet

Adopting a fabless manufacturing model, NVIDIA concentrates on GPU design and development and outsources fabrication to third parties like TSMC and Samsung. This strategy allows NVIDIA to maintain minimal physical assets relative to its competition, securing NVIDIA's leadership position in GPU performance and profitability.
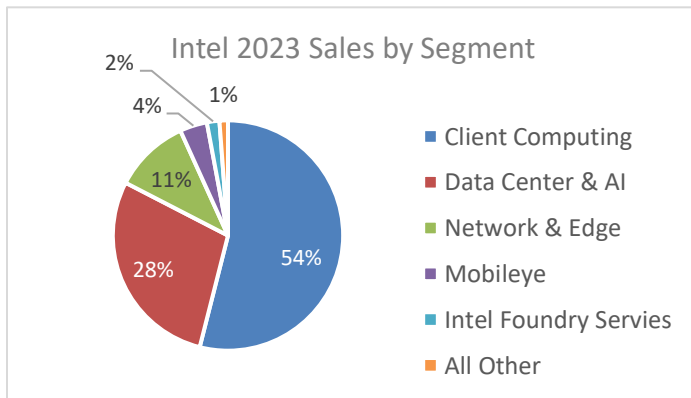
RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

AMD: AMD entered the GPU market in 2006 with the acquisition of ATI.[5] Since then, AMD has steadily increased its GPU offerings. AMD mainly markets its products to users who balance price with performance. AMD also offers a lineup of CPUs. AMD employs a fabless model, relying on third-party fabricators to manufacture its chips. AMD reports its revenue in four segments: Data Center, Client, Gaming, and Embedded.[6] AMD records GPU sales for AI training and inference under the Data Center category. In 2023, Data Center comprised 29% of total sales.

AMD 2023 Sales by Segment

- Data Center — 29%
- Gaming — 27%
- Embedded — 23%
- Client — 21%
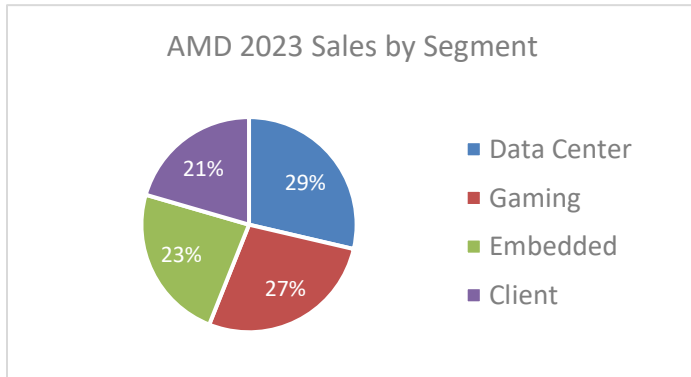
$ in millions   Source: FactSet

Intel: Intel vertically integrates the design and manufacture of its GPUs (and CPUs). This integration allows Intel greater control over its production processes and hedges against dependency risk. However, vertical integration requires significant capital investment in facilities and equipment, which increases Intel's asset base and depresses the firm's profitability and operational efficiency. Intel's operating segments include Client Computing Group, Data Center and AI, Network and Edge, Mobileye, Intel Foundry Services, and All Other.[7] GPUs used for AI accounted for 28% of total sales last fiscal year.


Intel 2023 Sales by Segment

- Client Computing — 54%
- Data Center & AI — 28%
- Network & Edge — 11%
- Mobileye — 4%
- Intel Foundry Servies — 2%
- All Other — 1%

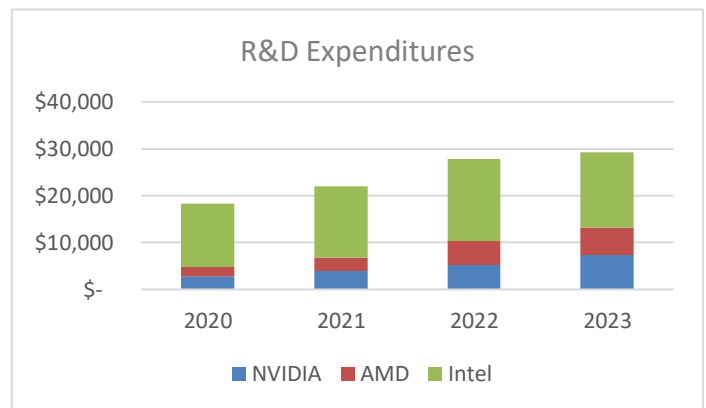*$ in millions*   Source: FactSet

Research and Development

Moore's Law: Moore's Law, named after Intel co-founder Gordon Moore, states that the number of transistors on a microchip will double every two years, increasing computational power.[8] As illustrated in the chart below, Moore's Law has endured, enabling computing power once believed unfathomable. Moore's Law may soon face its end as GPU makers approach physical limitations, such as the atomic size barrier and quantum effects.

Transitstors Per Chip
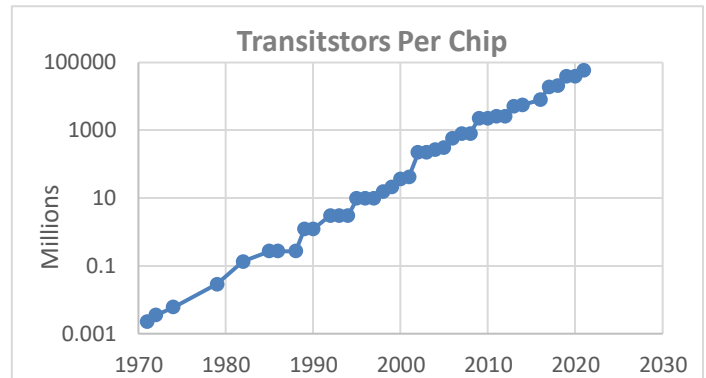
Source: Our World in Data

The industry invests heavily in research and development (R&D) to prolong Moore's Law and continue chip advancement. As a percentage of sales, R&D is often the most considerable operating expense and a significant barrier to entry. Last year alone, the combined big three spent $29.2 billion on R&D or 28% of their total revenue.


R&D Expenditures

NVIDIA · AMD · Intel

$ in millions   Source: FactSet

In relative terms, NVIDIA spends considerably less on R&D than AMD and Intel. Yet, NVIDIA continues to outperform its competitors in the GPU market, primarily due to its strategic focus on design and development while outsourcing manufacturing. By not having to invest in

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

manufacturing capabilities like Intel, NVIDIA allocates more resources toward innovation and enhancing the performance of its GPUs. This approach has translated into a superior product lineup with leading performance benchmarks, evidenced by NVIDIA's GPUs occupying the top thirteen spots in Geekbench's OpenCL rankings (see Market Leadership section below). However, NVIDIA's competitive edge could face challenges if AMD and Intel's higher R&D investments lead to significant advancements in GPU architecture.

R&D Expenditures
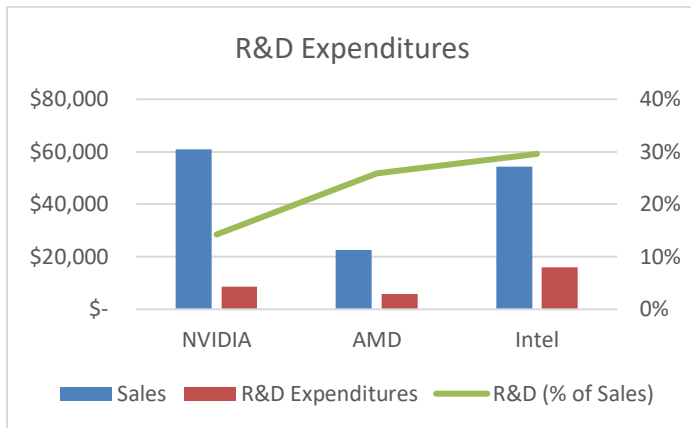
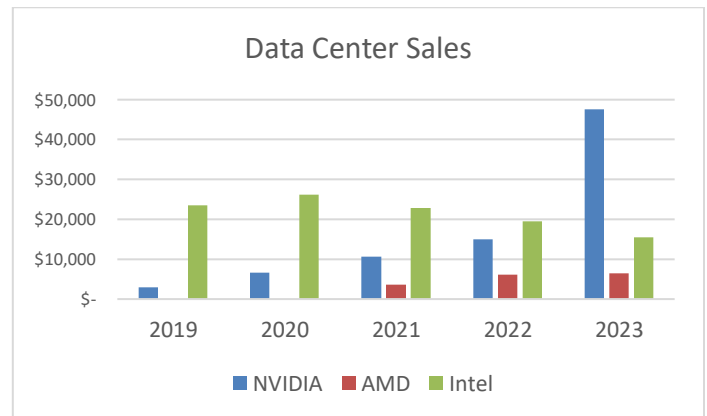*$ in millions*   Source: FactSet

Market Leadership

Amongst industry competitors, NVIDIA is the best positioned for the AI boom. Using Geekbench's OpenCL as a benchmark, the top thirteen GPUs belong to NVIDIA. As displayed in the following graphic, the difference between NVIDIA and its competition is stark. For context, the following scores are calibrated against a baseline score of 2,500. Higher scores are better, with double the score indicating double the performance.[9]
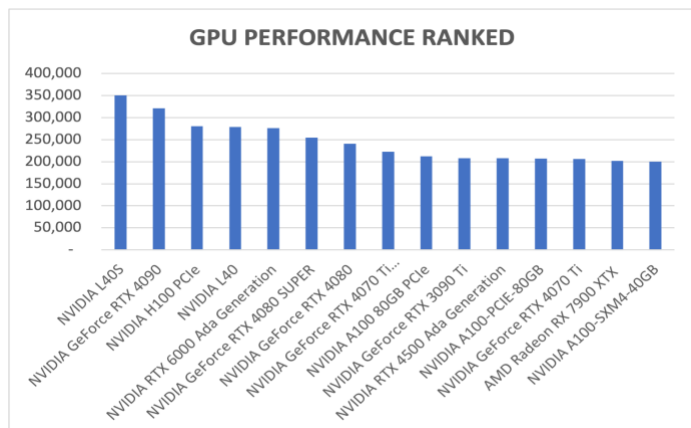
Source: Geekbench

NVIDIA's lead in GPU technology and performance is causing significant growth in Data Center sales. As seen in the graphic below, this growth is eroding Intel's once-dominant position in GPU sales despite Intel's higher R&D expenditures when compared to AMD and NVIDIA. This indicates a market shift where Intel's R&D investment has yet to translate into maintaining, let alone growing, its market share within this high-growth segment. While AMD began segmenting data center revenue in 2021 and has shown a positive trajectory since, Intel's decline in sales from 2021 to 2023 starkly contrasts with NVIDIA's surge, suggesting that Intel's strategies may not be as effectively aligned with market demands as its fabless competitors. This divergence raises critical concerns about the effectiveness of Intel's R&D spending and innovation strategy, emphasizing a need for Intel to reassess and realign its approach to regain its market share in the increasingly vital data center market.

Data Center Sales

$ in millions   Source: FactSet
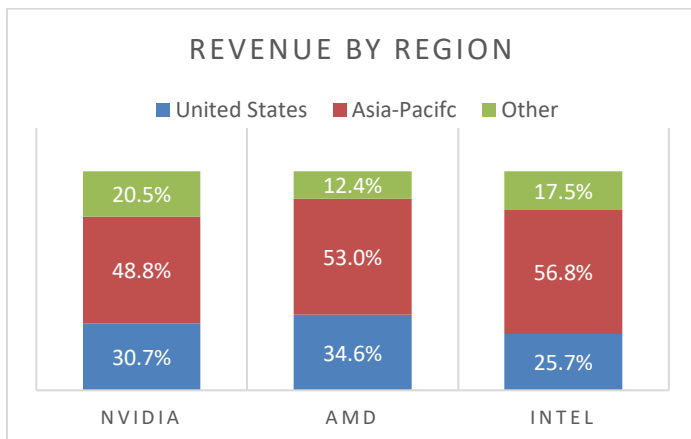
Market Segmentation

The GPU market is segmented by region, type, application, and end-user.

By Type: GPUs are categorized into discrete and integrated units. Discrete GPUs are standalone chips dedicated to graphics processing, offering higher performance. Integrated GPUs, in contrast, are built into the same chip as the CPU, providing cost-effective and energy-efficient processing. Discrete GPUs are, on average, more powerful than integrated GPUs and comprise most of the demand for AI training hardware.

By Region: Customers in the United States and the broader Asia-Pacific region generate the most GPU sales. China,

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

Taiwan, Singapore, and Japan comprise the Asia-Pacific region. Any disruptions to supply chains and shipping in the Asia-Pacific region would adversely affect revenue and overall profitability. Regional sales are based on the end-user's geographic area.

REVENUE BY REGION

	United States	Asia-Pacifc	Other
NVIDIA	30.7%	48.8%	20.5%
AMD	34.6%	53.0%	12.4%
INTEL	25.7%	56.8%	17.5%

Source: FactSet

By Application: Gaming, professional visualization, and data centers are the primary uses of GPUs. In gaming, GPUs deliver high-resolution graphics and fast rendering. Professional visualization utilizes GPUs for tasks such as CAD design and scientific simulations. Data centers sell GPU computing as a service to accelerate various computational tasks, including AI and big data analytics. Data centers are the most critical application for GPUs and will stimulate growth over the short and long term. In the automotive sector, GPUs are integral for advanced driver-assistance systems (ADAS) and infotainment.

By End-User: Consumers, enterprises, cloud service providers (CSPs), and automotive manufacturers are the main customers for GPUs. CSPs harness the power of GPUs to deliver computing power as software as a service and are responsible for the recent growth in the data center segment.

INDUSTRY TRENDS

Generative AI

AI is poised to accelerate productivity gains across the entire economy. Today, the most useful application is generative AI, a form of AI that creates new content, including text, images, audio, and video. Large language models (LLMs) such as ChatGPT fall into this category.

LLMs respond to a user's natural language inputs. A practical application for LLMs is computer programming. Previously, programming required technical know-how and hours of labor. Today, using an LLM, code can be generated in seconds. The use cases are far greater. Some models create original artwork in response to user descriptions. Video models such as OpenAI's Sora create photorealistic high-definition video.[10]

Natural language is the next step forward in computer software. Natural language will overtake graphical user interfaces (GUIs) like GUIs obsoleted command-line interfaces. For example, we foresee a future where a user asks Excel to create a graph specifying certain features rather than navigating a series of menus to achieve the desired result. Beyond productivity gains, the ability to understand and execute natural language instruction will democratize knowledge far beyond the printing press, creating unforeseen externalities. As the adoption of LLMs and generative AI expands, the demand for training will also increase, likely leading to a direct rise in GPU sales.

Autonomy

Integrating self-driving technology into the automotive industry significantly boosts the demand for advanced GPUs, essential for processing the extensive data required for machine learning and real-time decision-making while driving. The GPU industry is poised to benefit from this trend. The synergy between self-driving advancements and GPU sales is clear: vehicles becoming more autonomous necessitate increasingly powerful GPUs.

Ray Tracing and Advanced Graphics

Ray tracing is a technique used in computer graphics to simulate the physics of light. This process produces detailed shadows, reflections, and photorealistic images. GPUs are indispensable for this rendering technology due to their parallel architecture, which is well-suited for the computationally intensive tasks involved in ray tracing. GPUs can process multiple rays simultaneously, significantly speeding up the rendering process and making real-time ray tracing feasible for interactive applications like augmented reality (AR) and virtual reality (VR).

The increased use of ray tracing in AR and VR applications presents a significant opportunity. As extended reality (XR)

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

devices (incorporating both AR and VR) amass greater consumer adoption, buoyed by Apple's entrance into the space, investors should expect GPU sales to rise. The XR market is predicted to grow to $100.8 billion by 2026.[12]

Extended reality (XR) market size worldwide from 2021 to 2026 ($ in billions USD)
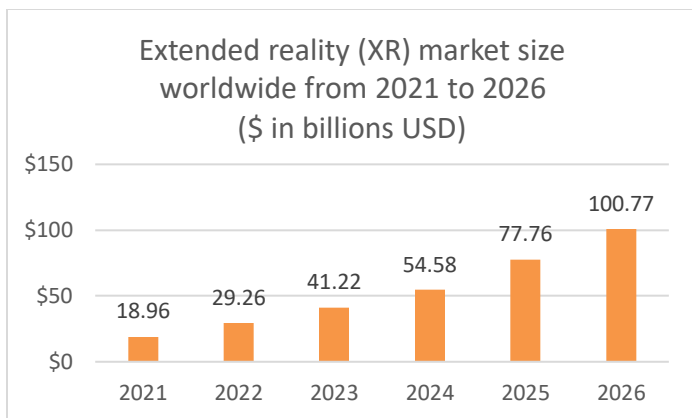
Source: Statista

Cloud Gaming

Cloud gaming, or game streaming, allows players to access video games via the internet on various devices without high-end hardware. Games are hosted at remote servers, and gameplay is streamed to the player's device, with inputs returned to the server in real time. This approach enhances accessibility by reducing the need for hardware upgrades and game installations. A fast, stable internet connection with low latency is necessary for the optimal experience. NVIDIA GeForce Now, Amazon Luna, Microsoft Xbox Cloud Gaming, and Sony PlayStation Now are popular game streaming services.

Cloud gaming is expected to grow at a CAGR of 33.6%, reaching a market value of $22 billion by 2028.[13] This will serve as a tailwind for both data center and gaming GPU sales. Data center sales will rise as Amazon, Microsoft, Sony, and others increase server capacity to accommodate gamers.
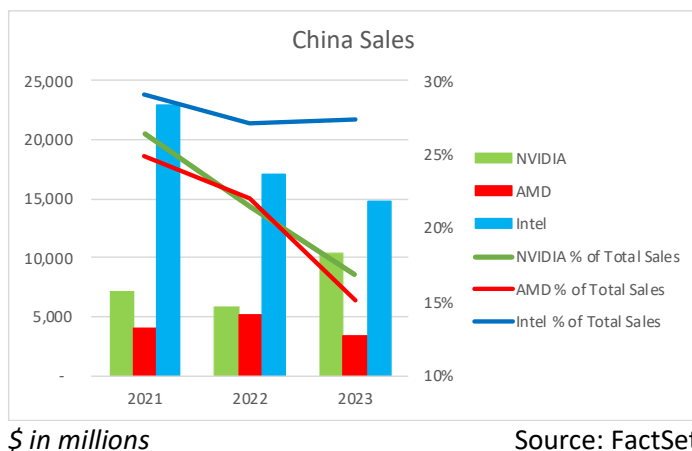
Export Controls

Effective October 2022, the United States government (US) implemented export controls to curb the People's Republic of China's (PRC) access to advanced computing infrastructure. This includes the most advanced GPUs. The export controls impede the PRC's ability to purchase and manufacture high-end chips used in military applications.[14] Chips for consumer items such as automotive and mobile

phones receive a presumption of approval, whereas more powerful semiconductors require a prior notice of sale and licensure.[15]

In October 2023, the US updated guidance to close loopholes used to circumscribe the export controls. For example, NVIDIA reduced the transfer speeds of its flagship GPU in half to comply with US regulations. The US shifted from interconnected speed to performance density as a threshold parameter to strengthen the rule. The US also expanded restrictions to include subsidiaries of companies headquartered in China, Macau (a semi-autonomous region in China), and other countries subject to arms embargos.[15]

The GPU industry will experience a fluid regulatory environment. One should look no further than US Commerce Secretary Gina Raimondo's commentary in December 2023 as evidence: "I am telling you if you redesign a chip around a particular cutline that enables [China] to do AI, I am going to control it the very next day."[16]

The new export control rules will likely adversely affect GPU sales and depress future revenue from Chinese customers. Moreover, the export controls may prompt China to accelerate its efforts towards domestic self-sufficiency in high-end GPUs used for AI training. The chart below illustrates a significant downturn in sales from China as a share of total sales due to trade constraints, primarily affecting NVIDIA and AMD. Intel's sales in China have remained more stable, mainly because CPU sales, which fall outside the trade restrictions, have not been impacted.

China Sales

$ in millions   Source: FactSet

Despite the adverse effects of the trade restrictions, the industry will continue to see revenue growth in its lower-

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

tier GPU offerings because export controls are limited to only the most advanced GPUs.

Mergers and Acquisitions

February 7, 2022, NVIDIA announced the termination of a stock purchase agreement for Arm Limited, a British semiconductor company. Unlike NVIDIA, Arm does not design semiconductors but licenses its proprietary technology to other firms. NVIDIA cited significant regulatory challenges as the impetus for the termination. The Federal Trade Commission's (FTC) complaint alleged the deal would substantially lessen competition and give NVIDIA unfair access to sensitive information provided to Arm by its licensees.[17] Moreover, it would incentivize NVIDIA to exercise its ownership of Arm's intellectual property at the expense of competition.[17] The market initially responded favorably, boosting NVIDIA 9.6% three days after the announcement before falling 10.6% during the last two trading days of the week. A week later, the stock jumped on the release of a statement by the FTC regarding the cessation of the acquisition.[18] Given the GPU industry's concentration level and the current administration's hostility towards large corporate mergers, it is reasonable to assume future merger activity will be limited to smaller acquisitions.

Market Reaction to NVIDIA-Arm Merger Termination
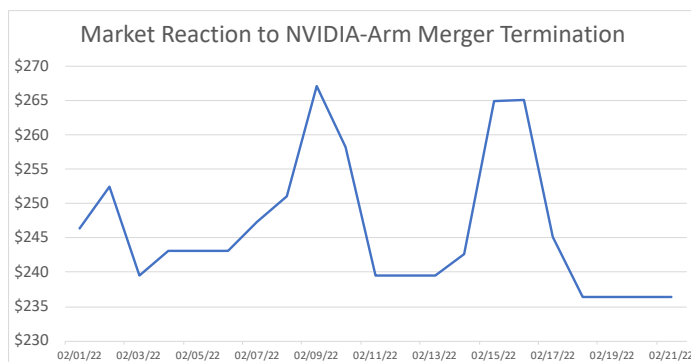
Source: FactSet

PEER COMPARISONS

Financial Metrics

The chart below compares various profitability measures over the trailing 12 months (TTM):

	NVIDIA	AMD	Intel
Revenue	$ 60.9	$ 22.7	$ 54.2

	NVIDIA	AMD	Intel
EBITDA	$ 34.5	$ 4.2	$ 9.6
Net Income	$ 29.8	$ 0.9	$ 1.7
FCF	$ 27.0	$ 1.1	$ (14.3)
Operating Margin	54.1%	2.8%	0.1%
Profit Margin	48.9%	3.8%	3.1%

$ in billions   Source: FactSet

Analyzing the financial metrics above, NVIDIA demonstrates superior financial health and performance compared to AMD and Intel. NVIDIA's EBITDA, which dwarfs AMD's and significantly leads Intel's, implies a stronger profit generation from operational activities. This is further confirmed by NVIDIA's net income, which exceeds its competitors, reflecting effective expense management and operational efficiency. NVIDIA's healthy free cash flow (FCF), as opposed to Intel's negative FCF also highlights a strong liquidity position, suggesting financial flexibility. High operating and profit margins for NVIDIA compared to its peers further indicate a combination of pricing power and efficient operations. Intel's fab model manifests in its financials with narrow margins and negative FCF, underscoring the capital-intensive nature of its integrated approach. Coupled with its superior GPU offerings, NVIDIA's relative financials cement its status as the market leader in the GPU industry.

Operating Metrics

The chart below compares various operational efficiency metrics:

	NVIDIA	AMD	Intel
ROA	55.7%	1.3%	0.9%
ROE	91.5%	1.5%	1.6%
Inventory Turnover	3.2	3.5	2.7
Asset Turnover	1.1	0.3	0.3
D/E	22.3%	4.0%	44.8%

Source: FactSet

The metrics above reveal a stark contrast in operational performance between NVIDIA, AMD, and Intel, attributable to their differing business models. NVIDIA's fabless model yields an ROA of 55.7% and an ROE of 91.5%, underscoring its capability to generate significant profit

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

with minimal assets and equity. Although AMD, also fabless, leads with the highest inventory turnover at 3.5, it does not convert this efficiency into profitability as effectively as NVIDIA. This discrepancy hints at NVIDIA's superior market execution and efficiency in its operations. Intel's traditional fab model manifests in lower ROA and ROE figures alongside a higher debt-to-equity ratio of 44.8%, reflecting its capital-intensive operations. The higher leverage could indicate potential vulnerability to market volatility and interest rate risks. NVIDIA's moderate 22.3% debt-to-equity ratio strikes a balance, facilitating growth without overextending financially. The metrics collectively suggest NVIDIA's fabless approach enhances profitability and affords greater agility, making it the most efficient operator among the three.

Valuation Metrics

The chart below compares various valuation metrics:

	NVIDIA	AMD	Intel
Market Cap	$ 2.3T	$ 288.7B	$ 179.4B
Enterprise Value	$ 2.3T	$ 285.9B	$ 203.6B
Trailing P/E	76.6	337.1	106.1
Forward P/E	38.0	53.2	31.8
P/S	37.4	12.8	3.3
P/B	53.2	5.2	1.7
EV/S	37.3	12.6	3.8
EV/EBITDA	63.8	68.9	18.1

Source: Yahoo! Finance

In evaluating the current financial landscape of key players in the GPU market—NVIDIA, AMD, and Intel—we observe significant variance in market valuations. NVIDIA's leading market capitalization and enterprise value of approximately $2.3 trillion reflect its dominant industry position and investor's high growth expectations. Conversely, while smaller, AMD's market cap of $288.7 billion reflects a promising growth trajectory in the eyes of investors. With a market cap of $179.4 billion, Intel appears to be perceived as a more mature entity with comparatively modest growth prospects.

The trailing and forward Price-to-Earnings (P/E) ratios reveal market sentiment regarding future profitability. NVIDIA's high trailing P/E ratio of 76.6 coupled with a forward P/E of 38.0 reflects a premium on its current

earnings but also suggests that investors expect its earnings to grow significantly. Similarly, AMD's trailing P/E of 337.1 with a forward P/E of 53.2 indicates a market assumption of considerable earnings growth, which, if unmet, could lead to a downward correction. Intel's trailing P/E ratio of 106.1 alongside a forward P/E of 31.8 depicts a market that is pricing in stabilization or improvement in earnings relative to its current performance.

The forward-looking metrics provide valuable insight into market expectations. NVIDIA's lower forward P/E relative to its trailing P/E may suggest an anticipation of solid growth in earnings, which could validate its current market premium if realized. In AMD's case, the high trailing P/E reflects an aggressive growth forecast, which may carry more risk if such growth fails to materialize. Conversely, Intel's forward P/E suggests that while it may not be expected to grow at the pace of its peers, there is an expectation of solidifying or increasing earnings power, making it an attractive proposition for value investors.

A similar trend is apparent in the Price-to-Sales (P/S) and Price-to-Book (P/B) ratios. NVIDIA maintains a commanding premium, further underscoring the market's bullish outlook on revenue and asset value growth. Intel's lower ratios in these categories signal an undervaluation of its sales and net assets, which could result in an opportunity for investors if the company can leverage its assets more effectively than the market currently anticipates.

The Enterprise Value-to-Sales (EV/S) and Enterprise Value-to-EBITDA (EV/EBITDA) ratios reinforce these narratives. NVIDIA and AMD are perceived to have their robust growth trajectories already factored into their stock prices, with their high ratios reflecting this sentiment. On the other hand, Intel's lower ratios may suggest a potential market undervaluation or a conservative view of its growth capabilities.

ECONOMIC OUTLOOK

United States: Inflation

Inflation's rise to 9.1% in 2022 caused the Federal Reserve (Fed) to increase interest rates rapidly.[19] Inflation has since fallen to 3.2%, and the market expects rate cuts in the near future.[20] As observed in data from the CME's FedWatch

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

tool below, the market is expecting multiple rate cuts consistent with Federal Reserve Chair Jerome Powell's recent commentary.[21] Currently, the target rate is 525 – 550.

Target Rate Probabilities for 12/18/24 FED Meeting
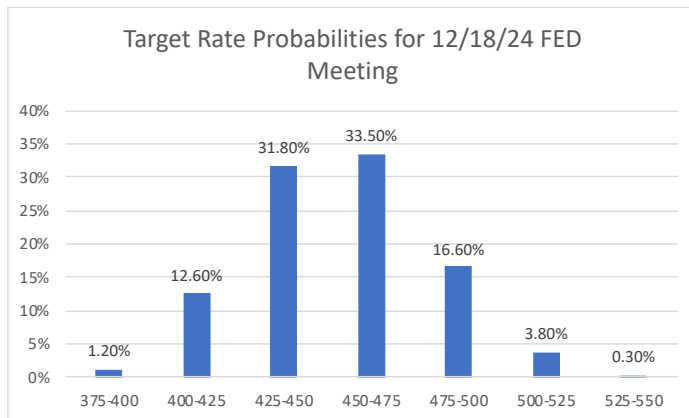
Source: CME Group

While we expect cuts, the distribution of expectations is too optimistic, with a bias towards the 400 – 450 range. In our estimation, a tight labor market, increased government expenditures, and shifting global supply chains will keep inflation stubbornly high. This will sustain a longer period of high-interest rates than expected. The target rate is likely to remain above 475 basis points. Higher-than-expected interest rates will impact all firms in the GPU industry.

In an environment where interest rates remain elevated for an extended period, Intel, NVIDIA, and AMD's valuations will likely come under pressure due to the higher discount rates applied to their future earnings. NVIDIA and AMD, with their high valuation multiples reflecting significant growth expectations, could experience more pronounced adjustments as the present value of their future earnings decreases with rising discount rates. While also susceptible to these macroeconomic conditions, Intel may see a relatively less severe impact on its valuation due to its more moderate valuation multiples. Overall, the sustained high-interest rate environment could necessitate a recalibration of stock valuations across the sector, with growth-oriented companies like NVIDIA and AMD potentially facing steeper revaluations.

RISKS

Valuation

The release of ChatGPT by OpenAI on November 30, 2022, marked a pivotal moment in the AI sector, demonstrating to the world the first practical and interactive application of generative AI technology. Recognizing the promise of this technology, investors have flooded the industry with capital, especially GPU firms. NVIDIA's stock has appreciated 457%, AMD's stock 131%, and Intel's stock 42%. Combined, the three share a market capitalization of $2.8 trillion. This raises the question of whether investor enthusiasm has propelled valuations beyond their intrinsic worth, potentially leading to limited upside and heightened volatility, should the market correct. While it's evident that high growth rates are anticipated and seemingly priced into the industry's current valuation, it's crucial to acknowledge the potential for novel applications and use cases that remain unforeseen today. Such developments could sustain demand and, in retrospect, justify the current market optimism despite concerns over potential overvaluation. Nevertheless, in our estimation, current prices exceed the industry's intrinsic value, justifying our underweight market rating.

Supply Chain Concentration

Photolithography: Photolithography is a critical step in the semiconductor manufacturing process in which ultraviolet light is used to etch nanometer-sized transitions into silicon wafers. ASML, a Dutch firm, primarily produces photolithography machines that make the most advanced GPU chips. No competitors match ASML's state-of-the-art machinery. ASML is a single point of failure in the industry's supply chain. Should an unforeseen event occur that paralyzes ASML's business, the industry as a whole will directly suffer. The risk is low. The Netherlands' neighborhood of nations is relatively peaceful, and its NATO membership guarantees security.

Fabrication: Outsourcing manufacturing operations to Asia carries significant geopolitical and supply chain risks for the GPU industry. The primary fabricators, TSMC and Samsung, are both exceptionally vulnerable. If hostilities turn kinetic across the Taiwan Strait or Korean Peninsula, supply will be severely disrupted. The US CHIPS and Sciences Act seeks to provide supply chain redundancy by

RESEARCH
HenryFund.Org

Tippie
College of
Business
THE UNIVERSITY OF IOWA

subsidizing the building out of domestic factories. If so, Intel's manufacturing abilities may erode NVIDIA's and AMD's market share.

Despite being a financial burden that impacts its margins and operational efficiency, Intel's capacity to manufacture chips in-house offers a form of redundancy that could become a competitive edge if NVIDIA's or AMD's supply chain is disrupted. This scenario suggests that, while Intel's current approach to chip fabrication may weigh on its financial performance, it also positions Intel to capitalize on market share opportunities potentially should NVIDIA or AMD face manufacturing constraints due to geopolitical tensions or other disruptions in its reliance on Asian fabricators.

SUMMARY

Investors navigating the GPU sector must weigh the promising potential of generative AI against the sector's lofty current valuations. Despite our conviction in AI's transformative impact and large total addressable market driving robust GPU demand, the market's current valuations exceed our comfort level. Investors should adopt a cautious stance, awaiting a market correction that aligns more closely with the sector's intrinsic value, consistent with its long-term growth and profitability prospects.

In the event of such a market adjustment, we recommend companies with substantial competitive advantages and enduring customer loyalty. NVIDIA emerges as the frontrunner, owing to its unparalleled GPU offerings crucial for training cutting-edge AI models. NVIDIA's dominance is further cemented by its CUDA software ecosystem, which optimizes GPU performance by enabling developers to fine-tune their models to NVIDIA hardware. This platform enhances performance and creates a dependency on its proprietary functions, thus ensuring user retention within the NVIDIA ecosystem. Lastly, despite the supply chain risks of the fabless model, outsourcing manufacturing to third-party fabricators enables NVIDIA to focus on improving GPU architecture while simultaneously facilitating a lean, agile, and more profitable business model. For similar reasons, we prefer AMD over Intel as a second choice for investors seeking exposure to the GPU industry.

REFERENCES

1. ReportInsights
2. Statista
3. Statista
4. CNBC
5. Avast
6. AMD 10-K (2023)
7. Intel 10-K (2023)
8. Intel
9. Geekbench
10. OpenAI
11. Wall Street Journal
12. Statista
13. Statista
14. U.S. Department of Commerce
15. Center for Strategic & International Studies
16. Reuters
17. Federal Trade Commission
18. Federal Trade Commission
19. AP News
20. U.S. Bureau of Labor Statistics
21. Wall Street Journal
22. FactSet

DISCLAIMER