Case Study: NVIDIA GPU

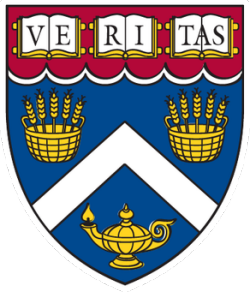
Prepared by

Group 21

Hunter Spence, Trishala Pugazhendhi, Jonathan Gardner, Yann-Daniel Adou, Guido Spotorno

A report presented for the course

ECON E-1615 Managerial Economics



Harvard Extension School

May 1, 2024

Contents

1 Executive summary **2**

2 Background **4**
2.1 Company & Mission 4
2.2 Market 4
2.3 Source of profits 4

3 How each force affects the organization **6**
3.1 Established Rivals 6
3.2 Consumers 9
3.3 Suppliers 10
3.4 Entrants 11
3.5 Substitute offerings 14

4 Organization Future **16**
4.1 Overview 16
4.2 Innovation Theory of Profit 16
4.3 NVDA Stock Overview 16
4.4 Moore's Law 19
4.5 Competitors 19
4.6 International Global Chip Conflict 20
4.7 TSMC and ASML Critical Role 20
4.8 Supply Chain 20

5 Advice to CEO **22**

Bibliography **23**

1 Executive summary

Nvidia is the market leader in Graphic Processing Units and Data Center accelerated computing products. They have expanded to create platforms for AI solutions, data science, AV, robotics, and 3D internet applications. NVIDIA plays a key role in the computing industry through its cutting-edge GPUs and related solutions and currently operates in four primary markets namely gaming, professional visualization, data center, and automotive.

Gaming and professional visualization are NVIDIA's legacy products. NVIDIA GPUs and software are ahead of the rest of the market. AMD and Intel are desperately trying to catch up. The company's future depends on its ability to innovate in emerging markets such as automotive and data center markets. NVIDIA holds a strong position but automotive revenue decline has impacted growth. The increase in Qualcomm's market share and its revenue growth in the automotive division cements Qualcomm to be their strongest competitor. However, in the data center market, NVIDIA dominates with a market share that analysts estimate above 95%.

NVIDIA's traditional competitors are AMD and Intel. NVIDIA's four largest customers are; Microsoft, Meta, Amazon, and Google. They account for 40% of NVIDIA's total revenue. All of these companies are actively developing their next-generation AI chips to compete against NVIDIA's products.

Though companies are developing products to compete, many factors block new entrants into the GPU manufacturing industry. The GPU manufacturing industry is a capital-intensive industry with high developmental and research costs. NVIDIA has around 11k intellectual property and patents. While the large economies of scale and patent protection act as barriers to new entry, restrictions on exports of GPUs could encourage new entrants.

Manufacturing Processing Units require a complex supply chain. To manufacture GPUs, NVIDIA must source, process, and transport raw materials to their facilities. These processes often are expensive to the buyer and the environment. Existing customers compete for these resources, with some companies having a large share of the distribution channels, potentially limiting access to efficient distribution channels for new entrants. While the complex supply chains protect NVIDIA from new entrants, they also pose considerable risks for the company. NVIDIA faces supply chain risks primarily due to its dependency on TSMC, a key supplier with advanced semiconductor facilities concentrated in Taiwan. This dependency limits NVIDIA's flexibility in supplier choice, enhancing TSMC's bargaining power and potentially affecting NVIDIA's production and pricing during demand surges. Taiwan is also compromised by China's regional policies, threatening national security. If China were to be successful in controlling trade from Taiwan, NVIDIA's supply chain would be compromised. A recommendation would be to increase supply chain security, by supporting TSMC manufacturing in other countries. If NVIDIA could integrate or purchase lithography machines from TSMC, they could manufacture their chips independently. NVIDIA could also explore strategies such as diversifying suppliers, fostering technological alternatives, or increasing vertical integration to mitigate supply chain risks.

With established rivals and high barriers for new entrants, NVIDIA currently stands at the forefront of the AI revolution and currently enjoys healthy profit margins from its 'Compute and network' and 'Graphics'. While there are substitutes to gaming for entertainment purposes, gaming has been evolving and also branching to include game streaming and competitive gaming. The growth of ESports has promoted the growth of GPU's popularity. While GPU can be substituted by cheaper integrated graphics card which is embedded in the CPU, the performance of GPU outweighs the price.

NVIDIA's stock price has increased by over 2,000% over the last five years. Making the company the hottest stock on Wall Street, but also the 3rd largest company in the US. NVIDIA's current

Market Capitalization is $2 Trillion. The CEO/Founder, Jensen Huang, positioned the company to strike emerging markets. As the world becomes more reliant on data, Processing Units that support data centers, AI, servers, and cloud networks are in higher demand. Nivida has been able to develop products that are tailored to these new services. NVIDIA's revenue was dominated by GPU sales, but since 2019, "Compute and Networking" NVIDIA has now added $45 Billion in revenue per year. Eclipsing their legacy GPU products revenue. NVIDIA's ability to pivot and develop new technology for emerging markets is the reason they are so successful. NVIDIA faces little threat from competitors and has the resources to develop faster than the rest.

Overall, NVIDIA is expected to continue being the market leader for the next decades. The CEO must keep the magic that put NVIDIA in this position. AI technologies will transform how computers process and analyze data. NVIDIA's head start in this market should allow them to keep their competitive advantage.

2 Background

2.1 Company & Mission

NVIDIA is a leader in visual computing technologies and the inventor of the GPU (Graphics Processing Unit), a high-performance processor used to accelerate compute-intensive workloads. NVIDIA plays a key role in the computing industry through its cutting-edge GPUs and related solutions. With a mission to "develop high-performance computers that scientists, researchers, artists, and creators from around the world use to create the future and improve lives," NVIDIA leverages its expertise in GPU technology to be an innovation leader in the computing industry (NVIDIA Corporation, 2024).

2.2 Market

NVIDIA's core business revolves around designing and manufacturing GPUs for gaming, professional visualization, data centers, and automotive markets. Their GPUs are widely used to accelerate compute-intensive workloads in these markets. NVIDIA's legacy is established in the gaming and professional visualization markets, while data center and automotive solutions are predicted to be the future of NVIDIA (McDowell, 2024).

2.2.1 Gaming Market

NVIDIA provides a variety of products for gamers, ranging from high-performance GPUs to gaming platforms. Its GeForce line of GPUs is established to target both PC and console gaming.

2.2.2 Professional Visualization Market

NVIDIA also provides industry professionals with tools for their daily jobs like design, engineering, architecture, and media. NVIDIA provides GPUs to support industry applications that interact with 3D content like CAD/CAM software, virtual reality (VR) experiences, etc.

2.2.3 Data Center Market

High-performance computing is an area with significant growth due to the rise of AI. Products like Tesla GPUs, CUDA, and TensorRT are designed to aid tasks like deep learning and big data processing.

2.2.4 Automotive Market

NVIDIA DRIVE platform includes hardware and software components for perception, mapping, planning, and control in autonomous vehicles. These components provide solutions for advanced driver assistance systems (ADAS) and shape the future of transportation.

2.3 Source of profits

NVIDIA's profits are derived from a combination of hardware sales, software licenses, and strategic partnerships across various industries. Revenue for fiscal year 2024 was $60.9 billion up 126% from the prior year with a 15% growth in Gaming. The source of its greatest amount of revenue is its 'Compute and Networking' business segment, which includes artificial intelligence (AI). Compute

Reportable Segments

Revenue by Reportable Segments

	Year Ended			
	Jan 28, 2024	Jan 29, 2023	$ Change	% Change
	($ in millions)			
Compute & Networking	$ 47,405	$ 15,068	$ 32,337	215 %
Graphics	13,517	11,906	1,611	14 %
Total	$ 60,922	$ 26,974	$ 33,948	126 %

Figure 1: Revenue by Reportable Segments (NVIDIA Corporation, 2024)

& Networking segment grew 266% due to higher shipments of the NVIDIA Hopper GPU computing platform.

The revenue from sales outside of the United States of America accounted for 56% of total revenue for the year 2024. In the Compute & Network segment, sales to one customer represented 13%, and indirect customer sales attributed to 19% of total revenue, unlike previous years where there were no customers with 10% or more of total revenue.

2.3.1 Compute and Networking Segment

According to NVIDIA's Form 10-K, the Compute and Networking segment is made up of:

- Data Center accelerated computing platforms and end-to-end networking platforms including Quantum for InfiniBand and Spectrum for Ethernet
- NVIDIA DRIVE automated-driving platform and automotive development agreements
- Jetson robotics and other embedded platforms
- NVIDIA AI Enterprise and other software
- DGX Cloud software and services (NVIDIA Corporation, 2024)

The Compute and Networking segment accounted for nearly 78% of NVIDIA's total revenue and delivered $47.4 billion in FY 2024, up 215% from FY 2023.

2.3.2 Graphics Segment

According to NVIDIA's Form 10-K, the Compute and Networking segment is made up of:

- The GeForce GPUs for gaming and PCs, the GeForce NOW game-streaming service and related infrastructure, and solutions for gaming platforms
- The Quadro/NVIDIA RTX GPUs for enterprise graphics design
- Virtual GPU (vGPU) for cloud-based visual and virtual computing
- Automotive platforms for infotainment systems
- Omniverse Enterprise software for the construction and operation of metaverse and 3D internet apps (NVIDIA Corporation)

The Graphics segment delivered around 22% of NVIDIA's total revenue at $13.51 billion in FY 2024, up 14% from FY 2023.

3 How each force affects the organization

3.1 Established Rivals

NVIDIA chips mainly compete in four specific markets: gaming, professional visualization, data center accelerators, and automotive. Gaming and professional visualization are NVIDIA's traditional markets where the firm had and has a stronghold. Datacenter accelerators and automotive are the markets with higher growing potential. In the data center accelerators market –a market that includes chips used in AI– NVIDIA has an even higher market share than in Gaming and professional visualization; this dominance is explained partially by the "moat" of NVIDIA's software CUDA; however, in the automotive market, the competition is more intense.

3.1.1 Gaming and Professional Visualization GPUs Competition

NVIDIA enjoys a strong dominance in its "legacy" markets: gaming and professional visualization. The graph below shows NVIDIA, AMD, and Intel's share of discrete GPU shipments toward AIB partners from Q4'22 to Q4'23, according to the latest John Peddie's Research survey[15]. AIB partners are known for their great-performance products powered by discrete GPUs. One example of an AIB partner is MSI, which produces high-performance gaming and professional visualization products (e.g., motherboards, PCs, and laptops).

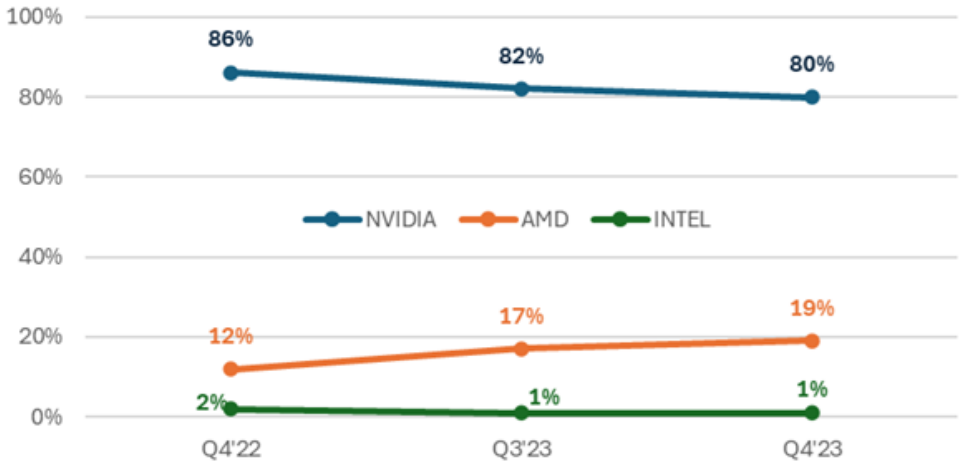


Figure 2: Discrete GPU shipments to AIB partners market share

As of Q4'23, NVIDIA holds 80% of the GPU shipments share to AIB partners, whereas AMD follows far behind with a gap of more than 60pps at 19%. Intel, with a very small portion of the business, has only 1% of the market share. This graph is a good representation of NVIDIA's strength in the gaming and professional visualization GPU markets.

Steam is the largest video game digital distribution platform for PC gaming developed by Valve Corporation with 120 million registered users. There are 33.5 million users at peak time and an average of 30 million users on Steam. The company conducts a monthly survey to collect data about what kinds of computer hardware and software their customers are using. The latest Steam hardware and software survey results are another evidence of NVIDIA's dominance in the GPU gaming market. We can see in the graph below that 78% of Steam's surveyed customers use NVIDIA graphics cards with the top 10 rank graphic card usage being held by NVIDIA graphic cards.

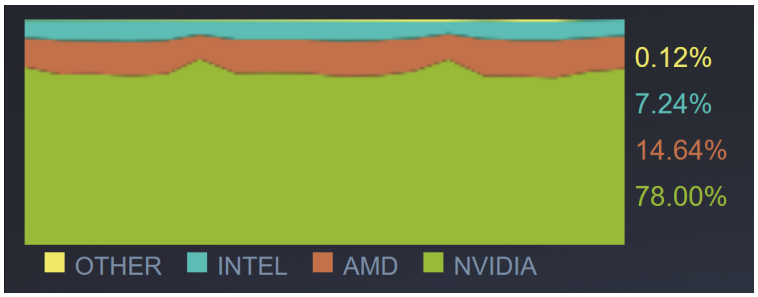
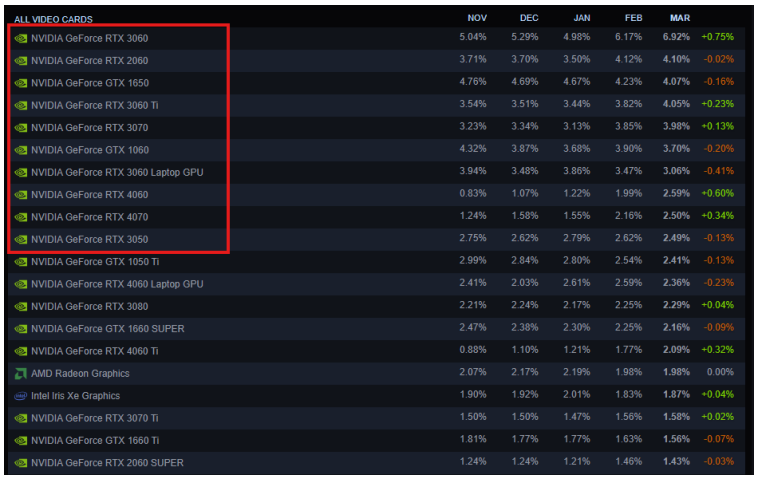


Figure 3: Video Card Usage by Manufacturer for October 2022 - March 2024



ALL VIDEO CARDS	NOV	DEC	JAN	FEB	MAR	
NVIDIA GeForce RTX 3060	5.04%	5.29%	4.98%	6.17%	6.92%	+0.75%
NVIDIA GeForce RTX 2060	3.71%	3.70%	3.50%	4.12%	4.10%	-0.02%
NVIDIA GeForce GTX 1650	4.76%	4.69%	4.67%	4.23%	4.07%	-0.16%
NVIDIA GeForce RTX 3060 Ti	3.54%	3.51%	3.44%	3.82%	4.05%	+0.23%
NVIDIA GeForce RTX 3070	3.23%	3.34%	3.13%	3.85%	3.98%	+0.13%
NVIDIA GeForce GTX 1060	4.32%	3.87%	3.68%	3.90%	3.70%	-0.20%
NVIDIA GeForce RTX 3060 Laptop GPU	3.94%	3.48%	3.86%	3.47%	3.06%	-0.41%
NVIDIA GeForce RTX 4060	0.83%	1.07%	1.22%	1.99%	2.59%	+0.60%
NVIDIA GeForce RTX 4070	1.24%	1.58%	1.55%	2.16%	2.50%	+0.34%
NVIDIA GeForce RTX 3050	2.75%	2.62%	2.79%	2.62%	2.49%	-0.13%
NVIDIA GeForce GTX 1050 Ti	2.99%	2.84%	2.80%	2.54%	2.41%	-0.13%
NVIDIA GeForce RTX 4060 Laptop GPU	2.41%	2.03%	2.61%	2.59%	2.36%	-0.23%
NVIDIA GeForce RTX 3080	2.21%	2.24%	2.17%	2.25%	2.29%	+0.04%
NVIDIA GeForce GTX 1660 SUPER	2.47%	2.38%	2.30%	2.25%	2.16%	-0.09%
NVIDIA GeForce RTX 4060 Ti	0.88%	1.10%	1.21%	1.77%	2.09%	+0.32%
AMD Radeon Graphics	2.07%	2.17%	2.19%	1.98%	1.98%	0.00%
Intel Iris Xe Graphics	1.90%	1.92%	2.01%	1.83%	1.87%	+0.04%
NVIDIA GeForce RTX 3070 Ti	1.50%	1.50%	1.47%	1.56%	1.58%	+0.02%
NVIDIA GeForce GTX 1660 Ti	1.81%	1.77%	1.77%	1.63%	1.56%	-0.07%
NVIDIA GeForce RTX 2060 SUPER	1.24%	1.24%	1.21%	1.46%	1.43%	-0.03%

Figure 4: Video Card Usage Rank for November 2023 - March 2024

For the last two and a half years, NVIDIA's card has been the clear choice of most Steam users. In the last survey of March 2024, NVIDIA's cards were the choice of 78% of the users. Far behind, AMD's cards were the choice for 14.64% of users. Finally, Intel cards were chosen by only 7.24% of the users

3.1.2 Data center accelerator GPUs competition

As big as NVIDIA's dominance is in the gaming and professional visualization markets, it's not as big as its dominance in the data center accelerators market, where analysts estimate that its market share is around 95-98%. Data center accelerators are used in High-Performance Computing (HPC) and cloud-based services. High-performance computing is needed for applications such as machine learning, AI, weather forecasting, big data analyses, etc.

Traditional NVIDIA rivals like Intel and AMD are building powerful chips to compete against NVIDIA's in this market. Both Intel and AMD are chip manufacturers that traditionally focused on phone and personal computer markets. However, since the boom of AI, these companies have been redeploying resources to construct chips for AI. Last December, AMD launched its MI300 GPU, which directly competes against NVIDIA's H100, whereas Intel is building its GAUDI3 AI GPU that will also compete head-to-head against NVIDIA's GPU.

AMD is trying to get market share by offering its GPU at a significant discount: between $10,000 and $15,000 per unit. Compared to the NVIDIA H100 chip, which costs $40,000, AMD's price is tempting. However, NVIDIA's stronghold on the data accelerators market comes not only

from the high quality of its chips but also from its CUDA software stack, which has been refined and continually improved for AI applications and other types of workloads.

Intel is also developing the GAUDI3 data accelerator chip that will launch in the second half of 2024 and compete with NVIDIA's H100. Intel will also face the same challenges AMD faces against NVIDIA's high-quality hardware and software.

But NVIDIA faces not only threats from its traditional rivals; firms like Microsoft, Google, Amazon, and Meta are also investing in building their own chips specialized in AI. Their chips won't be used in gaming, professional visualization, the automotive industry, or other important applications of NVIDIA's GPUs, but they will focus specifically on AI applications.

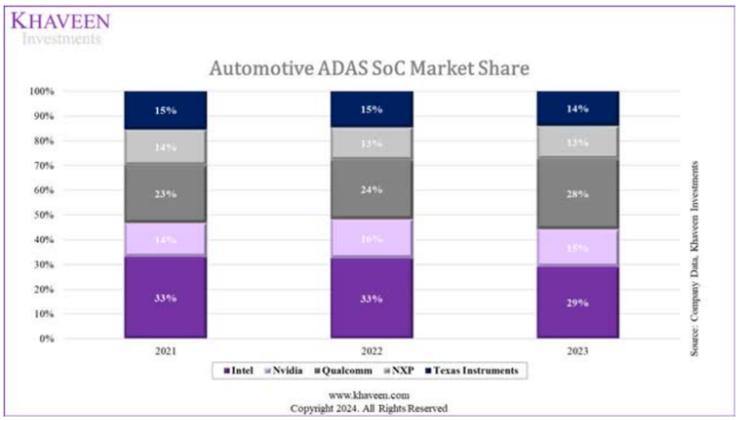
NVIDIA still has the advantage of having been in the business of developing and manufacturing high-powered chips for decades. They are ahead of the competition, but the question is whether they can keep the edge in the next decade.

Analysts estimate that the AI chip market will be $400 billion by 2027 and $1 trillion by 2030 (as a reference, the entire chip industry in 2023 was $500 billion). Many players besides NVIDIA, AMD, and Intel are expected to try to capture part of this market. It won't be surprising that more firms—big and small—will join them.

3.1.3 Automotive GPUs Competition

In the last earnings call, NVIDIA informed that its automotive division revenue was $281 million in Q4, representing a fall of 4% year on year. These results show a different story than the one of success, growth, and dominance that NVIDIA enjoys in the data center accelerator market.

The automotive market is strategic for NVIDIA; the company believes this market will be $300 billion over the long term (as big as they believe the market for data accelerators will be). In contrast with the data accelerator market, NVIDIA is not the leader in the automotive market, and the market does not rely on its CUDA software stack, making NVIDIA a weaker competitor in the automotive market than in the data center accelerators market. The graph below shows the market for the last three years in the Automotive ADAS SoC market (ADAS SoC stands for advanced driver-assistance system on chip; ADAS chips are like the brain of a car.)



Company Data, Khaveen Investments

Figure 5: Automotive ADAS SoC Market Share

This graph shows that NVIDIA is the 3rd biggest player in this market and that its market share has not grown in the last 3 years, fluctuating between 14% and 16%. Qualcomm, on the other hand, has grown its market share from 23% to 28% from 2021 to 2023 and has reported in its

latest earning calls that its automotive division grew +31% year-over-year. Intel, the lead player in ADAS, also seems to be struggling and losing market share against Qualcomm.

Qualcomm's growth in the ADAS market is bad news for NVIDIA. Qualcomm is NVIDIA's most direct competitor in the automotive market because both companies are the only ones that provide ADAS, digital cockpits, and infotainment solutions. With technology moving towards a more integrative solution for the automotive market, the future of car chips looks more like data center accelerator GPUs than traditional automotive GPUs. This future is a good fit for companies like Qualcomm and NVIDIA. Intel and AMD are also creating integrated solutions for this market, but they are less comprehensive than NVIDIA's and Qualcomm's. However, NVIDIA's automotive future doesn't look that bright: NVIDIA does not have a dominant software stack like it has in other markets. NVIDIA's business is stagnant while Qualcomm's is growing, and Qualcomm has a legacy advantage since it has been on connectivity solutions in the automotive industry for a long time.

3.2 Consumers

Throughout history, advances in technology have always been at the forefront of human evolution and way of life. As technology evolved and technology companies began to take firmer stances in cultural and societal transformations, one cannot help but wonder what sort of reactionary forces the clientele of such companies have on their day-to-day operations and what impact the consumers have on these companies. Are the goods and services provided by technology companies shaping our societies, or are there the needs and demands of societies that shape and dictate the production of goods and services from technology companies?

Currently standing as a leading pioneer in the field of Artificial Intelligence and chip technologies is NVIDIA. According to Yahoo Finance, the "demand for NVIDIA chips is so far above supply" that its CEO Jensen Huang "had to discuss how 'fairly' the company decides who can buy them"[22].

Microsoft, Meta (formerly Facebook), Amazon, and Alphabet top the list of customers contributing to NVIDIA's remarkable success, contributing to "nearly 40% of its revenue" according to the LA Times. NVIDIA's strategic partnerships with these industry giants may, however, be costly in the future. In parallel to the strong and mutually beneficial alliances that these tech giants have with NVIDIA, they also aim to emerge as pioneers of innovation in the tech industry.

Apple, for example, began making its own chips and developing its homegrown AI technology. According to CNBC.com [24], Apple's silicon team has grown to thousands of engineers working across labs all over the world, including in Israel, Germany, Austria, the U.K., and Japan. Within the United States, the company has facilities in Silicon Valley, San Diego, and Austin, Texas and, unlike traditional chipmakers, Apple is not making silicon for other companies. Apple aims to focus on the product and optimize the chips for their products. Meta also follows a similar trajectory.

According to Reuters[16], Meta plans to deploy a new version of a custom data center chip referred to internally as "Artemis" that will help reduce its reliance on NVIDIA. Once again, the story repeats itself with Microsoft. During the Microsoft Ignite Conference 2023, Microsoft announced its homegrown chip called "Maia". The chips are "custom-designed computing chips" whose mission is "bringing key technologies in-house" according to Reuters. Amazon, another principal consumer of NVIDIA's technologies, similar to other tech giants, is also developing its chips AWS Graviton4 and AWS Trainium2.

As perfectly summarized by the New York Times, "Chafing at their dependence, Amazon, Google, Meta, and Microsoft are racing to cut into NVIDIA's dominant share of the market" [28].

The general public also qualifies as one of NVIDIA's customer base and on that front, little remains to be desired on NVIDIA's ability to project unchallenged and sustainable long-term

profitability at levels well above that of its competitors. For years, NVIDIA and AMD, as well as Intel to a lesser extent, have been locked in a constant battle for the title of best graphics card manufacturer. They are "the choices" every PC user faces when working on a new build, buying a desktop, or a laptop according to DigitalTrends [17]. The constant introduction of substitute goods from AMD and Intel provides the general public with alternatives, which ultimately translates into a lower quantity demanded for NVIDIA's goods. Unless NVIDIA can achieve a monopoly in its market, its general public customers will continue to flock to providers of substitute goods, in search of cheaper or better alternatives, thus hurting the company's profitability.

NVIDIA's strategic partnerships with its main consumers of goods and services, even though currently beneficial, undoubtedly pose a threat to its supremacy in the Tech industry. The willingness of these consumers to become self-reliable will almost certainly contribute to NVIDIA's perceived riskiness from investors and would negatively affect its ability to maintain high-profit margins in the future.

3.3 Suppliers

NVIDIA, an innovator in visual computing and the progenitor of the GPU, is heavily dependent on a complex supply chain to procure high-performance processors integral to its products. The firm's reliance on a restricted set of suppliers, particularly for semiconductor manufacturing, significantly increases the bargaining power of these suppliers. For instance, NVIDIA's semiconductor fabrication is predominantly commissioned to Taiwan Semiconductor Manufacturing Company (TSMC), the preeminent independent semiconductor foundry globally. TSMC operates multiple advanced facilities across the island, including four 12-inch wafer GIGAFAB® fabs and several 8-inch and 6-inch wafer fabs all located in Taiwan. TSMC also has a 12-inch wafer fab in Nanjing, China, and is involved in manufacturing operations in the United States and Japan, but the core of its manufacturing remains in Taiwan ("TSMC Fabs"). TSMC's cutting-edge technology and prodigious production capabilities are crucial for NVIDIA, positioning the chipmaker as a supplier with considerable leverage.

Given the geographical concentration of NVIDIA's main supplier, TSMC, NVIDIA as a key customer—faces potential risks in the event of significant geopolitical disruptions affecting Taiwan. The region's geopolitical tensions, particularly with China, could impact TSMC's operations, leading to supply chain disruptions for companies like NVIDIA that heavily rely on its semiconductors.

The concentration of essential suppliers like TSMC underscores that any changes in pricing, availability, or technological enhancements directly affect NVIDIA's operational efficacy. The semiconductor sector is cyclical, where surges in demand can precipitate shortages and escalated costs, potentially decelerating production and impairing NVIDIA's capacity to satisfy market demand.

The scarcity of substitutes for the components needed by NVIDIA underscores suppliers' elevated leverage, enabling them to influence terms and pricing. The specialized nature of high-performance GPUs mandates cutting-edge technology that only a select few global suppliers can furnish. This lack of alternatives further amplifies suppliers' influence, permitting them to dictate terms with greater authority. NVIDIA's ability to switch suppliers is limited due to TSMC's unique capabilities in advanced semiconductor manufacturing. TSMC is a leading company in the industry, pioneering technologies like 5-nanometer and 7-nanometer processes, which are crucial for manufacturing high-performance chips like those used by NVIDIA. This specialization makes finding alternative suppliers with similar technological capabilities challenging.

Furthermore, the costs associated with transitioning suppliers are problematic. Migrating to an alternative semiconductor manufacturer entails substantial fiscal expenditure, material investment, and risk. Prospective suppliers must be evaluated to ascertain compliance with NVIDIA's quality

and reliability standards. Such transitions necessitate adjustments in production processes and the cultivation of new supply relationships, entailing potential disruptions that could defer product releases and inflate operational costs.

The critical nature of inputs from suppliers like TSMC illuminates another dimension of risk – the potential for forward integration. While currently minimal, the threat remains that suppliers might opt to enter NVIDIA's market by developing competing products or technologies. TSMC's concentrated manufacturing presence in Taiwan poses a risk to NVIDIA due to geopolitical tensions. Moreover, the difficulty in substituting TSMC with another supplier due to technological specialization heightens this risk, potentially impacting NVIDIA's supply chain resilience. Although mitigated by NVIDIA's robust market position and brand reputation, this possibility remains a pertinent strategic consideration.

The significant power wielded by suppliers necessitates that NVIDIA maintains robust, collaborative relationships with its principal suppliers. Long-term contracts and strategic alliances can help stabilize supply terms and pricing, ensuring a steady flow of innovation that is crucial for maintaining NVIDIA's technological leadership.

NVIDIA can implement several strategic measures to mitigate the risks associated with substantial supplier power. Promoting competition among suppliers, where feasible, can diminish dependency on any single provider. Strategic measures involve fostering alternative technologies or encouraging more suppliers to enhance their technological capabilities. Encouraging or investing in the geographic diversification of TSMC's operations could be beneficial. TSMC itself has been expanding its manufacturing footprint outside of Taiwan, such as in the United States. NVIDIA could support or incentivize further expansion in geopolitically stable regions. Furthermore, NVIDIA could consider augmenting its vertical integration by investing in or acquiring capabilities further up the supply chain, thereby securing more control over its critical components. Continued investment in research and development is also vital, enabling NVIDIA to discover new materials or methodologies that could provide alternative solutions to its needs.

3.4 Entrants

The threat of new entrants refers to how a new entrant would affect the existing firm in an industry and affect the competition in the market. In an industry where it is difficult for new entrants to enter the market, there will be very few firms in the industry. When there are few firms, for example, the firms are in a monopoly or duopoly, the firms would have more power over the pricing of the products. This would lead to higher profit margins.

Since 2019, NVIDIA and Advanced Micro Device (AMD) Inc. have been in a duopoly with NVIDIA controlling about 77.3% of the market and AMD controlling the remaining 22.7% in 2019. Intel has been in work to break this duopoly but hasn't yet been successful in breaking this. Many factors need to be looked into to gain an understanding of why it is hard to enter the GPU manufacturing industry, namely capital investments, development and research, brand loyalty, distribution, and technology.

3.4.1 Capital investments

The GPU manufacturing industry is a capital-intensive industry. GPUs are silicon layered with tantalum and palladium transistors and capacitors for storage and perform operations like addition and multiplication and copper, boron, cobalt, and tungsten for starters. The aluminum, tantalum, and tungsten need to be heavily processed to be usable in the GPU. Procurement and processing of these natural metals is high-cost. The GPU manufacturing process has multiple phases including

wafer fabrication, assembly, testing, and packaging. Each of these phases consumes a lot of resources and needs a high capital investment.

NVIDIA utilizes a contracting manufacturing strategy for all phases of the manufacturing process to utilize the expertise of industry-leading suppliers that are certified by the International Organization for Standardization. In the 2024 10-K report, NVIDIA and AMD reported property, plant, and equipment of $3,914 million and $1,589 million respectively.

	Jan 28, 2024
	(In mil
Property and Equipment:	
Land	$ 218
Buildings, leasehold improvements, and furniture	1,816
Equipment, compute hardware, and software	5,200
Construction in process	189
Total property and equipment, gross	7,423
Accumulated depreciation and amortization	(3,509)
Total property and equipment, net	$ 3,914

Figure 6: NVIDIA Property, Plant and Equipment

Property and Equipment, net	December 30, 2023
	(In mil
Land, building and leasehold improvements	$ 821
Equipment	2,346
Construction in progress	209
Property and equipment, gross	3,376
Accumulated depreciation	(1,787)
Total property and equipment, net	$ 1,589

Figure 7: AMD Property, Plant and Equipment

As seen above, both companies have spent more than a billion dollars on land, building, and leasehold equipment and more than two billion dollars on equipment. The capital-intensive nature of the GPU manufacturing industry acts as a barrier to entry for new firms.

3.4.2 Development and Research

Innovation is at the core of all businesses of NVIDIA. The GPU is a highly parallel processor architecture, composed of processing elements and a memory hierarchy. Designing a general processing unit starts with a document called product requirement specification (PRS) which outlines in detail the features the new GPU must have and what tasks the GPU should perform. Generally, it takes around six months to complete the PRS with thousands of engineers along with management staff being involved in the process. After the PRS is complete, more teams of engineers work on two distinct areas of design namely GPU architecture/floorplan and coding the circuit.

The above two steps need a lot of development and research. As of the end of fiscal year 2024, NVIDIA had approximately 29,600 employees in 36 countries out of which 22,200 were engaged in research and development. It is reported about 14% of the revenue is spent on research and development by NVIDIA which amounts to $8,675 million. AMD also spent $5,872 million on R&D in 2024.

3.4.3 Access to Supply and Distribution Channels

GPU manufacture involves the procurement and transport of raw materials from source to destination. It is essential to effectively distribute the products to reduce the costs. NVIDIA utilizes a contracting manufacturing strategy where they employ and partner with key suppliers that are certified by the International Organization for Standardization for all phases of the manufacturing process. This reduces the significant costs and risks associated with owning and operating manufacturing operations. The suppliers are responsible for the procurement of raw materials and processing and NVIDIA handles only the procurement of certain raw materials. To produce semiconductor wafers, they utilize foundries such as Taiwan Semiconductor Manufacturing Company Limited, TSMC, Samsung Electronics Co., Ltd., or Samsung, and purchase memory from Micron Technology, Inc., SK Hynix Inc., and Samsung. They engage CoWoS technology for semiconductor packaging and independent subcontractors and contract manufacturers such as Hon Hai Precision Industry Co., Ltd., Wistron Corporation, and Fabrinet to perform assembly, testing, and packaging of our final products. AMD and Intel also have similar distribution channels.

If firms were to enter the GPU manufacturing industry, they would need to access these distribution channels also to reduce costs and produce their products efficiently. Since NVIDIA and other existing companies have a large share of the distribution channels, they also have large control over the channels. New entrants would be charged higher to access these channels and also would have more restrictive contracts which would decrease their profit margins by large.

3.4.4 Intellectual property and Patents

NVIDIA has a total of 15034 patents globally with 11490 patents being currently active related to their products and the technology used in connection with our products.

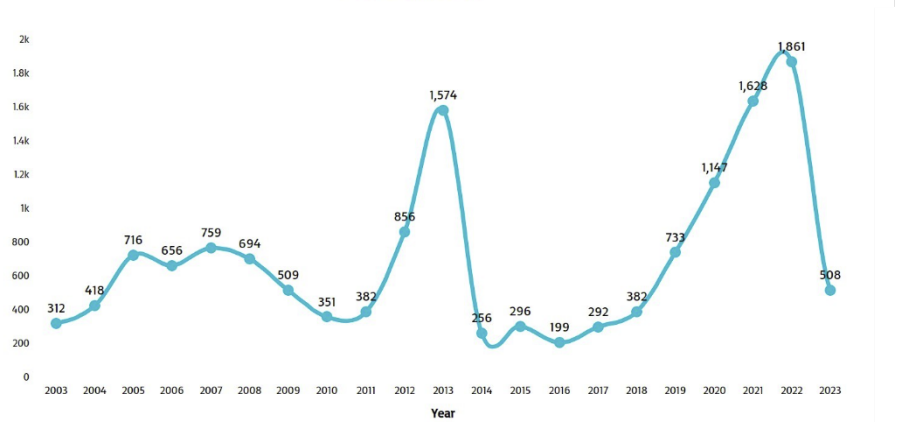


Figure 8: NVIDIA Patent Portfolio - Patent filing trend

NVIDIA relies on a combination of patents, trademarks, trade secrets, employee and third-party nondisclosure agreements, and licensing arrangements to protect our IP in the United States and internationally and confer a competitive advantage. With NVIDIA's development and research, they have patented products that are superior, and efficient, provide better functionality, and have more design wins than its competitors. Patents, trademarks, portfolios, and intellectual property cements NVIDIA's dominance in the market.

3.4.5 Government Regulations and Licensing Requirements

Government regulations and licensing requirements could potentially increase or decrease the barrier for new entrants. When more standards and conditions are set by the local government and the International Organization for Standardization, the cost to meet these standards increases. The high costs related to meeting and maintaining these standards could increase the barrier for new entrants.

On the other hand, Government regulations and licensing requirements could also affect the supply of GPU products. Licensing requirements were announced by the U.S. government during the third quarter of fiscal year 2023 that impact exports of A100 and H100 integrated circuits or any systems/boards that incorporate either of these two circuits to China and Russia. Additional licensing requirements were added for certain customers and regions including some countries in the Middle East. Going forward, more regulations and restrictions could adversely affect the export of GPU products to these regions. When imports are banned in certain regions, local producers in these regions would be promoted/motivated with incentives to manufacture GPUs for their local markets. After the local producers establish themselves as a profitable firm in their respective local markets, they would look into expanding their business. As these local manufacturers grow and enter the global market, they could become potential competitors to NVIDIA. Thus, Government regulations and licensing requirements could also potentially lower the barrier for new entrants.

3.5 Substitute offerings

Substitutes play a key role in potential profitability within the market. When customers find substitutes for the products/services, with an increase in the price of the product, customers would move to substitutes. With customers moving to substitutes, the profit margin of the company would decrease.

NVIDIA's Graphics segment includes GPUs for enterprise workstation graphics, gaming, cloud-based visual, and virtual computing with gaming being the main target. Gaming is the largest entertainment industry with PC gaming as the predominant platform [2]. Gaming platforms use NVIDIA GPUs to improve the gaming experience with smoother, higher-quality graphics. NVIDIA GPUs for the gaming market include GeForce RTX and GeForce GTX GPUs for gaming desktops and laptop PCs, GeForce NOW cloud gaming for playing PC games on underpowered devices, as well as SoCs and development services for game consoles[2]. There are many substitutes for gaming for entertainment purposes namely online streaming, movies, art, dance, music, sports, etc.. Technology is changing the way we consume information and entertainment. Due to the low cost of online streaming services like Netflix, Disney, Amazon Prime. Due to its affordability, convenience, and scalability, online streaming has grown popular. Video streaming is expected to grow by 8.27% with a projected revenue of $108 billion in 2024.

Though video streaming and other substitutes are growing popular, the threat of these substitutes is offset by the gaming industry leveraging the growing popularity of its substitutes and shifting away from its traditional roots. The games industry was originally about buying games and playing them as an individual unit. Currently, the gaming industry has branched out and merged with video streaming and sports. Game streaming has become increasingly popular with many gamers moving towards streaming themselves playing games. The growth of game streaming and the development of online multiplayer competitive games has led to the skyrocketing growth of the eSports industry. Esports is a form of competition using video games with organized tournaments where players and teams compete for a specific goal/price. Twitch launched in 2011 as a live streaming service focused on gaming and eSports hit 250 million active users on their website

with 7 million streamers publishing content monthly[25].

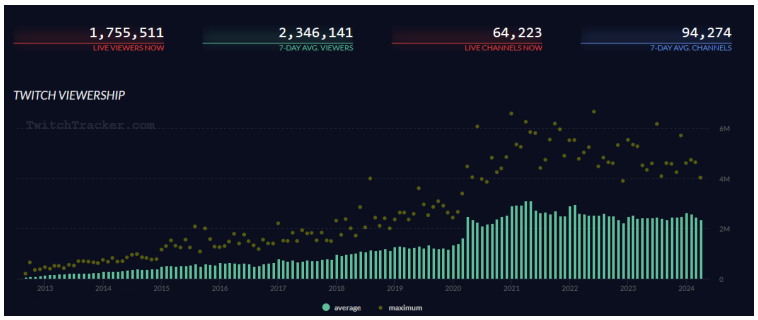


Figure 9: Twitch usage for games streaming on 21^{st} April 2024[11]

While a graphic processing unit is a processor for graphic processing operations, the central processing unit (CPU) is a hardware component that's the core computational unit in a server. CPUs have integrated graphics which are built into the processor and use system memory that is shared with the CPU. Discrete graphics cards are the graphic cards that NVIDIA manufactures while the integrated graphic cards are manufactured by AMD and Intel and embedded in its CPU. The integrated graphics are similar in function to discrete graphics and the availability of the close substitute acts as a potential threat to the GPU manufacturing industry. When potential close substitutes are available, customers might move towards the substitutes and the threat of substitutes is high.

As discrete graphics are separate from the processor, they consume more power and generate significant heat. Integrated graphics typically use less power than discrete graphics, create less heat, and have longer battery life. Though integrated graphics are cheaper, discrete graphic cards provide higher performance and better graphics visuals than their substitute. Many gamers, especially professional gamers, would prioritize performance over price. As the relative performance is higher for GPUs which more than compensates for its costs, currently the threat of substitutes is low. With development and research, the threat of substitutes would increase if integrated graphics were to become on par with discrete graphics.

GPUs excel in parallel processing through several cores or arithmetic logic units (ALU) as they simply take a high volume of the same instructions and push them through at high speeds[19]. Hence they are used to mine cryptocurrencies and used for AI models. However as they were initially designed for 3-D gaming acceleration, they also can be more efficient[20]. As discussed earlier, many companies are working on algorithm-specific AI chips. Tensor Processing Unit (TPU) is an AI accelerator application-specific integrated circuit (ASIC) developed by Google for neural network machine learning, using Google's own TensorFlow software. Such algorithm-targeted AI chips could potentially affect demand for NVIDIA GPUs in the long run. To keep pace with its competitors and continue its dominance in the chip market, NVIDIA continues to research and develop AI chips and also on how to optimize their GPUs. The recent launch of their Blackwell AI chip, "the world's most powerful chip as cited in [14]", is a testament to NVIDIA's efforts to understand their substitutes and remain on top of the GPU market.

4 Organization Future

4.1 Overview

NVIDIA has been the lead innovator for GPU technology since the 90s. A trend we will not likely see shift in the coming years. As our society becomes more data and computationally-driven, companies that provide the products required for these processes are in higher demand. Processing Units have become essential to modern life. Since NVIDIA is currently the third largest US company in market capitalization, it will have the resources to keep its market advantage.

4.2 Innovation Theory of Profit

The GPU (Graphical Processing Unit) is the visual processor used in most computers. GPUs have a vital role in the performance of specific computer applications. Software that utilized GPUs became more advanced, forcing NVIDIA to innovate frequently. Developing faster and higher-quality processing units year after year, satisfying consumer market demand, made NVIDIA an industry leader and a company that can aggressively create new processing units. NVIDIA used these techniques to tackle big data problems—systems like Cuda enable users to run programs on NVIDIA GPUs with increased performance. NVIDIA employs the Innovation Theory of Profit and has demonstrated its ability to do that faster than its competitors.

4.2.1 AI and Data Centers

NVIDIA has leveraged its skills and expertise to develop products for larger companies. In the data era, the world's computational demands are increasing daily. Big data centers are facilities where servers run platforms like Facebook, bank transactions, data analytics, and AI. These centers require computers that can handle extensive data efficiently. NVIDIA innovated existing technology to develop products that meet new market demand.

Artificial Intelligence (AI) is a breakthrough in computational ability that enables developers to tackle complex problems. Machine learning uses large data sets to train models to find patterns in data. This process requires powerful computers and is computationally intensive. NVIDIA created chipsets that were specially designed for AI problems. This innovation will help NVIDIA capitalize on the rise of AI. NVIDIA Blackwell Architecture, released on 03/18/24, is "the world's largest GPU, built with the specific purpose of handling data center-scale generative AI workflows with up to 25X the Energy Efficiency of the prior NVIDIA Hopper GPU generation". This GPU has over "208 billion transistors"[14].

4.2.2 Consumer Products

Computer enthusiasts recognize NVIDIA as the best GPU manufacturer. As mentioned previously, 70% of PC enthusiasts were using NVIDIA GPUs. NVIDIA and AMD are the only options for consumer-grade GPUs. The crowd favorite has proved that their team can deliver fast and better quality products. With their 10,20,30,40 series graphics cards, consumers can choose from a wide range of products that fit each system's requirements.

4.3 NVDA Stock Overview

On April 29, 2019, NVIDIA (NVDA) stock was worth $45.50 per share. Since then, NVDA price reached an all-time high of $942.25 on March 18, 2024. A total gain of around 2,000% in 5 years, or $1,000 invested into $20,708. This explosion in stock price has garnered the attention of everyone

that tracks financial markets. Investors have been riding high on the stock and expect the market to increase. NVDA's market capitalization is $2 Trillion, the third-largest American company.

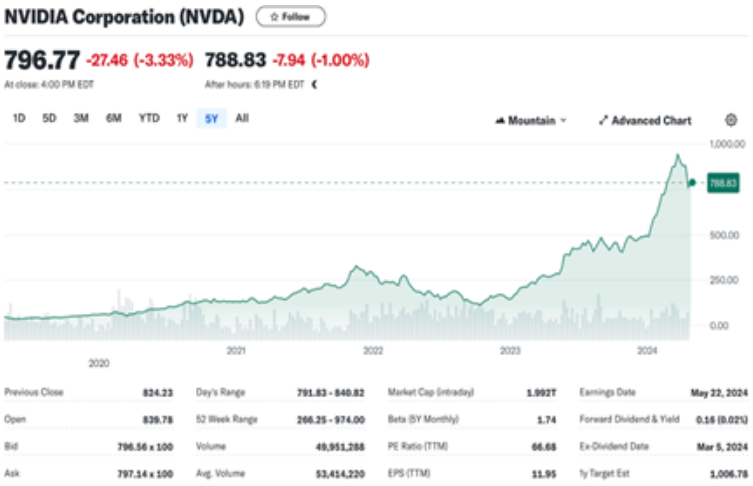


Figure 10: NVDA stock chart

Mickensy & Company projects the semiconductor industry's "aggregate annual growth could average from 6 to 8 percent a year up to 2030". The growth in the overall market will increase demand for NVDA products. NVDA will attempt to capitalize on the increasing value of the market. Due to the performance of tech stocks over the last two decades, portfolio managers will continue to invest in companies like NVDA to increase clients' exposure to the market. Jensen Huang, founder/CEO of NVDA, has grown this company faster than almost any other. Under Haung's guidance, NVDA should be able to keep using the strategies that have made them successful. Investors expect NVDA to grow and be the market leader.

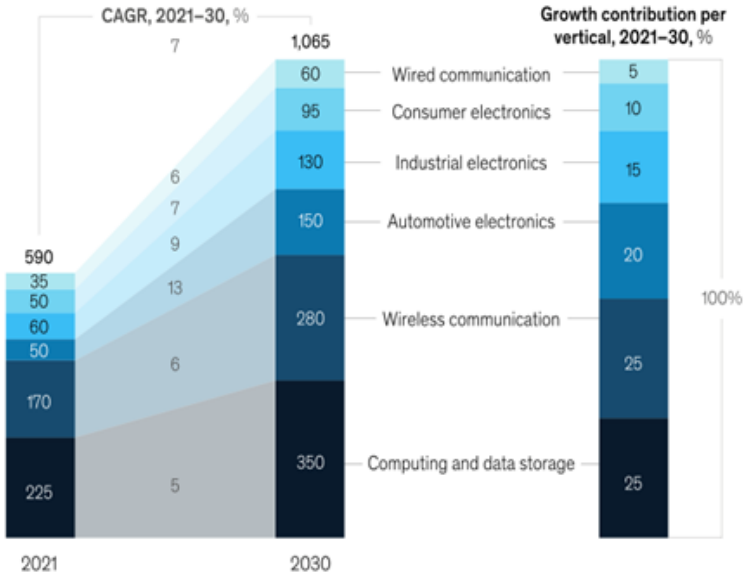


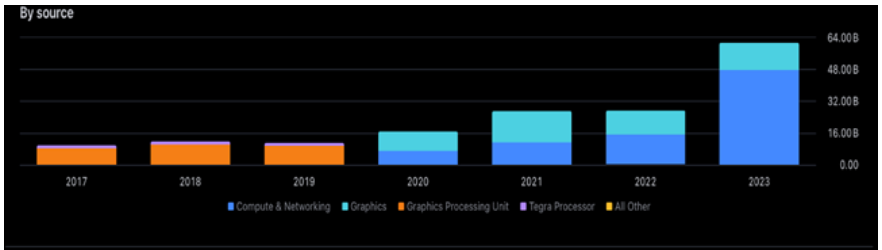
Figure 11: Global semiconductor market value by vertical, indicative, $ billion [18]

4.3.1 Financial Statements

NVDA in 2019 recorded $10.92 billion in Total Revenue, $6.78 billion in Gross Profit, and $2.80 billion in Net Income. NVDA in 2023 recorded $60.92 billion in Total Revenue, $44.30 billion in Gross Profit, and $29.76 billion in Net Income.

Currency: USD	2017 Jan 2018	2018 Jan 2019	2019 Jan 2020	2020 Jan 2021	2021 Jan 2022	2022 Jan 2023	2023 Jan 2024	TTM
Total revenue	9.71B +40.58%	11.72B +20.61%	10.92B -6.81%	16.68B +52.73%	26.91B +61.40%	26.97B +0.22%	60.92B +125.85%	60.92B
› Cost of goods sold	-3.88B	-4.50B	-4.14B	-6.28B	-9.44B	-11.62B	-16.62B	-16.62B
Gross profit	5.84B +43.21%	7.21B +23.59%	6.78B -5.96%	10.40B +53.27%	17.48B +68.09%	15.36B -12.13%	44.30B +188.49%	44.30B
› Operating expenses (excl. COGS)	-2.61B	-3.36B	-3.89B	-5.78B	-7.43B	-9.78B	-11.33B	-11.33B
Operating income	3.22B +63.69%	3.85B +19.45%	2.89B -24.91%	4.62B +59.67%	10.04B +117.53%	5.58B -44.46%	32.97B +491.21%	32.97B
› Non-operating income, total	-27.00M	46.00M	79.00M	-207.00M	-100.00M	-1.40B	846.00M	847.00M
Pretax income	3.20B +67.77%	3.90B +21.90%	2.97B -23.77%	4.41B +48.45%	9.94B +125.47%	4.18B -57.94%	33.82B +708.85%	33.82B
Equity in earnings	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
› Taxes	-149.00M	245.00M	-174.00M	-77.00M	-189.00M	187.00M	-4.06B	-4.06B
Non-controlling/minority interest	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
After tax other income/expense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
Net income before discontinued operations	3.05B	4.14B	2.80B	4.33B	9.75B	4.37B	29.76B	29.76B
Discontinued operations	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
Net income	3.05B +82.89%	4.14B +35.90%	2.80B -32.48%	4.33B +54.94%	9.75B +125.12%	4.37B -55.21%	29.76B +581.32%	29.76B

(a)



Currency: USD	2017	2018	2019	2020	2021	2022	2023
Compute & Networking	–	–	–	6.84B	11.05B	15.07B	47.41B
Graphics	–	–	–	9.83B	15.87B	11.91B	13.52B
Graphics Processing Unit	8.14B	10.18B	9.47B	–	–	–	–
Tegra Processor	1.53B	1.54B	1.45B	–	–	–	–
All Other	43.00M	–	–	–	–	–	–

(b)

Figure 12: NVIDIA's Total Revenue, Gross Profit, Net Income over the last 4 years

In 2019, NVDA had two revenue sources: GPU ($9.47 billion) and Tegra Processor ($1.45 billion). In 2020, NVDA changed its Revenue Sources. GPU was included in Graphics, and Tegra Processor was integrated into Compute and Networking. By 2023, the Total Revenue was $47.41 billion for Compute & Networking and $13.52 billion in Graphics. In 2022, NVDA's Net Income was $4.37 billion. In one year (2023), NVDA increased its Net Income to $29.67 billion +581.32% [21]. NVDA's drastic increase in Total Revenue, Gross Profit, and Net Income over the last four years demonstrates its ability to make profitable products. In 2020, the company was positioned

to target data/AI-related customers. Since that switch, "Compute and Networking" has eclipsed NVIDIA's legacy products (GPU) with $34 billion more in revenue in 2023.

4.4 Moore's Law

The observed number of transistors on computer chips doubles every two years, this is known as Moore's Law. In 1965, Gordon Moore (co-founder of Intel) observed this pattern and graphed a prediction of transistor growth. Moore's hypothesis was meant to observe growth 10 years in the future. Since then, Moore's law has "held true" for over half a century. This growth in transistors is correlated with the innovation progression NVIDIA has demonstrated. As mentioned in Section 4.2.1, the NVIDIA Blackwell Chip has 208 billion transistors. The chart above ends in 2020, at 50 billion transistors. If the principle continues to hold true, NVIDIA should be able to continue to innovate at exponential rates.

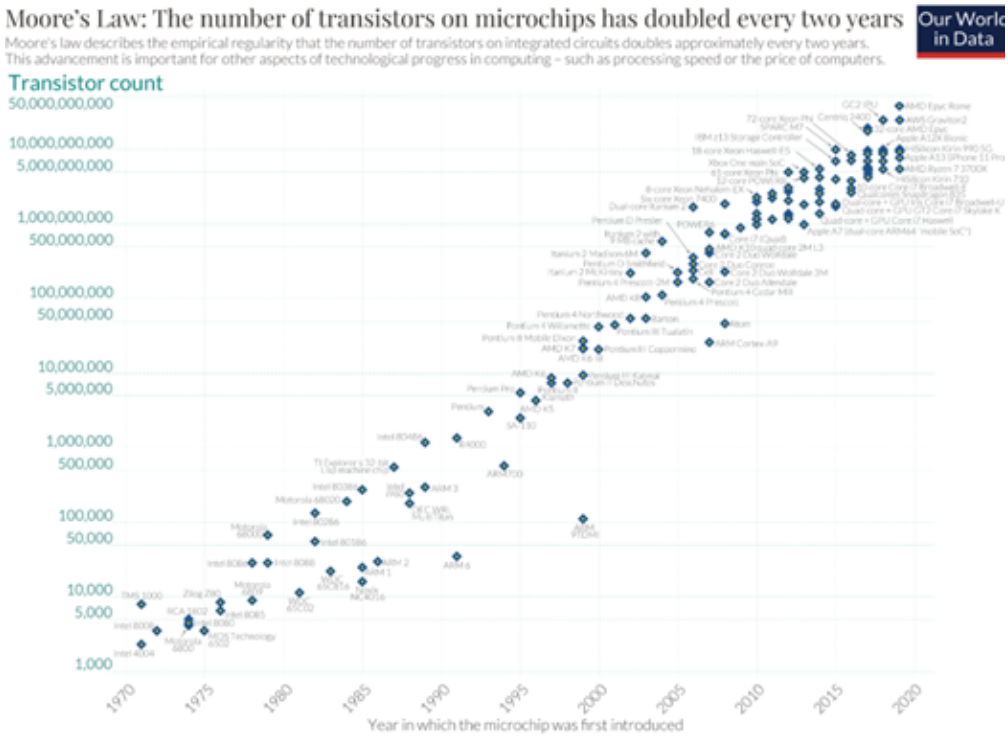


Figure 13: Number of transistors on microchips over the years [1]

4.5 Competitors

NVIDIA competes against well-established competitors on all of its products. AMD is its biggest competitor for consumer GPUs. Intel, and other tech companies, compete to bring products to market that rival NVIDIA data processing units. Though NVIDIA has a market advantage, the company must closely monitor its competitors.

4.5.1 AMD

The GPU market has only two options: AMD and NVIDIA. Unlike NVIDIA, AMD produces consumer CPUs (Central Processing Units) and has been in a tight battle against Intel. Over the

last five years, AMD has developed CPUs that rival Intel processors. If AMD can translate that success, they could compete with NVIDIA on consumer GPUs and AI/Data processing units.

4.6 International Global Chip Conflict

The world has become reliant on computers. Technology companies are now the backbone of our developing society, and this relationship has become a national security risk for all nations.

4.6.1 United States Policy

The United States has become a global superpower in part due to the innovation of computers and technology. Companies like NVIDIA and Intel have allowed the US to be at the forefront of science, military, economy, and modernization. The United States must maintain chip dominance to be competitive against the rest of the world. The US must decrease its dependency on processing units manufactured outside the US. President Biden passed the Chip Act, which incentivizes manufacturing processing units in the US. ASML, TSCM, and Intel are now building factories in the US because of government incentives. The government understands that processing units are essential to national security and is investing to secure the supply chain [23].

4.6.2 China's Policy (Chinese Communist Party)

BRI (Belt and Road Initiative) is a program the CCP has created to increase their trade, economic, and national security interests. The BRI creates infrastructure projects worldwide, particularly connecting East Asia to Europe. The CCP is attempting to increase trade, economic, and military opportunities by creating projects to build infrastructure along this belt: roads, ports, electricity, cellular towers, etc. While doing this, the CCP can leverage and utilize its currency to fund these projects. Taiwan is among the countries threatened by CCP expansion. If the CCP were to take control of trade in Taiwan, the global chip shortage would be exacerbated, and other countries would not have the resources to continue modernization[30].

4.7 TSMC and ASML Critical Role

ASML and TSMC use proprietary technology, lithography, to develop semiconductors for processing units. These companies' performance and security will have implications on the whole market. ASML is a Netherlands-based company that sells lithography machines to manufacturers such as TSMC. TSMC then manufactures chips based on the designs from companies like NVIDIA and Apple. "NVIDIA announced that TSMC and Synopsys are going into production with NVIDIA's computational lithography platform to accelerate manufacturing and push the limits of physics for the next generation of advanced semiconductor chips". NVIDIA relies on TSMC's manufacturing processes to make its processing units and ASML's lithography.

4.8 Supply Chain

The supply chain for NVIDIA to design and manufacture any Processing Units requires collaboration from multiple companies. NVIDIA designs chips that are then manufactured by companies, like TSMC. The complex supply chain can add tightened stress when market demand is high. This reliance amongst corporations is not uncommon but can be concerning when projecting future stability[36].

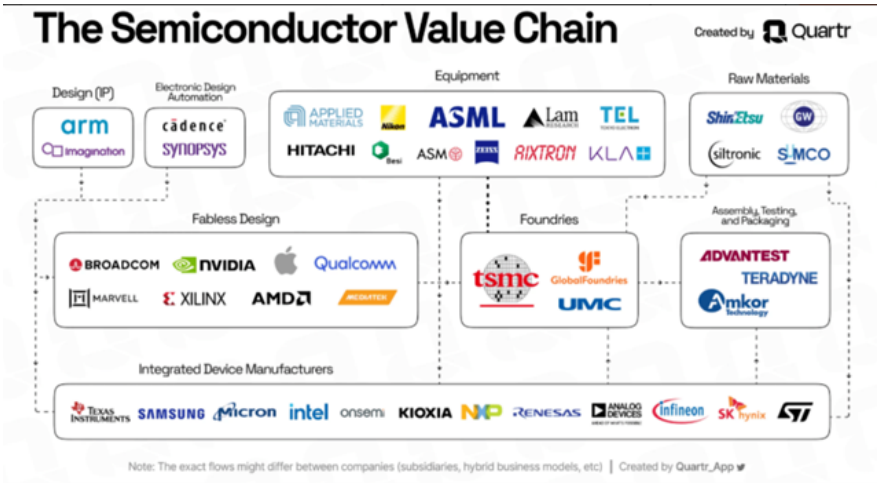


Figure 14: Semiconductor value chain

4.8.1 Raw Materials

Securing, mining, refining, and transporting raw ore is one of the first weak points in the supply chain. The locations where raw ore is located can be scarce and isolated. Depending on geopolitical issues, these locations can be compromised. The mining processes used to extract ore are not friendly to the environment or workers. Around the world, humanitarian rights violations have plagued mining operations in underdeveloped regions. Refining raw ore often uses processes that could be more environmentally sustainable. Once the raw ore is mined and refined, the material must be shipped to locations worldwide, furthering environmental impact.

4.8.2 Factories

TSMC currently does not have enough factories worldwide to be de-risked from the threat of geopolitical issues. If Tawain were to be cut off from Western countries, the supply of Processing Units would not be able to sustain the demand and growth of these nations, affecting the whole market, including NVIDIA. The assembly locations also require significant capital investment and slow build times. Fortunately, countries and investors around the world have supported plans to build facilities and increase production.

5 Advice to CEO

The CEO/Founder of NVIDIA, Jensen Huang, has a trusted and proven track record of excellence. Huang has a great rapport with reporters and investors. The CEO's frequent public appearances bring confidence to the market. The company is in good hands even though NVIDIA will eventually have to separate from its founder.

A recommendation for the CEO would be to increase supply chain security by integrating the TSMC manufacturing process. A partnership at a joint facility would increase stability, costs, and geopolitical constraints. NVIDIA could also purchase lithography machines from ASML to integrate into their manufacturing process. Both options would give NVIDIA more control over an essential process for their products.

Many upcoming startups in the data center market namely SambaNova, Cerebras, GraphCore, Groq, and xAI are working on processing large language models (LLMs) and other elements of AI[20]. Another recommendation to the CEO would be to invest and form strategic partnerships with such startups. Partnering with startups helps NVIDIA explore new technologies and ideas that may complement or enhance its existing technologies, thus obtaining access to innovation, expanding its market, and spreading risk by sharing resources, knowledge, and expertise.

NVIDIA has taken the world by storm and developed essential technology for emerging markets. Modernization across the world is rapidly expanding. Companies that fuel modernization can expect increasing demand. NVIDIA has the financial resources to spend more on R&D than its competitors while also elevating the capital requirement needed to enter the market. The financial statements demonstrate a corporate ability to innovate, profit, and shift in a developing market. NVIDIA has increased profits, revenue, net income, and market capitalization while increasing gross margin by 13% since 2019. Nations are creating incentives to increase the global and national supply of processing units. This shift in policy will benefit companies like NVIDIA and TSMC. The global conflict for processing units has no resolution and could threaten growth if the issue escalates. NVIDIA can develop products that exceed the demanding market requirements. If NVIDIA can keep its team together and perform at the standard they set, the company will continue to be the market leader.

Bibliography

[1] Max Roser, Hannah Ritchie, and Edouard Mathieu. "What is Moore's Law?" In: *Our World in Data* (2023). https://ourworldindata.org/moores-law.

[2] NVIDIA Corporation. *NVIDIA. Form 10-K 2024. U.S. Securities and Exchange Commission.* https://www.sec.gov/ix?doc=/Archives/edgar/data/1045810/000104581024000029/nvda-20240128.htm. Jan. 2024.

[3] *Qualcomm: Automotive Segment Driving Market Domination In ADAS.* https://seekingalpha.com/article/4680579-qualcomm-automotive-segment-driving-market-domination-in-adas. Mar. 2024.

[4] URL: https://docs.nvidia.com/deeplearning/performance/dl-performance-gpu-background/index.html#gpu-perf.

[5] URL: https://www.forbes.com/sites/forbestechcouncil/2023/08/21/shaping-the-future-of-gaming/?sh=12864db67e08.

[6] URL: https://www.investopedia.com/terms/p/porter.asp.

[7] URL: https://store.steampowered.com/hwsurvey/Steam-Hardware-Software-Survey-Welcome-to-Steam.

[8] URL: https://www.designlife-cycle.com/nvidia-gpu.

[9] URL: https://beonair.com/video-streaming-skyrocketing-in-popularity#:~:text=Technology%20is%20changing%20the%20way,also%20shifting%20to%20mobile%20platforms.

[10] URL: https://www.statista.com/outlook/dmo/digital-media/video-on-demand/video-streaming-svod/worldwide.

[11] URL: https://twitchtracker.com/statistics.

[12] URL: https://www.investopedia.com/terms/m/macroeconomic-factor.asp.

[13] URL: https://www.cfr.org/backgrounder/chinas-massive-belt-and-road-initiative.

[14] URL: https://resources.NVIDIA.com/en-us-blackwell-architecture.

[15] URL: https://www.jonpeddie.com/news/shipments-of-graphics-add-in-boards-increase-for-third-quarter-in-a-row/.

[16] URL: https://www.reuters.com/technology/meta-deploy-in-house-custom-chips-this-year-power-ai-drive-memo-2024-02-01/.

[17] URL: https://www.digitaltrends.com/computing/amd-vs-intel-vs-NVIDIA-i-dont-care-anymore/?amp.

[18] URL: https://www.mckinsey.com/industries/semiconductors/our-insights/the-semiconductor-decade-a-trillion-dollar-industry.

[19] URL: https://aws.amazon.com/compare/the-difference-between-gpus-cpus/.

[20] URL: https://www.networkworld.com/article/1306554/startups-pursue-gpu-alternatives-for-ai.html.

[21] URL: https://finance.yahoo.com/quote/NVDA/financials.

[22] URL: https://finance.yahoo.com/news/customer-demand-nvidia-chips-far-013826675.html.

[23] *CHIPS for America Act FABS Act.* https://www.semiconductors.org/chips/.

[24] CNBC. *How Apple makes its own chips for iPhone and Mac, edging out Intel.* https://www.cnbc.com/2023/12/01/how-apple-makes-its-own-chips-for-iphone-and-mac-edging-out-intel.html.

[25] Brian Dean. *Twitch Usage and Growth Statistics: How Many People Use Twitch?* https://backlinko.com/twitch-users.

[26] R. Dow. *Shipments of graphics add-in boards increase for third quarter in a row.* https://www.jonpeddie.com/news/shipments-of-graphics-add-in-boards-increase-for-third-quarter-in-a-row.

[27] *GPU Performance Background User's Guide.* https://docs.NVIDIA.com/deeplearning/performance/dl-performance-gpu-background/index.html#gpu-perf.

[28] Mike Isaac. *NVIDIA's Big Tech Rivals Put Their Own A.I. Chips on the Table.* https://www.nytimes.com/2024/01/29/technology/ai-chips-nvidia-amazon-google-microsoft-meta.html.

[29] Ian King. *NVIDIA gives upbeat forecast as AI reaches 'tipping point.'* https://www.latimes.com/business/story/2024-02-21/NVIDIA-gives-upbeat-forecast-as-ai-reaches-tipping-point.

[30] Lindsay Maizland and Andrew Chatzky. *China's Massive Belt and Road Initiative.* https://www.cfr.org/backgrounder/chinas-massive-belt-and-road-initiative.

[31] Taiwan Semiconductor Manufacturing. *TSMC Fabs.* https://www.tsmc.com/english/aboutTSMC/TSMC_Fabs.

[32] S. McDowell. *Is NVIDIA Lagging In High-Growth Automotive Segment?* https://www.forbes.com/sites/stevemcdowell/2024/02/22/is-nvidia-lagging-in-high-growth-automotive-segment/?sh=5af7c7495992.

[33] J. Nguyen. *What you need to know about NVIDIA and the AI chip arms race.* https://www.marketplace.org/2024/03/08/what-you-need-to-know-about-NVIDIA-and-the-ai-chip-arms-race/.

[34] J. Norem. *Analysts Estimate NVIDIA Owns 98% of the Data Center GPU Market.* https://www.extremetech.com/computing/analysts-estimate-NVIDIA-owns-98-of-the-data-center-gpu-market.

[35] NVIDIA Corporation. *CORPORATE SOCIAL RESPONSIBILITY (CSR) DIRECTIVE.* https://www.nvidia.com/content/dam/en-zz/Solutions/about-us/documents/nvidia-corporate-responsibility-directive.pdf.

[36] Amit Peretz. *TSMC Vs. ASML: Buy The Semiconductor Supply Chain (NASDAQ:ASML).* https://seekingalpha.com/article/4651969-tsmc-vs-asml-buy-the-semiconductor-supply-chain.

[37] Amazon Press Release. *AWS-Unveils-Next-Generation-AWS-Designed-Chips.* https://press.aboutamazon.com/2023/11/aws-unveils-next-generation-aws-designed-chips.

[38] Charlotte Trueman. *AMD data center segment sets internal revenue records, as GPU sales exceed expectations.* https://www.datacenterdynamics.com/en/news/amd-data-center-segment-set-internal-sales-records-in-2023/.

[39] H. Wilson. *NVIDIA Facts and Statistics.* https://www.investing.com/academy/statistics/NVIDIA-facts-and-statistics/.