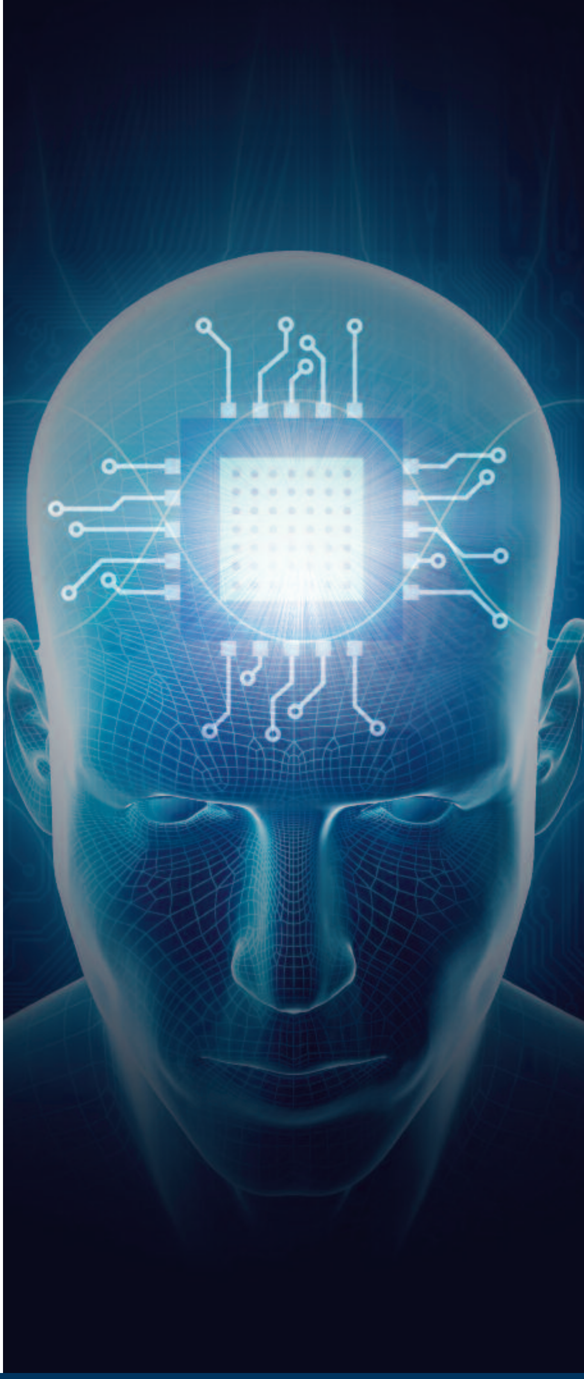
Goldman Sachs

Profiles in Innovation Revisited

AI Hardware

Mapping the $100bn IT spending opportunity



Artificial intelligence is reshaping the hardware landscape, driving demand for devices to accelerate application development and crowning new hardware leaders in the process. In the latest in our **Profiles in Innovation series** on AI's disruptive potential, we discuss how increased adoption of AI could translate to massive growth in the hardware market, as well as architectural and structural change. We believe continued growth in data availability, compute power, and the developer ecosystem is likely to support increased levels of hardware spend, resulting in a $109bn AI hardware market by 2025.

Leading the pack will be those whose architectures have emerged as standard compute accelerators for AI application development (GPUs, FPGAs, and ASICs at the expense of CPU manufacturers) as well as manufacturers of AI hardware. In addition, we expect memory and storage vendors to benefit.

Toshiya Hari
(646) 446-1759
toshiya.hari@gs.com
Goldman Sachs & Co. LLC

Charles Long
(212) 902-7494
charles.long@gs.com
Goldman Sachs & Co. LLC

Daiki Takayama
+81 3 6437-9870
daiki.takayama@gs.com
Goldman Sachs Japan Co., Ltd.

Donald Lu, Ph. D.
+86 10 6627-3123
donald.lu@ghsl.cn
Beijing Gao Hua Securities Company Limited

Mark Delaney, CFA
+1 212 357-0535
mark.delaney@gs.com
Goldman Sachs & Co. LLC

Rod Hall
+1 415 249-7437
rod.hall@gs.com
Goldman Sachs & Co. LLC

The Goldman Sachs Group, Inc.

Table of contents

Contributing Authors: Toshiya Hari, Charles Long, Daiki Takayama, Donald Lu, Mark Delaney, Doug Clark, Rod Hall, Heather Bellini, Heath Terry, Giuni Lee, Timothy Sweetnam, Chris Jeon, Balaji Krishnamurthy, Matt Fahey

PROFILES IN INNOVATION

See below for more related reports on Artificial Intelligence or visit our portal for the full coverage:

- Artificial Intelligence - AI, Machine Learning and Data Fuel the Future of Productivity, Nov 14, 2016
- Americas Technology: Semiconductors: Updating estimates to reflect AI report findings; reiterate Buy for XLNX (on CL) and NVDA, March 11, 2018
- China's Rise in Artificial Intelligence: The New New China, Aug 31, 2017
- Artificial Intelligence and Machine Learning: Key highlights from Intel's first AI Day, Nov 21, 2016
- Artificial Intelligence & Machine Learning: GS Technology & Internet Conference 2018 - What we learned, Feb 20, 2018

Portfolio Manager's summary

The proliferation of Artificial Intelligence and its implications for the overall technology landscape remain top of mind among investors. Below, we discuss how increased adoption of AI could translate to massive growth in the hardware TAM as well as architectural/structural change. Bottom line, we estimate the overall AI hardware TAM has potential to grow from $12bn in 2017 to $35bn/$100bn+ by 2020/2025 and would highlight **Nvidia (Buy)**, **Xilinx (Buy, Conviction List)** and **TSMC (Buy)** as potential beneficiaries in the Compute segment, as well as **Micron (Buy)**, **Samsung Electronics (Buy, Conviction List)** and **Hynix (Buy)** in Memory. Elsewhere, while the rising tide would allow **Intel (Neutral)** to grow AI-related revenue throughout the forecast period, we see limited potential upside to our modelled forecasts compared to peers as we expect the company is likely to remain a wallet share donor.

What's driving the change?

With everything from smartphones, to smartwatches, to smart meters spitting out data, the need to make sense of, and ultimately monetize, the world's "new oil" has never been stronger. Artificial Intelligence (AI) has been around for decades; however, limited access to large data sets and lack of appropriate computing architectures has restrained development in the field until recently. The emergence of *deep learning* and the adoption of parallel computing architectures have accelerated AI research and software development, driving a new wave of investment in talent, hardware, and infrastructure. This new wave of investment is geared towards leveraging the growing "reservoirs" of data across multiple cloud vendors, enterprises, and even small businesses in order to achieve such goals as increasing productivity and establishing competitive advantages.

What's the opportunity?

Companies are racing to be the first to develop sophisticated AI capabilities for a number of applications including speech, image, video, and text recognition. Demand for hardware that is capable of accelerating application development is also rising. Service providers have recognized such demand for acceleration and are increasingly investing in infrastructure to address it given the premium developers are willing to pay in order to gain access to such hardware. We believe continued growth of data, compute acceleration capability, and the developer ecosystem is likely to support increased levels of hardware spend going forward and, as a result, see AI-related hardware spend reaching over $100bn by 2025.

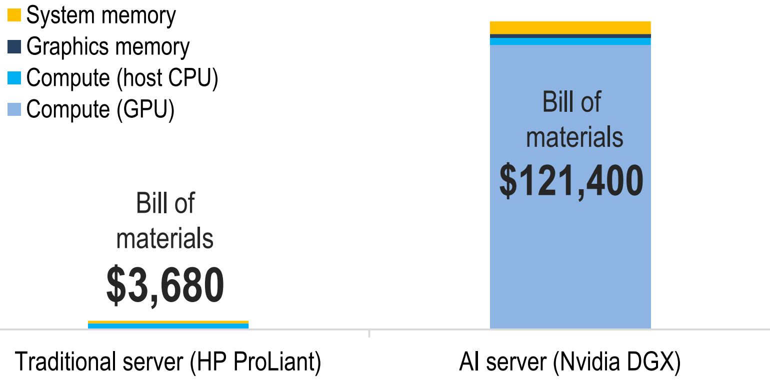
Who are the winners and losers?

Within the context of an expanding AI-hardware TAM, we see both leaders and laggards. Leading the pack are those whose architectures have emerged as standard compute accelerators for AI application development, namely Nvidia (with GPUs) and Xilinx, who's FPGAs are being adopted across multiple cloud vendors. We also expect the manufacturers of AI hardware to benefit and would highlight TSMC, who manufactures Nvidia's GPUs and Xilinx's FPGAs. In addition to the compute market, we expect memory and storage vendors to benefit as well. While most companies are likely to benefit from a rising hardware tide, we do expect some to fall behind, namely Intel, the leading CPU vendor, who we expect to be a net wallet share donor as the landscape shifts away from CPUs for AI training and inferencing.

Powering AI our hardware forecast, in brief

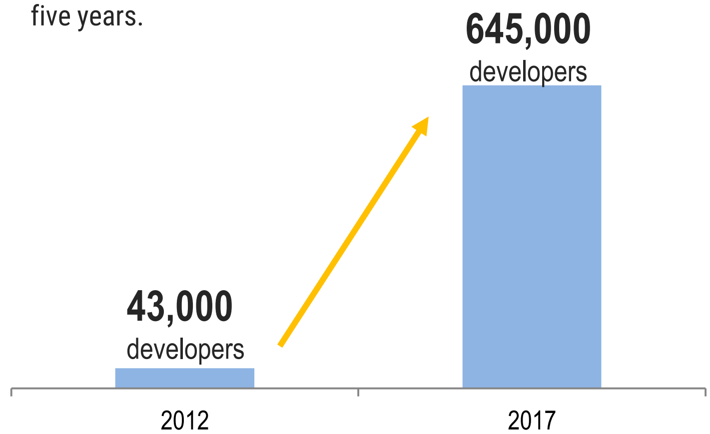
AI aspirations mean new hardware

The standard CPU isn't designed for complex AI workloads. More powerful (and more expensive) computers capable of parallel processing are needed for AI's heavy lifting. Cue the GPU—an electronic circuit fast emerging as the pillar of the AI hardware stack.



...and new developers

The developer base is evolving alongside the hardware stack. With greater access to large data sets and the opportunity to capture a piece of the AI economy, the number of programmers seeking to accelerate application development with GPUs has grown fifteen-fold in the last five years.

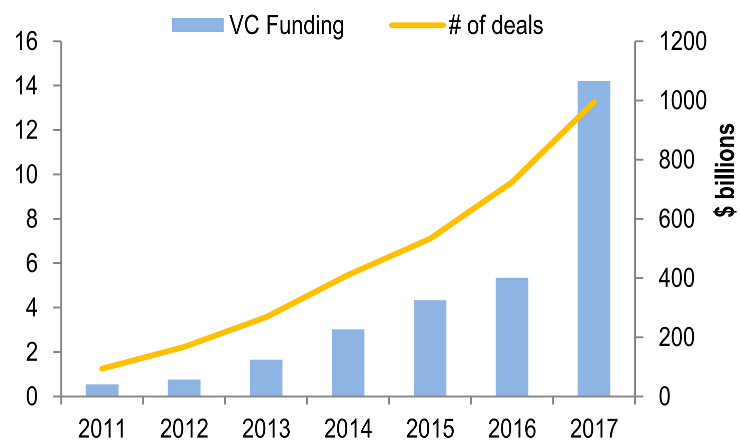


Funding pouring in

Cloud companies and venture capitalists are seizing on the interest in artificial intelligence.

$25BN → $76BN

We expect major cloud players (including Amazon, Facebook, Google, Microsoft, Alibaba, Tencent, Oracle and Salesforce) to spend a combined $76BN on capex by 2020, in part to build out their AI hardware infrastructure, vs. $25BN spent in 2014. VC dollars into AI development are also on the rise, more than doubling in the last two years (see chart on right).



Access priced at a premium

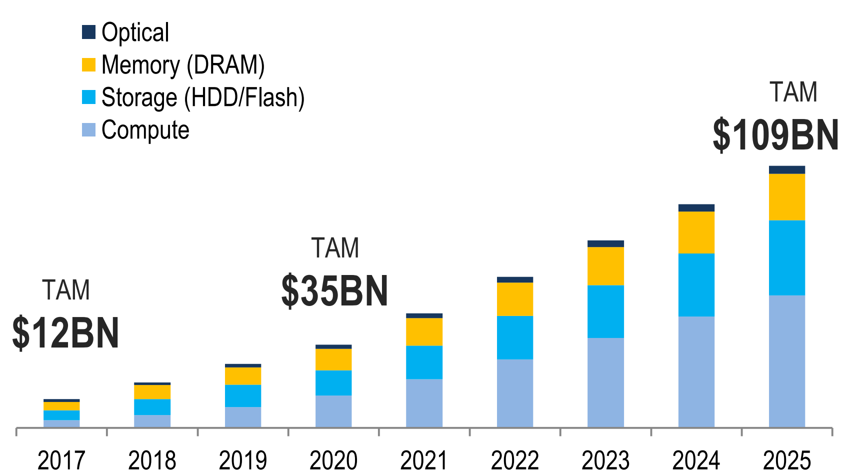
With so much demand for AI processing power, cloud companies are able to charge a premium for use of their hardware investments.

Google, for instance, charges $6.50 per hour for use of its application-specific Tensor Processing Unit (the high-end option for AI processing), vs. $1.60/hour for a mid-range Nvidia GPU and $0.06/hour for base CPUs.

$0.06→$1.60→$6.50

Price/hour for **standard CPU**	Price/hour for **Nvidia GPU** (Pascal)	Price/hour for **Google TPU**

The bottom line: A $109BN AI hardware market by 2025 (GSe)



The Ecosystem

Premium Artificial Intelligence Hardware - Key Players

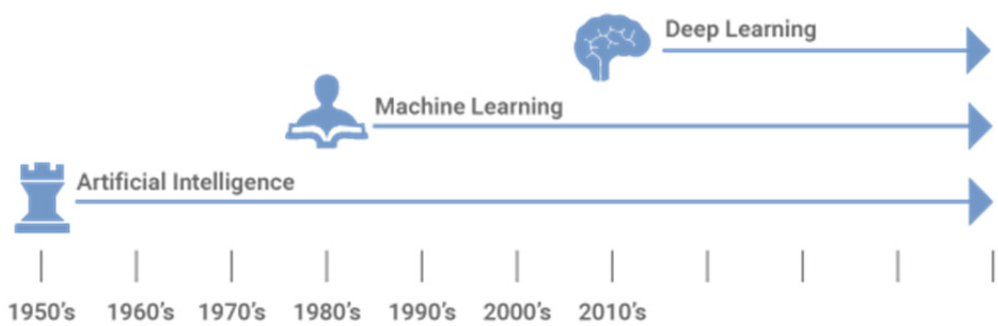
	Market Cap ($bn)	Key Tech	Market Share/Exposure
Compute ($55bn TAM)			
Nvidia	$146	GPU	~70% desktop GPU share
Xilinx	$19	FPGA	~60% FPGA share
Intel	$237	CPU, ASIC, FPGA	99% server CPU share, ~40% FPGA share
Broadcom	$101	ASIC	
Qualcomm	$92	CPU, SoC	~ 30% smartphone SoC share
AMD	$12	CPU/GPU	<1% server CPU share, ~30% desktop GPU share
Graphcore	Private	ASIC	
Cerebras	Private	ASIC	
Memory ($19bn TAM)			
Samsung	$329	DRAM	1% of revenue tied to datacenter AI
SK Hynix	$57	DRAM	4% of revenue tied to datacenter AI
Micron	$64	DRAM, other	4% of revenue tied to datacenter AI
Intel	$237	Other	
Storage ($31bn TAM)			
Samsung	$329	NAND	1% of revenue tied to datacenter AI
SK Hynix	$57	NAND	4% of revenue tied to datacenter AI
Micron	$64	NAND, other	4% of revenue tied to datacenter AI
Western Digital	$29	HDD, SSD	5% of revenue tied to datacenter AI
Seagate	$17	HDD, SSD	
Toshiba	$19	HDD, SSD	
Intel	$237	NAND, other	
Foundry ($16bn TAM)			
TSMC	$222		Market share at least 85% in 2018-20
Samsung	$329		
GlobalFoundries	Private		
Intel	$237		
Optical ($3bn TAM)			
Lumentum	$4	Optical components	
Finisar	$2	Optical components	
Acacia	$2	Optical components	
Networking			
Cisco	$214	Switches, routers, optical	
Broadcom	$101	Switches, routers	
Arista	$22	Switches	
Juniper	$9	Switches, routers	
Cavium	$6	Switches	
Mellanox	$4	Switches, adapters, interconnect	
Barefoot Networks	Private	Programmable switches	
Innovium	Private	Switches	

Note: Green highlights refer to leaders; Red highlights refer to laggards.

How is AI reshaping the hardware landscape?

Refresher: What is AI, what is Machine Learning, and how is all of this related to Deep Learning?

Exhibit 1: Deep learning is a subset of AI
AI timeline



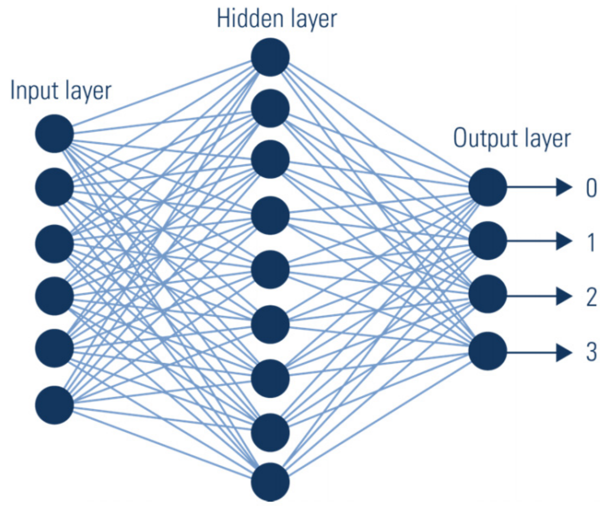
Source: Goldman Sachs Global Investment Research.

What is Artificial Intelligence? Artificial intelligence describes a science of simulating intelligent behavior in computers. It entails enabling computers to exhibit human-like behavioral traits including knowledge, reasoning, common sense, learning, and decision making. Classically, these include natural language processing and translation, visual perception and pattern recognition, and decision making, but the number and complexity of applications is rapidly expanding.

What is machine learning? Machine learning is a branch of artificial intelligence that reemerged in the 1980s and entails enabling computers to learn from data on their own. Traditionally a computer would be programmed to do something like recognize people in photos, but if it sees a photo of an object that only resembles a person (e.g. a baboon), a machine may falsely identify it as a human. Machine learning reduces the probability of this happening by using backpropagation (discussed later) to enable the computer to learn from far more data than any programmer could. By using a large set of examples of people and objects that only resemble people the computer is taught to better identify actual people, achieving a level of artificial intelligence.

Computers are able to learn various tasks due to neural networks (NNs), which is a type of software that simulates the structure of a human brain. Neural nets consist of connected nodes in aggregate that can solve more complex problems and learn, like the neurons in a human brain. The process of *backpropagation* is used in machine learning in order to adjust the weight of the neurons in a neural net, eventually strengthening the paths that produce a correct answer.

Exhibit 2: Neural networks simulate the structure of the brain
Multiple hidden layers would be characteristic of deep learning



Source: Michael A. Nielsen, "Neural Networks and Deep Learning", Determination Press, 2015, Goldman Sachs Global Investment Research

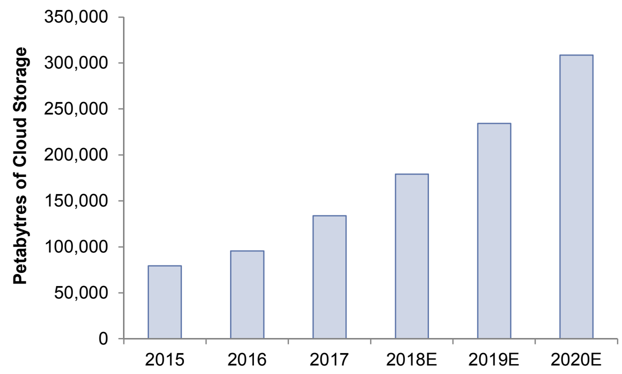
What is deep learning? Deep learning is a subset of machine learning that entails training a hierarchy of "deep layers" of large neural nets. While both Machine Learning and Deep Learning use the same underlying linear algebra, the defining characteristic is layer count, with Deep Learning systems applying >2 layers (think of layers as the filters in machine learning models, sorting the defining characteristics of the data that are being processed). Multiple layers are used to solve different aspects of a problem with model complexity increasing in relation to the difficulty of the problem at hand. For example, in order to recognize a person, a deep learning system would have a different layer for different features that define a person -- from having two eyes and a nose to walking upright on two legs -- with the program returning the result of "person" or "not a person" only after the layers collectively determine what is in a given image.

Going forward we expect AI development to shift more towards Deep Learning as we believe more complex problems, such as autonomous driving, are better addressed by machine learning models that are capable of taking into account more variables.

How is AI reshaping the hardware landscape?

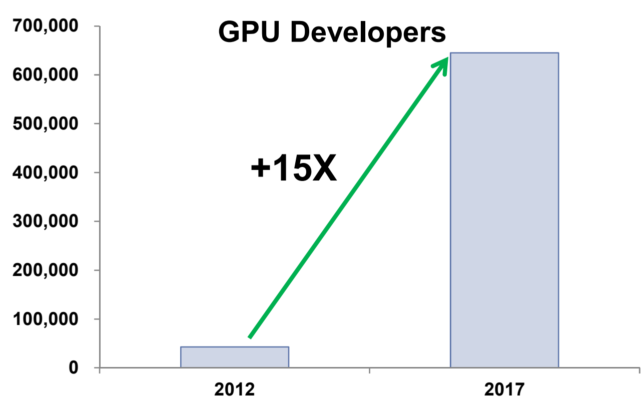
Why now? Before we dig into how AI is reshaping the hardware landscape, we should first discuss why the AI ecosystem is growing. AI research and application development has primarily been driven by cloud companies and academia, with spending disproportionately driven by cloud companies. A key driver of the growth of AI has been the availability of data which can be leveraged in order to develop and refine AI models, as evidenced in the growth of cloud storage (Exhibit 3). Growth in the availability of data has in turn contributed to an expansion of the AI ecosystem, particularly in the developer base and model complexity. Note the number of GPU developers has grown 15x from 2012-2017 and the number of downloads of Nvidia's CUDA parallel computing platform, which serves as the interface between AI models and the underlying hardware, has grown 5x over the same time period (Exhibit 5).

Exhibit 3: Cloud storage is expected to grow dramatically
Petabytes (PB) of cloud storage



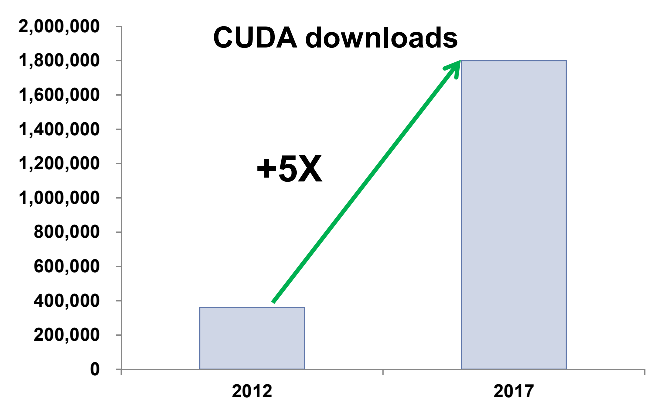
Source: IDC

Exhibit 4: The number of GPU developers has grown significantly…



Source: Nvidia

Exhibit 5: …driving demand for software to support AI development



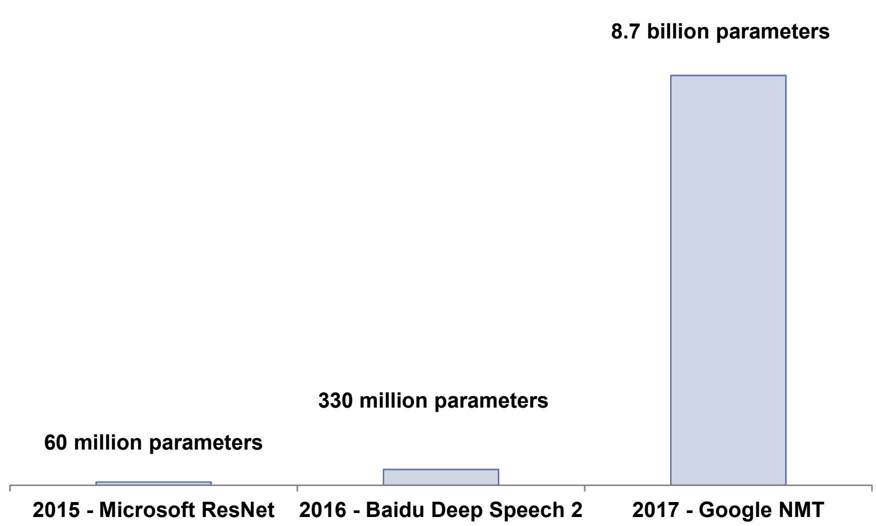
Source: Nvidia

How is AI reshaping the hardware landscape? Easier access to data has enabled an expansion of the AI ecosystem, growth in AI model complexity, and an increase in AI workloads. For instance in Spotify's recent F1 filing, the company stated that its "machine-generated playlists have been made possible by…investments in artificial intelligence and machine learning.... [Spotify] now program approximately 31% of all listening on Spotify across these and other playlists, compared to less than 20% two years ago." That said, such growth has also been enabled by adoption of newer compute architectures that are better suited for processing AI workloads. In the example outlined in the sections above, each layer identifies a different trait of a person, but before that happens the neural net needs to be trained how to do this. In deep learning, the important features are not predefined by humans, but learned and created by an algorithm that leverages large amounts of data using the process of backpropagation described earlier. Backpropagation is extremely compute-intensive, requiring days, weeks, and even months on CPUs to produce usable AI models and, if anything, this process has become even more compute-intensive as models have become more complex (Exhibit 6). However, with the discovery of the benefits of parallel computing architectures this has changed.

In simple terms, parallel computing architectures have a greater number of cores compared to CPUs, and can process tasks in parallel, whereas CPUs process tasks in serial order. Think of it in terms of making a sandwich – if you want to make 10 sandwiches the same way a CPU processes data, you would take a piece of bread, add meat, add cheese, then add the top layer of bread – 1 sandwich down, 9 more to go. However, if you were to do this in a fashion similar to the way GPUs process data, you would start with 10 pieces of bread and apply meat all 10 simultaneously, then add cheese to all 10 simultaneously, and then apply the 10 top pieces of bread all at the same time – 10 sandwiches down, 0 more to go.

Within the context of AI (at least for the training market), the ability to parallelize data is important given the vast amount of data that is being mined for insights. Parallelization enables workloads to be accelerated, reducing the time it takes to develop, refine, and use Deep Learning models. As a result of the improvement in training times, AI application development has accelerated, increasing demand for AI hardware, and creating a large opportunity for companies across the hardware landscape.

Exhibit 6: AI model complexity has been increasing rapidly
Number of parameters in various models

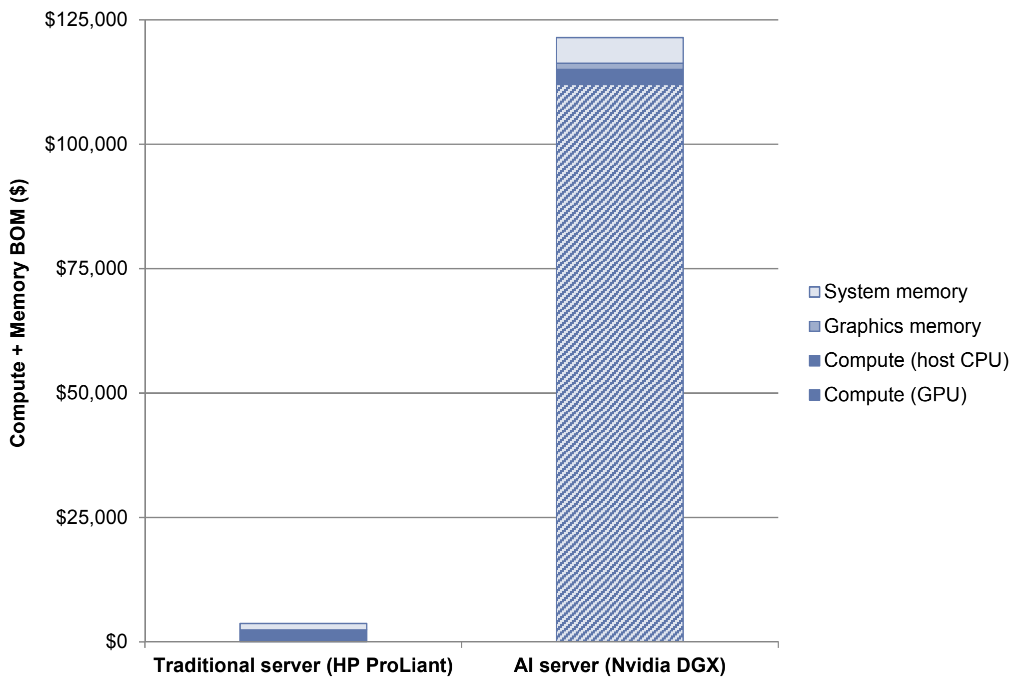


Source: Nvidia.

Server configurations have changed as the focus on AI has grown. Specifically, servers used for AI development have shifted from a configuration traditionally centered on CPUs as the primary compute engine to one featuring multiple GPUs. Take for example Nvidia's DGX AI server, which features 1 Intel host CPU (for $3,000), 8 Nvidia Tesla V100 GPUs (for as much as $14,000 each), 128GB of graphics DRAM (at ~$10/GB), and 512GB of system DRAM (at ~$10/GB). The bill of materials (BOM) for the compute and memory alone for the DGX-1 is $120,000+. For reference, a standard HPE ProLiant 2U rack server features 2 CPU processors (at approximately $1,200 each) and 128GB of DRAM (at ~$10/GB) or $3,700 for compute + memory. What should be clear by this point is that the demands of AI development are driving content higher across the hardware landscape (Exhibit 7) as developers move from traditional architectures to ones that leverage new technologies to accelerate workflow.

Exhibit 7: AI servers have significantly more compute and memory content relative to traditional servers

Server bill of materials comparison



Source: Company data, Goldman Sachs Global Investment Research.

Who is using what? Cloud companies are largely adopting similar hardware configurations to address the growing demand for AI applications (Exhibit 8). We found two commons threads when assessing the hardware stack of companies pursuing AI development: 1) all of them leverage GPUs and 2) they also leverage either FPGAs or ASICs.

That said, we believe the actual mix of hardware is disproportionately skewed towards GPUs given the revenue scale of GPU business relative to those of FPGA/ASIC companies. Note Nvidia's datacenter GPU segment was running at a $2.4bn run rate as of the end of CY4Q17 while the implied datacenter business for Xilinx, an FPGA manufacturer with ~60% market share, was at a run rate of approximately $30mn.

In some cases companies have chosen to develop their own hardware, as is the case at Google, who developed Tensor Processing Units (TPUs). However, we believe that in the medium term this strategy is more likely to be the case at companies with narrowly defined application workloads (i.e. search) that can benefit from the performance boost offered by application-specific chips, rather than at companies with diverse workloads where the high cost of chip development for each application is likely to outweigh the associated benefits.

Bottom line, almost all cloud companies are deploying heterogeneous infrastructure architectures to address a diverse set of application demands, which is why we believe the market for training and inference is likely to be fragmented going forward. However, we also believe that companies that have already established a foothold in the AI hardware environment, whether it be by having the best hardware, the best ecosystem, or both, are more likely to benefit from first-mover advantages than those that are attempting to acqui-hire their way into the market.

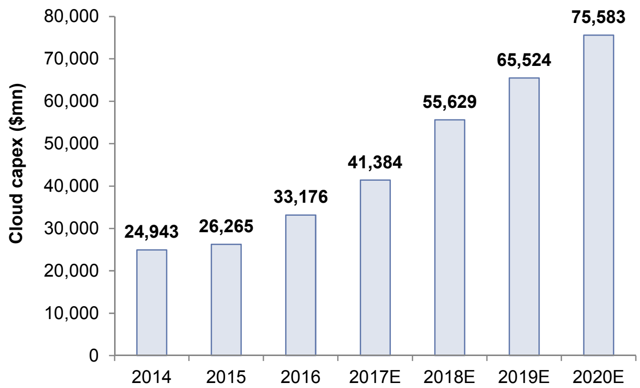
Exhibit 8: Cloud companies leverage multiple hardware architectures in their datacenters
Hardware used by various cloud companies based on public statements

	Hardware platform	Hardware providers	Announced AI/ML-related partnership	Comments
Google	GPUs, ASICs (TPU)	Nvidia, AMD, Intel	Nvidia	Google offers Nvidia-based services as well as services based on its own TPU
Amazon	GPUs, FPGAs	Nvidia, AMD, Xilinx, Intel	Nvidia, Xilinx	Amazon offers Nvidia GPU and Xilinx FPGA instances on AWS.
Microsoft	GPUs, FPGAs	Nvidia, AMD, Xilinx, Intel	Nvidia, AMD (host processor)	Microsoft offers Nvidia GPU services on Azure and has in the past discussed using FPGAs for hyperscale acceleration fabric.
Facebook	GPUs, ASICs	Nvidia, AMD, Intel	Nvidia, Intel	Facebook leverages Nvidia GPUs for its AI development servers (Big Basin) and has indicated that it is working with Intel to develop AI hardware.
Alibaba	GPUs, FPGAs	Nvidia, AMD, Xilinx, Intel	Nvidia, Xilinx	Announced accelerator partnerships with both Nvidia and Xilinx
Baidu	GPUs, FPGAs	Nvidia, AMD, Xilinx, Intel	Nvidia, Xilinx	Announced partnerships with Nvidia, Xilinx, and AMD
Tencent	GPUs, FPGAs	Nvidia, AMD, Xilinx, Intel	Nvidia, Xilinx	Announced accelerator partnerships with both Nvidia and Xilinx

Source: Company data, Goldman Sachs Global Investment Research.

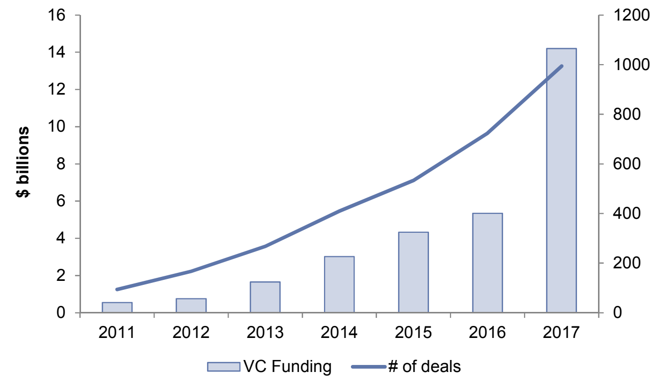
Why is so much investment going into AI? Investment in AI has been accelerating as evidenced by rising cloud capex levels (Exhibit 9) and growth in VC AI funding (Exhibit 10). So why are companies spending so much and why is the funding picture so robust? The simple answer is: because it pays to do so. Specifically, as the AI ecosystem continues to grow and research intensifies, demand for hardware that is capable of processing AI workloads at faster and faster rates is growing, enabling cloud companies to charge a premium for access to such hardware. Note pricing for an Nvidia Pascal (launched in 2016) GPU instance on Google's Compute Engine is ~3x that of an instance based on the Nvidia's older-generation Kepler GPUs and ~170x that of the base CPU offering (Exhibit 11). Even more dramatic is the difference in pricing for Google's TPU service, which is 4x that of the Nvidia Pascal instance and 677x that of the base CPU instance. We believe the combination of 1) constrained compute capacity and 2) growing demand for AI-acceleration is likely to support higher levels of cloud company investment as well as overall growth of the AI hardware TAM.

Exhibit 9: We expect cloud companies to spend $76bn on capex in 2020
Constituents include: Amazon, Facebook, Google, Microsoft, Alibaba, Tencent, Oracle, Salesforce, and others



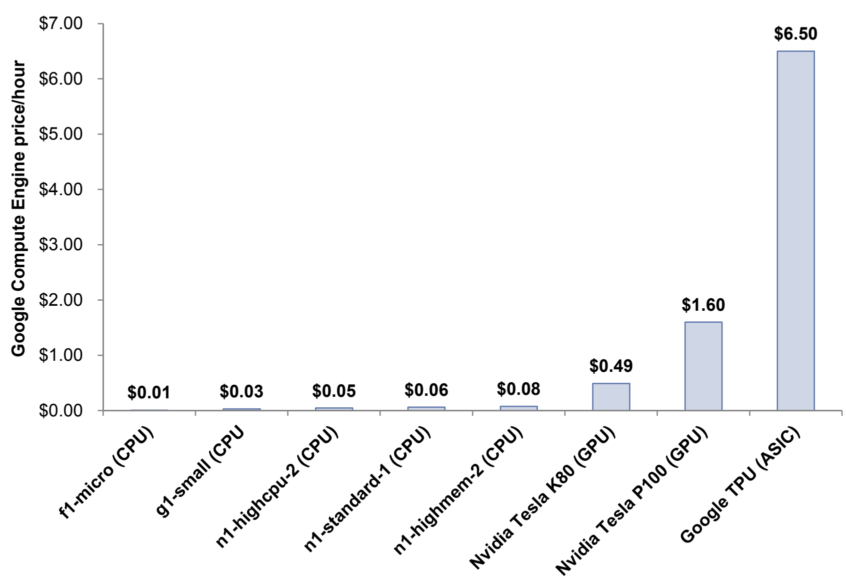
Source: Company data, Goldman Sachs Global Investment Research.

Exhibit 10: Global VC funding in Artificial Intelligence has accelerated
VC Funding (LHS) in $bn, # of deals (RHS)



Source: CB Insights, compiled by Goldman Sachs Global Investment Research

Exhibit 11: Offerings for AI command significantly higher prices
Google Compute Engine price/hour/single compute instance (i.e. per 1CPU, GPU, TPU, etc)



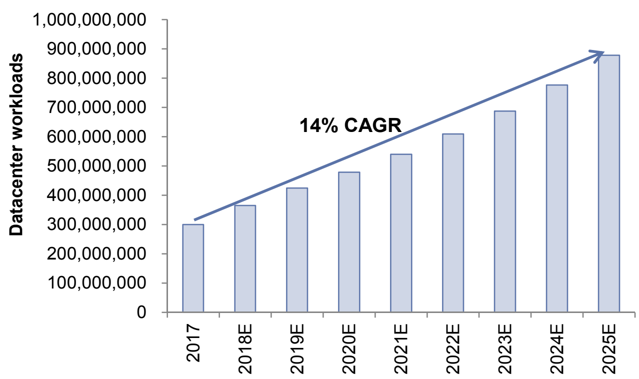
Source: Google, Goldman Sachs Global Investment Research.

How big is the hardware opportunity?

To arrive at our AI Hardware TAM estimates, we considered the following:

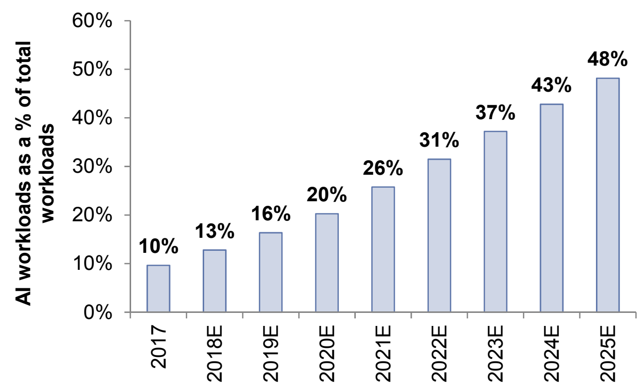
1. **Growth of AI workloads**. Note for the purposes of this report we use Ciscos' definition of "workload", which is a set of resources that are assigned to run a specific application or service.
2. **The ratio of AI servers/AI workload** (i.e. the number of servers required per AI workload).
3. **Server hardware component attach rates** (i.e. the number of GPUs/FPGAs/etc per server).

Exhibit 12: We expect total datacenter workloads to grow at a 14% CAGR through 2025
Datacenter workloads



Source: Cisco, Goldman Sachs Global Investment Research.

Exhibit 13: We expect the mix of AI workloads to grow from 10% in 2017 to 48% by 2025
AI workloads as a % of total workloads



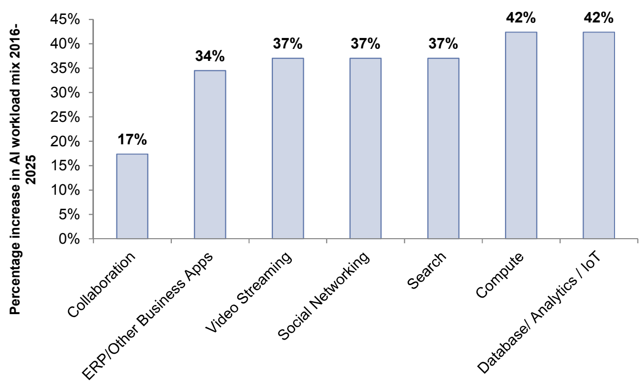
Source: Cisco, Goldman Sachs Global Investment Research.

Leveraging Cisco data we generated a long-term datacenter workload forecast. We assume a terminal workload growth rate that is consistent with the rate of workload growth exiting 2020, resulting in a datacenter workload forecast that grows at a 14% CAGR from 2017 to 2025 (Exhibit 12).

We then assume various AI penetration rates based on workload type. We expect AI to penetrate most workloads, with Database analytics (i.e. Big Data) and Compute (i.e. Cloud IaaS) seeing the greatest increase in AI workload penetration given our expectation for these segments of the datacenter to leverage AI applications the most in an effort to monetize their data (Exhibit 14). Based on workload mix and AI penetration, we see AI workloads growing at a 40% CAGR from 2017 to 2025 or from 10% of total workloads in 2017 to 48% of total workloads by 2025.

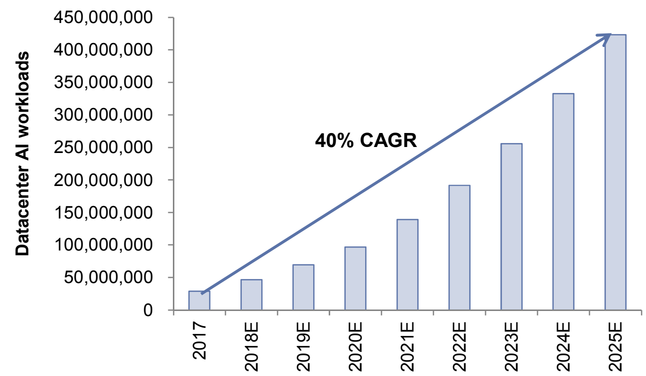
In order to sanity check our assumptions we compare our findings to the growth of both PCs and smartphones (from mainstream introduction to peak/current levels), using each as a proxy for historical growth rates of software platforms that eventually became ubiquitous (i.e. Windows and Android/iOS). Note that the 40% CAGR we forecast for AI workload growth is largely in line with the midpoint of the CAGRs for PCs and smartphones.

Exhibit 14: We expect Database analytics and Compute to see the greatest AI penetration
Net change in AI workload mix 2017 - 2025



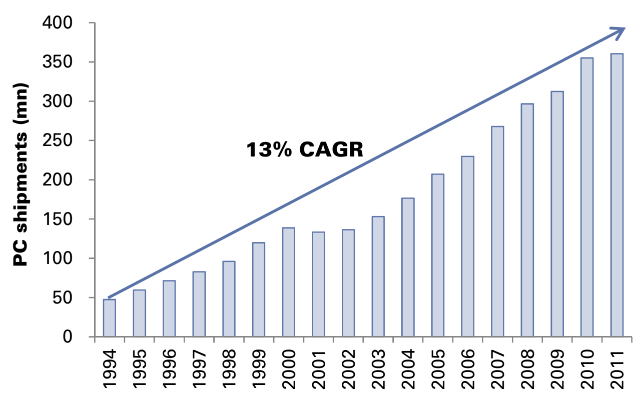
Source: Cisco, Goldman Sachs Global Investment Research

Exhibit 15: We expect AI workloads to grow at a 40`% CAGR through 2025
AI workload growth



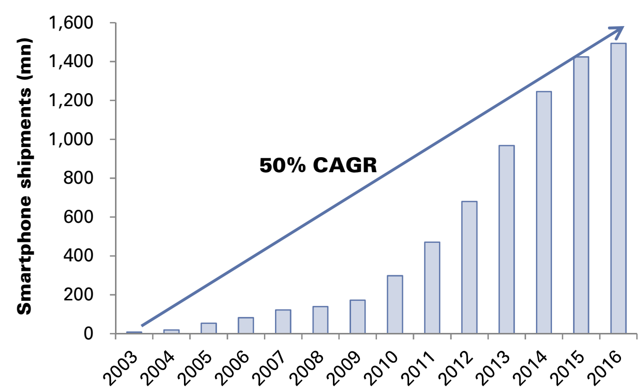
Source: Cisco, Goldman Sachs Global Investment Research

Exhibit 16: PCs grew at a 13% CAGR from 1994 to their peak in 2011
PC shipments 1994 - 2011



Source: IDC.

Exhibit 17: Smartphones have grown at a 50% CAGR since 2003
Smartphone shipments 2003 - 2016



Source: Gartner.

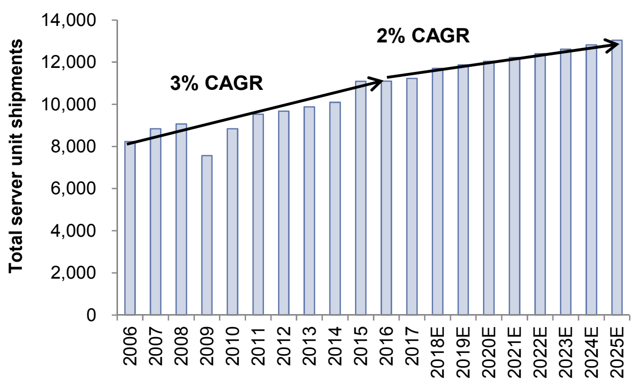
We leverage data provided by Intel and Cisco in order to translate total workloads into an approximate server forecast. Specifically, at its 2017 Investor Day, Intel noted that in 2016 7% of all servers were used for AI applications, which would imply ~777k AI servers. Similarly, Cisco data indicates that there were a total of 235mn workloads in 2016. If we assume that the mix of AI workloads is similar to that of servers, we come to a total implied AI server/AI workload rate of 0.05 servers/workload.

We use this ratio as the starting point for our forecast, taking the AI workload mix discussed above and backing into an implied AI server number. Note that our go-forward estimates account for efficiency improvements (i.e. over time servers should become more efficient and fewer should be required per workload).

Given an implied AI server unit number and assuming that the mix of AI servers is equivalent to the mix of AI workloads, we can back into a total server number (i.e. if there are 1mn servers and AI workloads are 10% of total workloads then a) AI servers account for 10% of total servers and b) there are 10mn total servers).

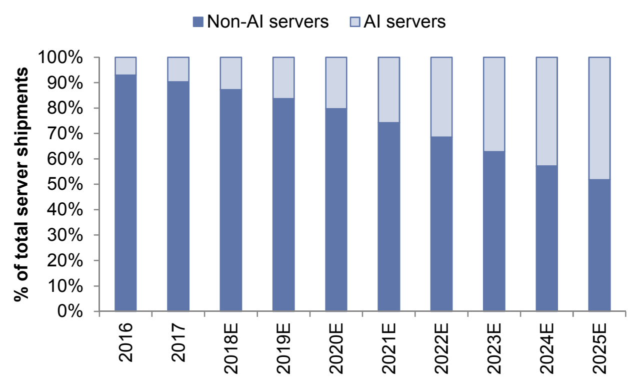
Overall, we expect total server units to grow from 11mn units in 2016 to 13mn units by 2025, which implies a 2% CAGR, or slightly below the rate witnessed from 2006 to 2016. Importantly, though, we would note that the long term growth rate is completely driven by AI servers, which we expect to grow at a 25% CAGR, rather than traditional servers, which we expect to largely decline over the next 8 years as workloads incorporate AI and hardware migrates. Within AI workloads and servers, we assume a gradual shift from training to inference as the "heavy lifting" of AI model development transitions to AI application deployment (i.e. inferencing).

Exhibit 18: We expect the overall server market to grow at a 2% CAGR 2017-2025, slightly below the 3% historical rate
Total server unit shipments



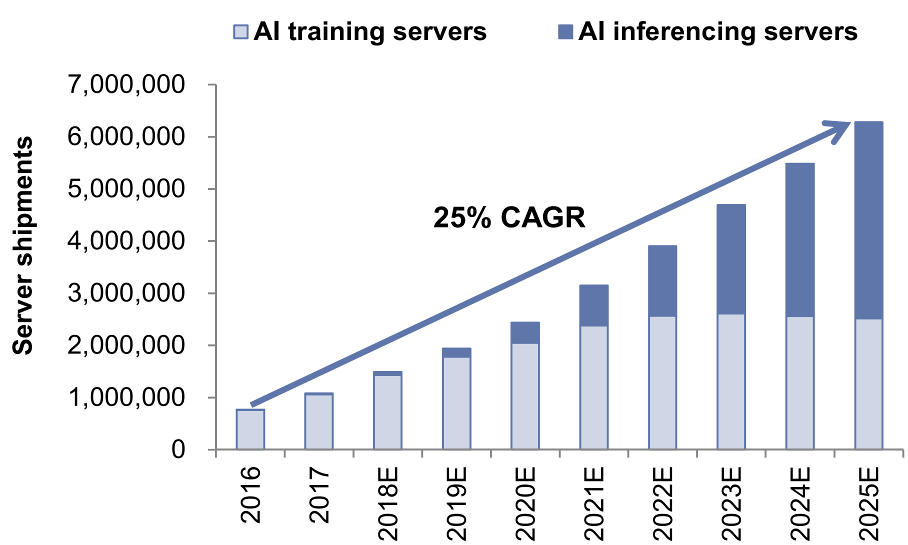
Source: Goldman Sachs Global Investment Research.

Exhibit 19: ....however, we expect the percentage of AI servers to increase over time
Mix of AI and Non-AI servers



Source: Goldman Sachs Global Investment Research.

Exhibit 20: We expect AI servers to grow at a 25% CAGR



Source: Goldman Sachs Global Investment Research.

AI datacenter Compute hardware expected to be $100bn+ opportunity

We view the growth of AI applications as a key enabler of the next leg of growth in the hardware space and expect those companies that are well positioned with premium hardware solutions to outgrow the broader market.

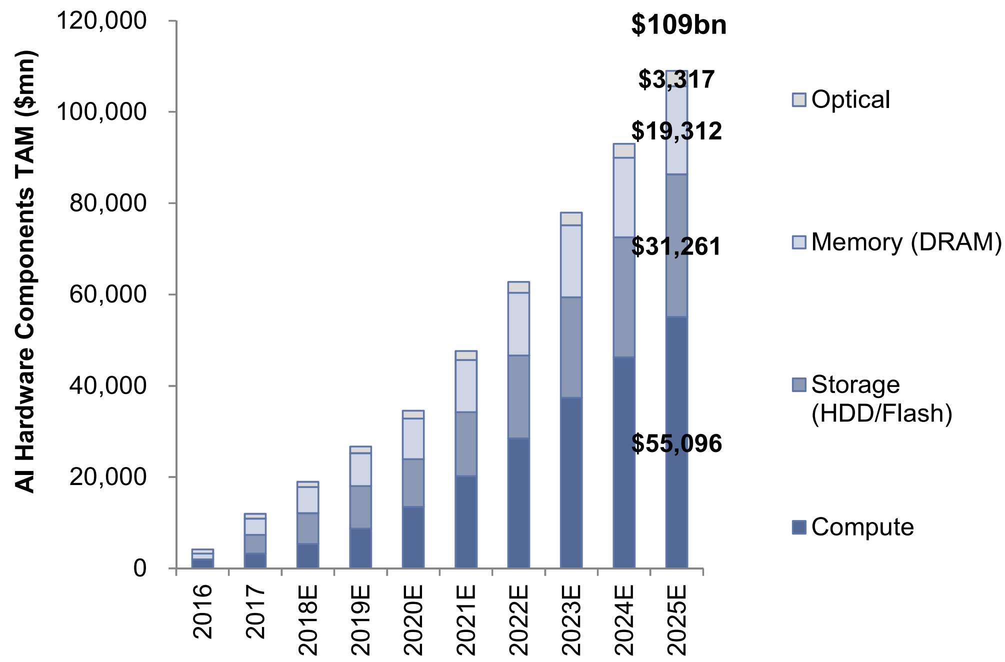
We forecast a total Datacenter AI hardware TAM of $109bn by 2025, which compares to a total AI hardware TAM of $12bn today. We see ASICs, FPGAs, and GPUs as the fastest growing segments of the hardware TAM (see Exhibit 22) and within the datacenter expect compute to account for the greatest amount of overall spend.

Below we go into detail around the opportunities within Compute, Memory, Storage, Foundry, and Optical. We also provide bull and bear cases for each hardware forecast based on 5 different scenarios that could present either upside or downside to our forecast:

- **Workloads grow faster than our baseline**: We assume datacenter workloads grow at a 14% CAGR through our forecast period, but would note that this largely assumes a flat year-over-year growth rate trajectory from 2021-2025. Therefore, there could be upside to our current forecast should AI prove to be additive to overall workloads, rather than replacing workloads, which is our current assumption.
- **Server density increases at a slower rate than expected**: Our forecast assumes an approximately 10%/year improvement in server efficiency (i.e. the number of servers required per workload declines). There could potentially be upside to our forecast if the rate of efficiency improvement were too aggressive.
- **AI accounts for a great % of overall workloads**: We assume AI accounts for 20% of total workloads by 2020 and 48% by 2025; however, there would be upside to our forecast if AI increased as a % of total workloads at a faster rate.
- **Workloads grow slower than our baseline**: We assume datacenter workloads grow at a 14% CAGR through our forecast period; however, there could be downside to our current forecast should workloads grow slower.
- **AI penetration doesn't change**: We assume AI accounts for 20% of total workloads by 2020 and 48% by 2025; however, there would be downside to our forecast if AI were to account for the same percentage through our forecast as it did in 2017 (i.e. 10%).

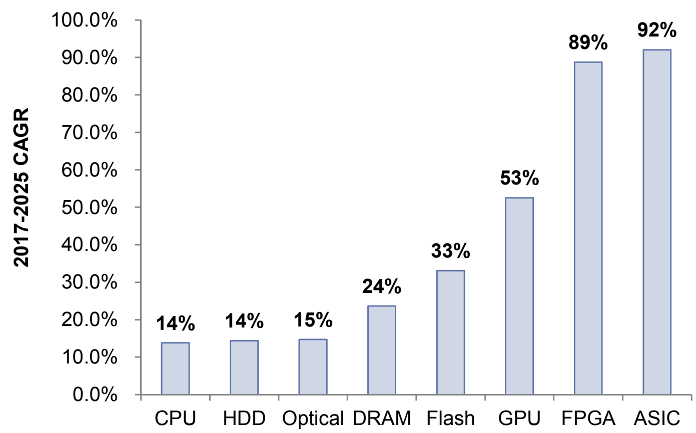
Bottom line, we find that there could be as much as 197% upside and 78% downside to our baseline forecast.

Exhibit 21: We forecast a total AI data center hardware TAM of $109bn by 2025
AI data center component revenue



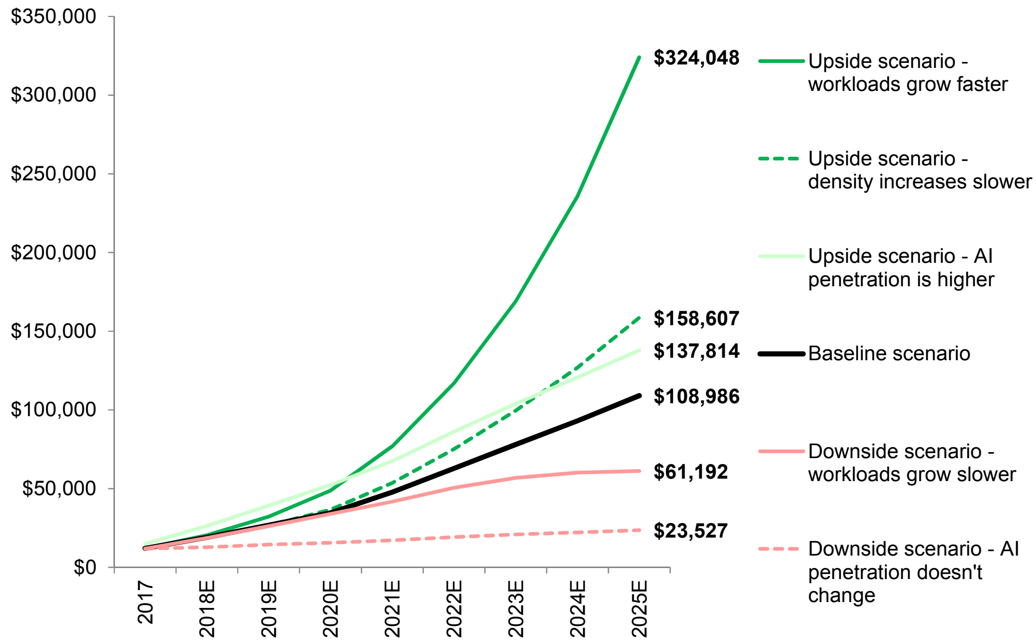
Source: Goldman Sachs Global Investment Research.

Exhibit 22: We expect ASICs, FPGAs, and GPUs to see the fastest growth
2017-2025E CAGR



Source: Goldman Sachs Global Investment Research.

Exhibit 23: We see as much as 197% upside and 78% downside to our baseline forecast
AI hardware TAM scenario analysis ($mn)

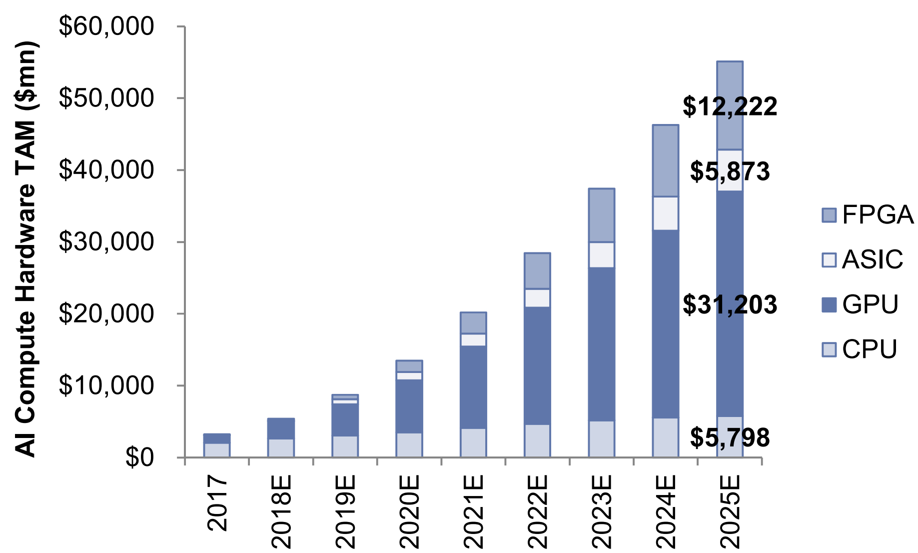


Source: Goldman Sachs Global Investment Research.

COMPUTE: $55BN TAM

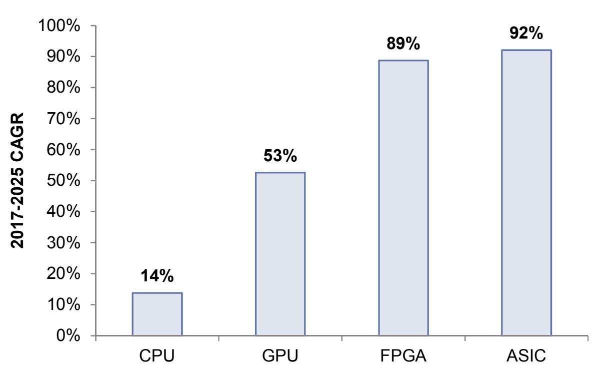
Our analysis indicates a potential AI Compute hardware TAM of $55bn by 2025. We define AI Compute hardware as hardware that is responsible for the processing of AI workloads within an AI system. We see the AI Compute market as being the most fragmented segment within the broader AI hardware landscape given the diversity of workloads that need to be processed. As a result, over the long term we forecast declining revenue share for the incumbent CPU architecture as new hardware solutions (i.e. GPUs and ASICs) are adopted due to their ability to provide relative performance enhancements over existing solutions (Exhibits 26 and 27). That said, we would note that we expect the overall market to grow and for the bulk of companies exposed to this segment to benefit from a rising AI hardware tide.

Exhibit 24: We forecast a total AI Compute hardware TAM of $55bn
AI Compute hardware TAM



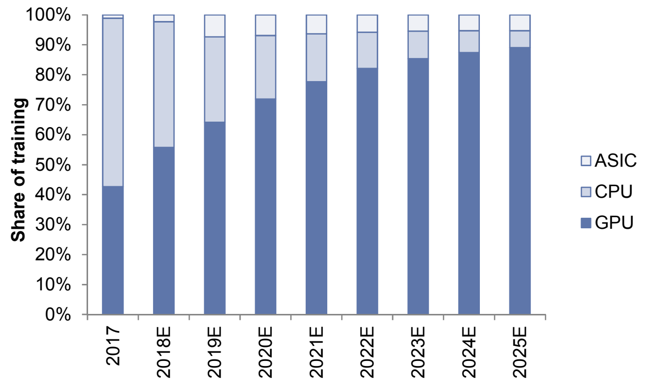
Source: Goldman Sachs Global Investment Research.

Exhibit 25: We expect FPGAs and ASICs to grow the fastest
2017 – 2025E revenue CAGR



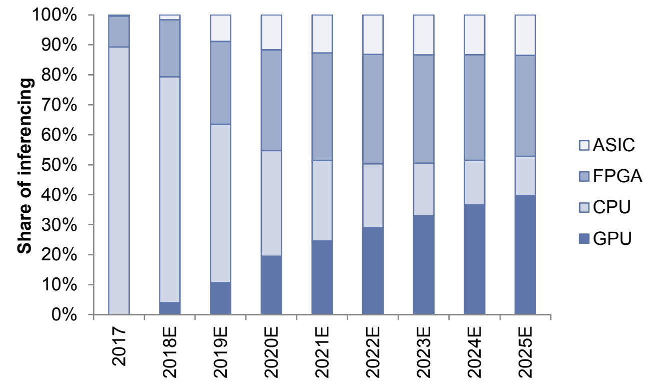
Source: Goldman Sachs Global Investment Research.

Exhibit 26: We expect CPUs to lose share in training…
Training revenue share



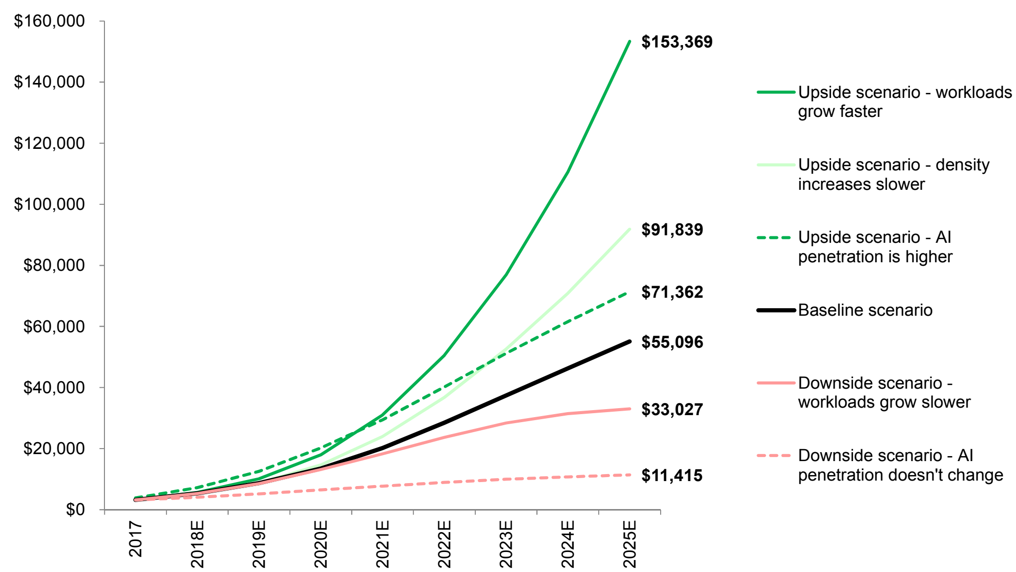
Source: Goldman Sachs Global Investment Research.

Exhibit 27: …as well as inferencing
Inferencing revenue share



Source: Goldman Sachs Global Investment Research.

Exhibit 28: We see as much as 178% upside and 79% downside to our baseline Compute forecast
Compute AI TAM scenario analysis ($mn)



Source: Goldman Sachs Global Investment Research.

Below we summarize our forecast by hardware type and discuss key drivers of growth going forward.

CPUs (growing from $2bn to $6bn by 2025)

We expect CPU unit share of training servers (i.e. primary compute, not as "host" processors) to decline from approximately 93% in 2017 to approximately 35% by 2025 and for inferencing server share to decline from 89% to 35% as we forecast share gains by competing solutions including GPUs, FPGAs, and ASICs. Recall that in recent years GPUs have been rapidly adopted for Deep Learning training applications given their parallel architecture, which is better suited for such applications. Our CPU share forecast reflects a continuation of this trend on the training side as well as the emergence of application-specific solutions for inferencing applications.

On a revenue basis, we expect AI CPU revenue to grow at a 14% CAGR through 2025, the lowest across the various architectures.

Note that our commentary above refers to both CPUs used specifically for AI workloads as well as CPUs that "host" AI workloads (i.e. that serve as control processors for servers, as opposed to being the hardware that workloads run on).

Key players in this area include Intel (who has ~99% server CPU share) and AMD.

GPUs (growing from $1bn to $31bn by 2025)

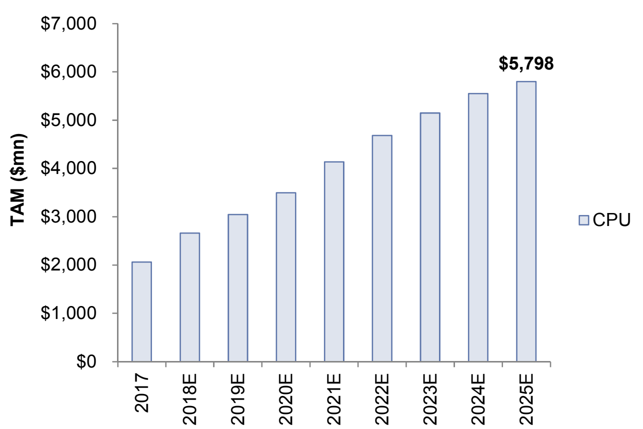
We expect GPU unit share of training servers to grow from approximately 6% in 2017 to approximately 48% by 2025 as we expect the trend of utilizing GPUs for the "heavy lifting" training workloads to continue.

On a revenue basis, we expect total AI GPU revenue to grow at a 53% CAGR through 2025, with training growing at 41% and inferencing growing at 126% (2018-25).

Note that our revenue forecast assumes a constant GPU/server attach rate of 4, which is consistent with our company checks. That said, there could potentially be upside to this attach rate given the fact that Nvidia's newest datacenter GPUs are capable of being assembled in server configurations featuring up to 8 GPUs/server.

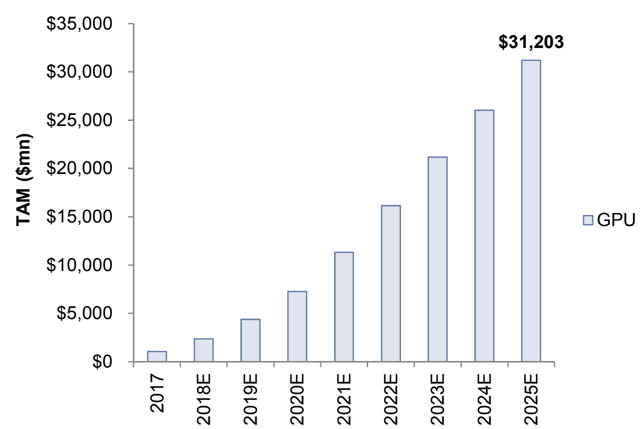
Key players in this area include Nvidia (who has ~70% discrete GPU share) and AMD, to a lesser extent.

Exhibit 29: We expect AI CPU revenue to grow to $6bn by 2025…
AI CPU TAM



Source: Goldman Sachs Global Investment Research.

Exhibit 30: …and for AI GPU revenue to grow to $31bn by 2025
AI GPU TAM



Source: Goldman Sachs Global Investment Research.

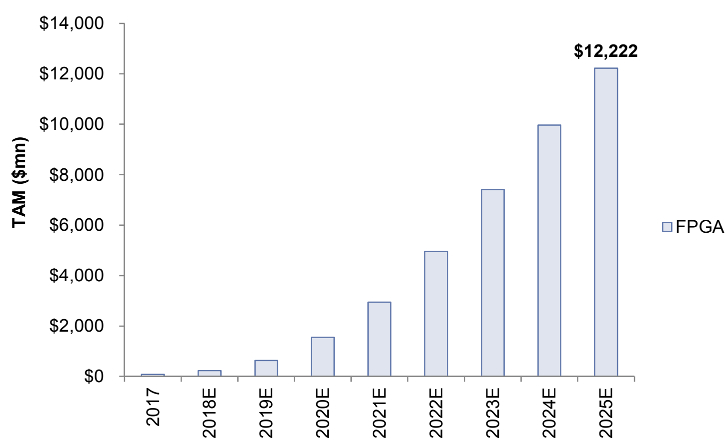
FPGAs (growing from less than $80mn in 2017 to $12bn by 2025)

We expect FPGAs to be utilized primarily in the inferencing market, where we believe share will grow from 10% of inferencing servers in 2017 to 18% by 2025. One barrier to broader adoption, in our view, is the high cost of FPGAs relative to alternative solutions such as ASICs and inference GPUs. Note Amazon utilizes Xilinx's Virtex VU9P line for its FPGA instances on AWS. According to prices posted on Avnet, the median price for the various SKUs of Xilinx's Virtex VU9P is $35,000, which compares to our forecast of $1,300-$1,500 for inference CPUs, $9,000-$13,000 for inference ASICs, and $3,000-$4,000 for inference GPUs.

Although our forecast assumes little relative share gain for FPGAs, we expect overall FPGA AI-related revenue to grow at a 95% CAGR through 2025 driven by increasing deployment of AI models in inferencing applications.

Key players in this area include Xilinx and Intel (via Altera, which was acquired by Intel in 2015)

Exhibit 31: We expect the FPGA TAM to be $12bn by 2025
AI FPGA TAM



Source: Goldman Sachs Global Investment Research.

ASICs (growing from almost nothing in 2017 to $6bn by 2025)

Over the past year investor focus on specialized silicon, or Application Specific Integrated Circuits (ASICs), has increased likely driven by a) deployment of internally developed chip solutions for AI workloads (i.e. Google's TPUs) and b) the emergence of a number of private companies seeking to enter the AI hardware market. The key value proposition of specialized silicon is the prospect of improved performance relative to solutions that are currently available (i.e. GPUs). This performance improvement is enabled by using chip architectures that are designed for specific applications (hence the name, Application Specific Integrated Circuit) and/or new architectures that are better suited for AI/ML workloads. As AI workloads are more broadly deployed (and thus have larger revenue bases to support hardware investment), we expect ASICs to gain share in both training and inferencing, driving a long-term revenue CAGR of 99%, the highest among the various Compute hardware segments.

That said, we would offer the caveat that to-date, adoption of ASICs has been limited. Google utilizes its internally developed TPUs and Intel announced that it was exploring custom chips solutions with Facebook; however, commercial availability of products from private companies like Graphcore and Cerebras is still forthcoming and large volume ramps aren't expected until 2019 at the earliest. Below we highlight a few of the key players in the ASIC space.

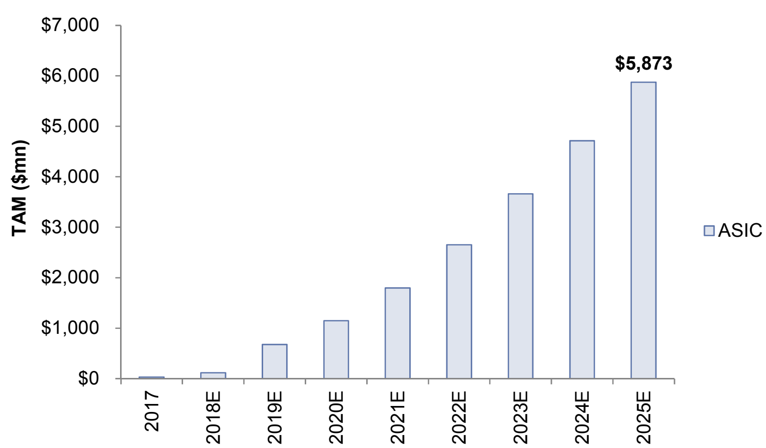
Tensor Processing Units (TPUs): Google's TPUs are internally developed ASIC, which the company uses for inference and training workloads. The company has noted that it will offer its TPUs as a service through Cloud; however, to-date TPUs are not available as commercial hardware offerings.

- Key providers: Google (commercially available, but only as a service on Google Cloud)

Specialized/custom silicon: There are a number of private and recently acquired companies that are working on specialized silicon solutions which feature chip architectures that are radically different from the parallel architectures featured in GPUs today. While details on the architectures have been sparse, partly owing to the fact that commercial products from most companies aren't yet available, company comments suggest that these new devices leverage on-core memory. In very broad terms what this means is that instead of having many cores that need to read/write in and out of memory (which incurs a performance penalty) the cores on these chips have their own memory, which reduces the time it takes to read/write in/out of memory and thus increases performance.

- Key players include:
  - Graphcore
  - Cerebras
  - Intel (Nervana, Movidius, Mobileye, Saffron)
  - Broadcom (custom ASIC IP)

Exhibit 32: We expect the ASIC TAM to grow to $6bn by 2025
AI ASIC TAM



Source: Goldman Sachs Global Investment Research.

Intro to Graphcore and Cerebras, two specialized silicon companies

Graphcore

Graphcore is private company focusing on specialized silicon products for machine learning applications. Graphcore specializes in silicon products for accelerating machine learning (ML) applications. The company was part of XMOS for two years before being established as a separate entity in 2016. Nigel Toon is Graphcore's CEO and has been at Graphcore since its inception and was previously a co-founder of Icera, a cellular modem company that was sold to Nvidia in 2011, and served as a VP and GM at Altera, where he spent 13 years. Graphcore's aim is to lower the cost of AI by improving the performance and efficiency of machine learning workloads with its Intelligent Processing Unit (IPU) accelerators.

Why is specialized silicon important? Graphcore's argument for why its solutions should be adopted is that its IPUs are better suited for ML workloads compared to other solutions like GPUs and FPGAs given the "share nothing" architecture of IPUs and potential cost savings. On the company's "share nothing architecture" the company believes that the IPU architecture is designed such that workloads do not need to be written to/fetched from shared external memory like a GPU and therefore do not face the same performance penalties and memory bandwidth limitations. Instead the company's IPUs feature many small processors with large amounts of individual memory. Graphcore believes that its architecture could enable its IPUs to achieve performance that is 1 to 2 orders of magnitude greater than current GPU solutions. Importantly, the company thinks that its products will enable customers to not only train ML models faster (vs current times of around 1-3 months), but also generate electricity and other savings as a result of shorter training times.

Product introduction and pipeline. Graphcore expects to have products sampling in early 2018, followed by mass deployments in 2H2018 with initial focus coming from cloud and large internet customers. In late 2016, the company was targeting 2H17 for customer availability (see our note *Intro to Graphcore: A specialized silicon company*, published 12/14/16). As noted in that report, the company was expecting its products to be manufactured on TSMC's 16nm FinFET process.

Other key points. Graphcore believes that it needs to at least achieve a 10X performance advantage over current solutions in order for its product to displace incumbent solutions (i.e. GPUs). Mr. Toon has in the past expressed the belief that data centers of the future will come to resemble large scale supercomputers, which will in turn drive demand for high performance networking and compute as well. Mr. Toon has also noted that custom-developed silicon may make more sense for companies with vertical applications like search and that companies with more heterogeneous workloads will likely go with best of breed solutions.

Cerebras

Headquartered in Los Altos, California, Cerebras was founded in 2016 with the purpose of developing an ASIC designed to address deep learning workloads. Andrew Feldman is Cerebras' CEO and co-founder, and the company is Mr. Feldman's 5th startup — the most recent being SeaMicro, which was sold to AMD for ~$350mn in 2012. Cerebras is scheduled to a) sample its first-generation product in early 2018, b) have product on sale in 6-9 months, and c) be in volume production in 2H18. From a funding perspective, Cerebras has raised north of $100mn cumulatively, with backing from several VCs.

Mr. Feldman has stated to us in the past that the company's first-generation product, scheduled to sample in early 2018 and be manufactured at TSMC (16nm), will be orders of magnitude faster than Nvidia's Volta GPU. From an architectural perspective, Cerebras' solution will have on-chip memory, which in Mr. Feldman's view will drive a large part of the performance gain vis-a-vis the GPU (which works with off-chip DRAM).

Single stock ideas for Compute

We expect the compute space to be a key battleground that will determine AI hardware winners and losers. Within our coverage we see Nvidia (Buy) and Xilinx (Buy, on CL) as outsized beneficiaries of growth in the AI ecosystem and believe risk/reward is skewed to the upside. Below we present a few scenario analyses to highlight the potential upside there could be to our current AI TAM estimates and earnings implications for Nvidia and Xilinx.

Key companies:

Nvidia (NVDA, Buy). Our base case AI forecast assumes that AI reaches 20% of total workloads by 2020, implying ~1.2mn AI servers and that 20% of these servers utilize GPUs for workload acceleration, equating to a total GPU opportunity of $7.3bn, which is 29% higher than our current CY2020 Nvidia datacenter estimate of $6.9bn. Note there is a discrepancy between our base case AI forecast and our current Nvidia Datacenter segment revenue estimates as we apply a discount to the implied AI market size as a result of uncertainty associated with such a rapidly growing market.

Below we present a scenario analysis that we believe illustrates Nvidia's opportunity in the AI training space. Using our baseline forecast as the midpoint of our scenario (i.e. AI accounts for 20% of total workloads and that GPUs take 20% share), we assume a range of outcomes with of AI workload mix and GPU share of AI servers as the key variables. Importantly, we would note that the low end of our AI mix range (i.e. 10%) assumes that over the next 3 years AI as a % of total workloads only increases 1% - a scenario we believe to be unlikely given the growing relevance of AI for businesses across a wide range of industries. Further, the low end of our GPU mix assumptions would imply that GPUs *lose* share from current levels, which we believe to be unlikely given the strong position they currently enjoy in the market.

Bottom line, we find that there could potentially be up to 105% upside to our current AI training GPU forecast (Exhibit 33) and up to 52% upside to our Nvidia CY2020 EPS estimate if we assume the high end of the scenario (Exhibit 34). We would note that at the high end of our scenario, Nvidia would trade at 13x CY20 EPS vs 22x based on our current EPS assumptions.

Exhibit 33: We see up to 105% upside to our GPU TAM forecast...
GPU scenario analysis

2020 GPU TAM (incl CY20 inferencing TAM estimate)

2020 AI mix	GPU share of training								
	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%	27.5%	30.0%
30.0%	5,561	6,727	7,892	9,058	10,223	11,388	12,554	13,719	14,885
27.5%	5,173	6,241	7,309	8,378	9,446	10,514	11,583	12,651	13,719
25.0%	4,784	5,756	6,727	7,698	8,669	9,640	10,611	11,583	12,554
22.5%	4,396	5,270	6,144	7,018	7,892	8,766	9,640	10,514	11,388
20.0%	4,008	4,784	5,561	6,338	7,115	7,892	8,669	9,446	10,223
17.5%	3,619	4,299	4,979	5,659	6,338	7,018	7,698	8,378	9,058
15.0%	3,231	3,813	4,396	4,979	5,561	6,144	6,727	7,309	7,892
12.5%	2,842	3,328	3,813	4,299	4,784	5,270	5,756	6,241	6,727
10.0%	2,454	2,842	3,231	3,619	4,008	4,396	4,784	5,173	5,561

Vs baseline estimate (incl CY20 inferencing TAM estimate)

2020 AI mix	GPU share of training								
	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%	27.5%	30.0%
30.0%	(23%)	(7%)	9%	25%	41%	57%	73%	89%	105%
27.5%	(29%)	(14%)	1%	15%	30%	45%	60%	74%	89%
25.0%	(34%)	(21%)	(7%)	6%	19%	33%	46%	60%	73%
22.5%	(39%)	(27%)	(15%)	(3%)	9%	21%	33%	45%	57%
20.0%	(45%)	(34%)	(23%)	(13%)	(2%)	9%	19%	30%	41%
17.5%	(50%)	(41%)	(31%)	(22%)	(13%)	(3%)	6%	15%	25%
15.0%	(56%)	(47%)	(39%)	(31%)	(23%)	(15%)	(7%)	1%	9%
12.5%	(61%)	(54%)	(47%)	(41%)	(34%)	(27%)	(21%)	(14%)	(7%)
10.0%	(66%)	(61%)	(56%)	(50%)	(45%)	(39%)	(34%)	(29%)	(23%)

Source: Goldman Sachs Global Investment Research.

Exhibit 34: ...and up to 52% upside to our CY20 EPS estimate
NVDA CY20 EPS scenario analysis

Vs current CY2020E NVDA Datacenter revenue

2020 AI mix	GPU share of training								
	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%	27.5%	30.0%
30.0%	(19%)	(2%)	15%	32%	49%	66%	83%	100%	117%
27.5%	(25%)	(9%)	7%	22%	38%	53%	69%	85%	100%
25.0%	(30%)	(16%)	(2%)	12%	27%	41%	55%	69%	83%
22.5%	(36%)	(23%)	(10%)	2%	15%	28%	41%	53%	66%
20.0%	(42%)	(30%)	(19%)	(7%)	4%	15%	27%	38%	49%
17.5%	(47%)	(37%)	(27%)	(17%)	(7%)	2%	12%	22%	32%
15.0%	(53%)	(44%)	(36%)	(27%)	(19%)	(10%)	(2%)	7%	15%
12.5%	(59%)	(51%)	(44%)	(37%)	(30%)	(23%)	(16%)	(9%)	(2%)
10.0%	(64%)	(59%)	(53%)	(47%)	(42%)	(36%)	(30%)	(25%)	(19%)

Implied NVDA CY2020E EPS

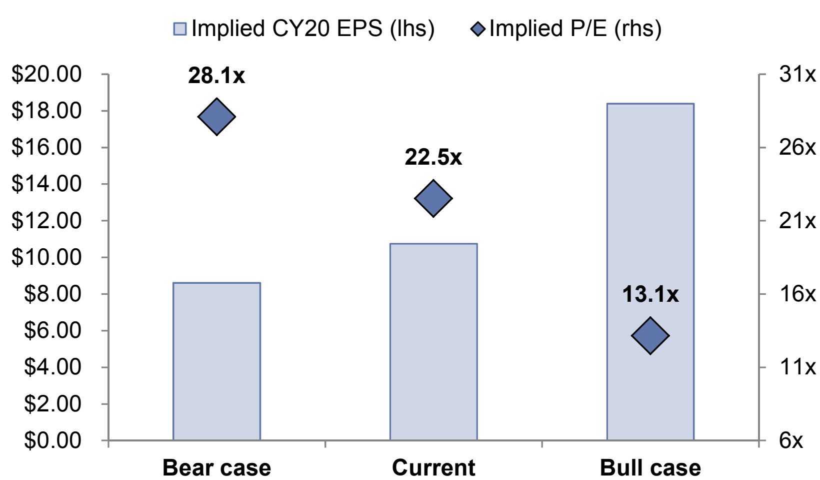
2020 AI mix	GPU share of training								
	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%	27.5%	30.0%
30.0%	$11.06	$11.98	$12.89	$13.81	$14.73	$15.65	$16.56	$17.48	$18.40
27.5%	$10.75	$11.59	$12.43	$13.28	$14.12	$14.96	$15.80	$16.64	$17.48
25.0%	$10.45	$11.21	$11.98	$12.74	$13.51	$14.27	$15.03	$15.80	$16.56
22.5%	$10.14	$10.83	$11.52	$12.21	$12.89	$13.58	$14.27	$14.96	$15.65
20.0%	$9.84	$10.45	$11.06	$11.67	$12.28	$12.89	$13.51	$14.12	$14.73
17.5%	$9.53	$10.07	$10.60	$11.14	$11.67	$12.21	$12.74	$13.28	$13.81
15.0%	$9.22	$9.68	$10.14	$10.60	$11.06	$11.52	$11.98	$12.43	$12.89
12.5%	$8.92	$9.30	$9.68	$10.07	$10.45	$10.83	$11.21	$11.59	$11.98
10.0%	$8.61	$8.92	$9.22	$9.53	$9.84	$10.14	$10.45	$10.75	$11.06

Vs current NVDA CY2020E EPS

2020 AI mix	GPU share of training								
	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%	27.5%	30.0%
30.0%	(8%)	(1%)	7%	14%	22%	30%	37%	45%	52%
27.5%	(11%)	(4%)	3%	10%	17%	24%	31%	38%	45%
25.0%	(13%)	(7%)	(1%)	6%	12%	18%	25%	31%	37%
22.5%	(16%)	(10%)	(5%)	1%	7%	12%	18%	24%	30%
20.0%	(19%)	(13%)	(8%)	(3%)	2%	7%	12%	17%	22%
17.5%	(21%)	(17%)	(12%)	(8%)	(3%)	1%	6%	10%	14%
15.0%	(24%)	(20%)	(16%)	(12%)	(8%)	(5%)	(1%)	3%	7%
12.5%	(26%)	(23%)	(20%)	(17%)	(13%)	(10%)	(7%)	(4%)	(1%)
10.0%	(29%)	(26%)	(24%)	(21%)	(19%)	(16%)	(13%)	(11%)	(8%)

Source: Goldman Sachs Global Investment Research.

Exhibit 35: Based on our scenario analysis Nvidia would trade at 13x in a Bull case
Pricing as of 3/7/18



Source: Goldman Sachs Global Investment Research.

Our 12-month price target of $300 is based on 43x FY19E (CY18E) EPS (incl. ESO) of $6.97. Key risks to our bull thesis include volatility in Ethereum pricing, increased competition particularly in Datacenter, and execution on new products.

Who or what could challenge Nvidia’s hold on the GPU?

Our AI compute hardware forecast implies GPUs have 57% share of the compute TAM by 2025 compared to 22% share for FPGAs and 11% share for ASICs. We would highlight AMD and Intel, to a lesser extent, as competitors in the GPU space that could challenge Nvidia’s position. AMD has graphics expertise in both integrated and discrete graphics chips and Intel has expertise in integrated graphics and has announced its intention to enter the discrete graphics market (for additional details see our note Intel entering discrete graphics market; some key observations, published 11/13/17). That said, we view Nvidia as the key beneficiary of growth in the GPU space given a) Nvidia’s strong competitive positioning as evidenced by its dominant market share in training (close to 100% today) and multiple partnerships with cloud companies, b) significantly larger R&D scale relative to competing GPU providers (note Nvidia’s CY17 R&D spend was $1.8bn and almost entirely focused on GPU development vs $1.2bn for AMD, who is also focused on CPUs, embedded devices, and console systems), and c) software moat (note Nvidia has been investing in its GPU computing platform, CUDA, since 2007).

While our forecast implies leading share for GPUs in the AI compute hardware market, we would note that we are not calling for GPUs to completely overtake this market. Rather, we expect GPU share to be capped by adoption of non-GPU technologies such as FPGAs and ASICs. Note from 2021-2025 we assume GPU TAM share essentially stays in the range of 56-57%. As a result, our forecast predicts that in the outer years (i.e. 2021-2025 and beyond) GPU growth is likely to be more a function of total market growth, rather than ongoing share gains.

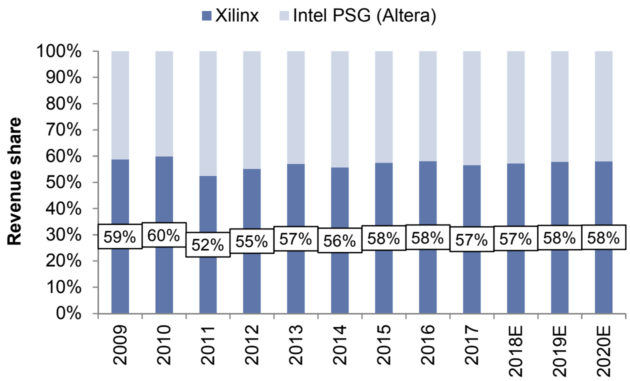
Xilinx (XLNX, Buy, on CL). Our base case AI forecast assumes that inferencing workloads account for 16% of AI workloads and AI servers by 2020E and that 13% inferencing servers utilize FPGAs for workload acceleration (with the balance made up of a mix of CPUs, GPUs, and ASICs), implying a total FPGA opportunity of $1.6bn. Note, our current CY2020 revenue estimate assumes $265mn of hyperscale acceleration contribution. Assuming our Xilinx estimate accounts for 65% share of the inferencing FPGA market equates to a total FPGA inferencing market of roughly $407mn compared to the $1.6bn market we forecast. Note there is a discrepancy between our base case AI forecast and our current Xilinx hyperscale revenue estimates as we apply a discount to the implied AI market size as a result of uncertainty associated with such a rapidly growing market.

Below we present a scenario analysis that we believe illustrates Xilinx’s opportunity in the AI inferencing market, which in our view has been underappreciated by investors based on feedback we’ve received post-our upgrade. Using our baseline forecast as the midpoint of our scenario (i.e. AI accounts for 20% of total workloads, inferencing accounts for 16% of AI servers, and that FPGAs take 13% share of inferencing servers), we assume a range of outcomes with of AI workload mix and FPGA share of inferencing servers as the key variables. We also assume that Xilinx takes 65% share of FPGA inferencing market, which is higher than the company’s current FPGA market share of 58%(Exhibit 36), but in line with Xilinx’s share on the 28nm node and below the company’s share on 16nm (see Exhibit 37).

Bottom line, we find that there could potentially be up to 185% upside to our current AI inferencing FPGA forecast (Exhibit 38) and up to 72% upside to our XLNX CY2020 EPS estimate (Exhibit 39). We would note that at the high end of our scenario, Xilinx would trade at 10x CY20 EPS vs 18x based on our current EPS assumptions.

Exhibit 36: We expect Xilinx to gain share going forward

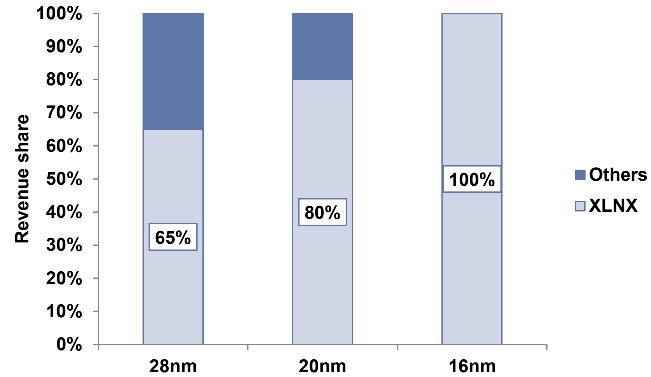
Revenue share



Source: Company data.

Exhibit 37: Xilinx has leading share on newer nodes

Xilinx node revenue share (based on 2017 Analyst Day data)



Source: Company data.

Exhibit 38: We see up to 185% upside to our FPGA TAM forecast…

Inferencing TAM scenario analysis

2020 inferencing FPGA TAM

	2020 FPGA share of inferencing								
2020 AI mix	5.0%	7.5%	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%
30.0%	883	1,324	1,766	2,207	2,648	3,090	3,531	3,973	4,414
27.5%	809	1,214	1,618	2,023	2,428	2,832	3,237	3,641	4,046
25.0%	736	1,103	1,471	1,839	2,207	2,575	2,943	3,310	3,678
22.5%	662	993	1,324	1,655	1,986	2,317	2,648	2,979	3,310
20.0%	589	883	1,177	1,471	1,766	2,060	2,354	2,648	2,943
17.5%	515	772	1,030	1,287	1,545	1,802	2,060	2,317	2,575
15.0%	441	662	883	1,103	1,324	1,545	1,766	1,986	2,207
12.5%	368	552	736	920	1,103	1,287	1,471	1,655	1,839
10.0%	294	441	589	736	883	1,030	1,177	1,324	1,471

vs baseline 2020 inferencing FPGA TAM

	2020 FPGA share of inferencing								
2020 AI mix	5.0%	7.5%	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%
30.0%	(43%)	(15%)	14%	42%	71%	99%	128%	156%	185%
27.5%	(48%)	(22%)	4%	30%	57%	83%	109%	135%	161%
25.0%	(53%)	(29%)	(5%)	19%	42%	66%	90%	113%	137%
22.5%	(57%)	(36%)	(15%)	7%	28%	49%	71%	92%	113%
20.0%	(62%)	(43%)	(24%)	(5%)	14%	33%	52%	71%	90%
17.5%	(67%)	(50%)	(34%)	(17%)	(0%)	16%	33%	49%	66%
15.0%	(72%)	(57%)	(43%)	(29%)	(15%)	(0%)	14%	28%	42%
12.5%	(76%)	(64%)	(53%)	(41%)	(29%)	(17%)	(5%)	7%	19%
10.0%	(81%)	(72%)	(62%)	(53%)	(43%)	(34%)	(24%)	(15%)	(5%)

Source: Goldman Sachs Global Investment Research.

Exhibit 39: …and up to 72% upside to our XLNX CY20 EPS estimate

XLNX CY20 EPS scenario analysis

vs current XLNX CY2020 datacenter revenue (assuming 65% share for XLNX)

	2020 FPGA share of inferencing								
2020 AI mix	5.0%	7.5%	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%
30.0%	43%	115%	187%	259%	330%	402%	474%	546%	617%
27.5%	31%	97%	163%	229%	294%	360%	426%	492%	557%
25.0%	20%	79%	139%	199%	259%	318%	378%	438%	498%
22.5%	8%	61%	115%	169%	223%	277%	330%	384%	438%
20.0%	(4%)	43%	91%	139%	187%	235%	283%	330%	378%
17.5%	(16%)	26%	67%	109%	151%	193%	235%	277%	318%
15.0%	(28%)	8%	43%	79%	115%	151%	187%	223%	259%
12.5%	(40%)	(10%)	20%	49%	79%	109%	139%	169%	199%
10.0%	(52%)	(28%)	(4%)	20%	43%	67%	91%	115%	139%

Implied XLNX CY2020 EPS (assuming 65% share for XLNX)

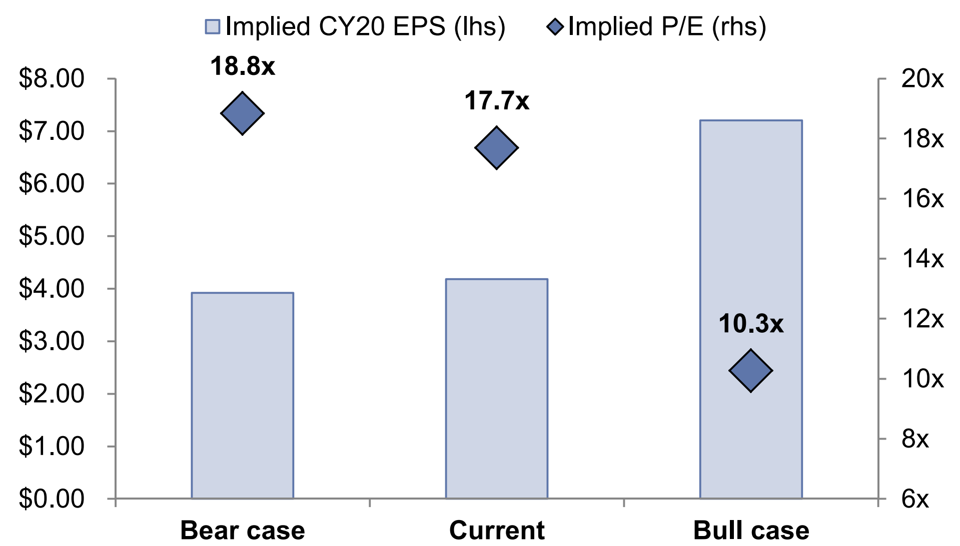
	2020 FPGA share of inferencing								
2020 AI mix	5.0%	7.5%	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%
30.0%	$4.39	$4.74	$5.10	$5.45	$5.80	$6.15	$6.50	$6.85	$7.20
27.5%	$4.33	$4.66	$4.98	$5.30	$5.62	$5.95	$6.27	$6.59	$6.91
25.0%	$4.28	$4.57	$4.86	$5.15	$5.45	$5.74	$6.03	$6.33	$6.62
22.5%	$4.22	$4.48	$4.74	$5.01	$5.27	$5.54	$5.80	$6.06	$6.33
20.0%	$4.16	$4.39	$4.63	$4.86	$5.10	$5.33	$5.56	$5.80	$6.03
17.5%	$4.10	$4.31	$4.51	$4.72	$4.92	$5.13	$5.33	$5.54	$5.74
15.0%	$4.04	$4.22	$4.39	$4.57	$4.74	$4.92	$5.10	$5.27	$5.45
12.5%	$3.98	$4.13	$4.28	$4.42	$4.57	$4.72	$4.86	$5.01	$5.15
10.0%	$3.92	$4.04	$4.16	$4.28	$4.39	$4.51	$4.63	$4.74	$4.86

vs current CY2020E EPS

	2020 FPGA share of inferencing								
2020 AI mix	5.0%	7.5%	10.0%	12.5%	15.0%	17.5%	20.0%	22.5%	25.0%
30.0%	5%	13%	22%	30%	39%	47%	56%	64%	72%
27.5%	4%	11%	19%	27%	35%	42%	50%	58%	65%
25.0%	2%	9%	16%	23%	30%	37%	44%	51%	58%
22.5%	1%	7%	13%	20%	26%	32%	39%	45%	51%
20.0%	(1%)	5%	11%	16%	22%	28%	33%	39%	44%
17.5%	(2%)	3%	8%	13%	18%	23%	28%	32%	37%
15.0%	(3%)	1%	5%	9%	13%	18%	22%	26%	30%
12.5%	(5%)	(1%)	2%	6%	9%	13%	16%	20%	23%
10.0%	(6%)	(3%)	(1%)	2%	5%	8%	11%	13%	16%

Source: Goldman Sachs Global Investment Research.

Exhibit 40: Based on our scenario analysis Xilinx would trade at 10x under the Bull case
Pricing as of 3/7/18



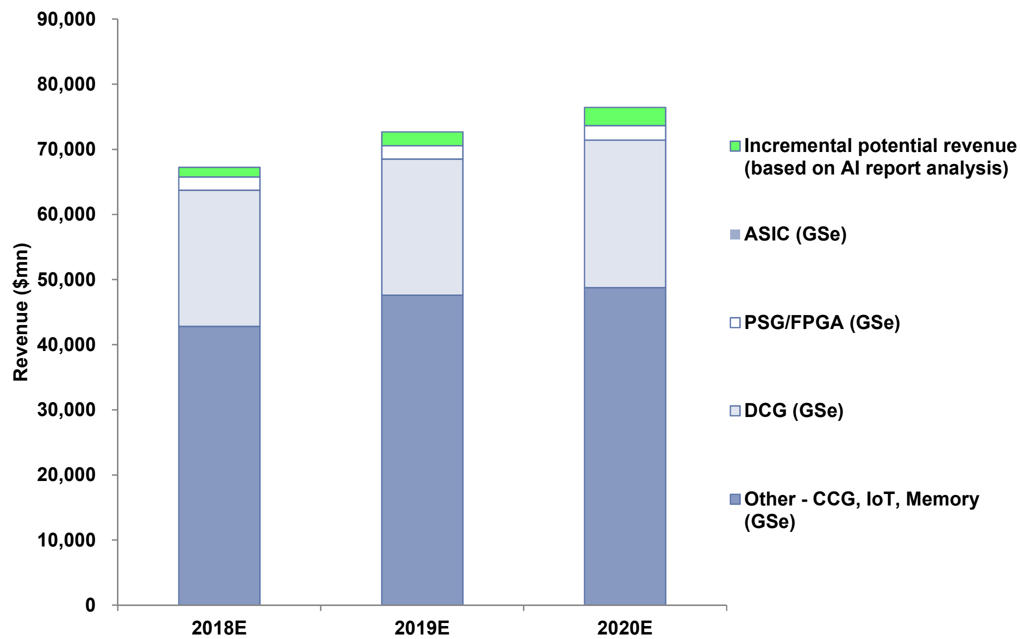
Source: Goldman Sachs Global Investment Research.

Our 12-month price target of $89 remains based 85% on a fundamental component (26x our normalized EPS forecast of $3.40) and 15% on an M&A component (22x normalized EBITDA). Key risks to our constructive thesis include, prolonged weakness in Comms, slow adoption of FPGAs in the Cloud, opex growth beyond what we already model, and intense competition.

Intel (INTC, Neutral). Our AI hardware forecast implies potential incremental revenue (from CPUs, FPGAs, ASICs, and NAND) of $1.5bn-$2.8bn. This would drive our CY18-20 Intel revenue estimates 2.2%-3.8% higher, with a 1.4%-2.9% potential impact to our EPS estimates (Exhibit 41 and 42). While the potential upside to our Intel estimates would be minimal, we think there is potential for downside to the company's valuation. Note Intel trades at 14x NTM EPS, which is 9%/6% above vs its median 2/5yr NTM P/E, driven in part we believe, by investor enthusiasm around new opportunities in the AI hardware market. Given the lack of potential upside implied by our forecast, we believe the stock could potentially de-rate as investors assign a multiple to the datacenter business (i.e. 30% of total revenue) that is more in line with lower growth enterprise companies. For reference, HPE trades at 14x NTM EPS, IBM trades at 11x NTM EPS, and HPQ trades at 12x NTM EPS.

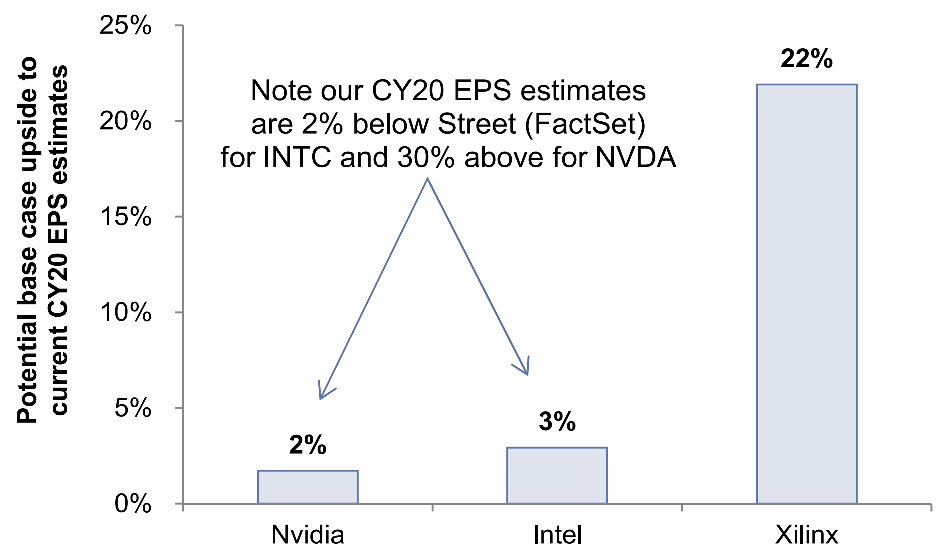
Our 12-month price target of $48 is based on 14x normalized EPS of $3.45. Key risks include, 1) any dramatic change in the PC demand environment, 2) market share shifts vis-à-vis AMD, 3) cloud capex growth, 4) any impact from Spectre/Meltdown, and 5) significant changes to capex.

Exhibit 41: We see minimal incremental contribution to our current Intel estimates
Intel potential upside scenario analysis



Source: Goldman Sachs Global Investment Research.

Exhibit 42: Intel has only modest upside relative to hardware peers
Comparison of potential base case CY20 EPS upside



Source: Goldman Sachs Global Investment Research

How could Intel respond to the threat? Our take.

One of the key conclusions that we draw from this report is that we do not expect Intel to benefit disproportionately relative to other companies that are part of the AI hardware ecosystem. However, we are not suggesting that Intel cannot avoid this outcome. Below we highlight a number of scenarios which we believe could drive upside to the conclusion we come to in this report. The scenarios below are not intended to be exhaustive and we acknowledge possibilities exist outside of what we discuss below.

1) **Opportunity at the Edge.** We would note that the analysis within this report largely pertains to datacenter applications and therefore does not consider Intel's opportunity in areas such as drones, autonomous vehicles, and smart cities. Note Intel has embedded processing capabilities and has also acquired a number of companies that develop edge-computing devices (i.e. Movidius for machine vision systems and Mobileye for autonomous vehicles). Therefore, we believe it is likely that Intel could benefit growth of AI-at-the-edge applications. Below we've outlined a scenario analysis in which we size the potential revenue opportunity related to autonomous vehicles, drones, and smart cities.

Exhibit 43: Intel has multiple opportunities at the edge
Intel edge opportunity scenario analysis

Autonomous vehicles		Comment
Worldwide automotive production CY2020 (mn)	98	GSe
Autonomous vehicle mix	2%	GSe
Intel content/vehicle	1,250	Consistent with server processor ASP
2020 revenue opportunity ($mn)	**2,450**	

Drones		Comment
Cumulative 5yr Drone forecast (mn)	1.8	From GS Profiles in Innovation: Drones
% with AI capability	50%	Midpoint of prior Nvidia analysis
Intel content/drone	110	Consistent with mobile CPU ASP
Cumulative revenue opportunity ($mn)	**100**	

Smart cities		Comment
Installed security cameras worldwide by 2020 (mn)	1,000	According to Nvidia
Assume % with AI functionality requiring high perf. processor	1%	GSe
Camera ASP	100	Consistent with what we believe tobe Movidius ASP
Cumulative revenue opportunity ($mn)	**1,000**	

Source: Company data, Goldman Sachs Global Investment Research.

2) **Buying growth.** One course of action that Intel could pursue in order to bolster its AI capabilities would be to acquire a company with AI capabilities that complement what it already has. We calculate that at 2x net leverage Intel would have $41.5bn of M&A firepower. According to FactSet there are approximately 300-350 semiconductor companies worldwide with a market cap below $41.5 (assuming a 20% premium to current market caps, which is consistent with recent deals over the last few years). That said, we believe there could be difficulty in pursuing this strategy given a number of factors including:

- A number of companies have expertise in technologies that would be unlikely to have the ability to leverage Intel's in-house manufacturing capabilities and IP. Such areas include analog and RF, which make up a large portion of the larger market cap names on the list; this which brings up another issue…
- The revenue scale of companies that do have expertise in areas that could leverage Intel's existing manufacturing capabilities and IP are generally to be too small to meaningfully move the needle for Intel. However, as we demonstrate above, such acquisitions could prove to have technologies that are capable of being deployed into large growing TAMs outside of the datacenter and therefore could

become meaningful over time.

- For companies that have significant overlap with Intel's existing business (particularly in CPUs, GPUs, and FPGAs), any potential deal could face intense anti-trust scrutiny.

3) **Foundry opportunity**. Finally, Intel's foundry business could provide an additional growth vector if the company were able to capture share of the AI hardware foundry market (see later in the report for additional details). Note though that foundry margins are materially lower than what we assume for the incremental opportunity associated with ASICs, FPGAs, and CPUs. Note we assume a 70% incremental gross margin in our Intel upside analysis on business related to these devices vs the "more than 50%" gross margin TSMC, the leading foundry, has guided to.

MEMORY AND STORAGE: A COMBINED $50BN TAM

Memory and Storage

We believe that memory and storage will be a key enabler of AI. We estimated memory and storage demand based off: 1) Underlying assumptions detailed in earlier in this report for both AI servers (to increase from about 10% of servers in 2017 to 20% in 2020 and 48% in 2025) and AI workloads. 2) Our estimates (and along with input from memory/storage company comments, and a report from Cisco) for DRAM capacity requirements per server and storage (HDD/SSD) needs per workload (done per workload rather than per server due to the impact of storage arrays), as well as typical ASP curves. We discuss our discussions more fully later in this section.

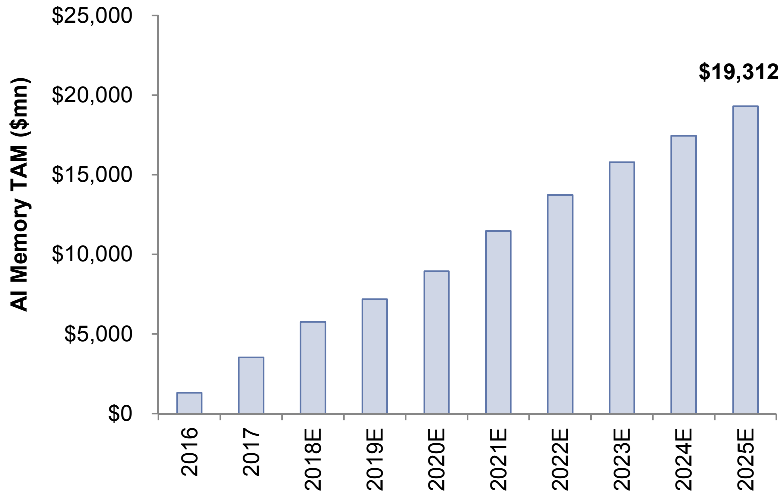
The bottom line is that we forecast that DRAM revenue (server and graphics DRAM) tied to AI will increase from $3.5bn in 2017 (a mid-single-digit percent of the total industry) to $8.9bn in 2020 and $19bn in 2025. We forecast storage revenue (HDDs and SSDs in total) to rise from $4.1bn in 2017 (a mid to high single digit percent of the industry) to $10.5bn in 2020 and $31bn in 2025. We believe **Micron** (DRAM and SSD exposure), **Western Digital** (leader in high capacity HDDs, as well as SSD exposure), **Samsung Electronics** (DRAM and SSD exposure), and **SK Hynix** (DRAM and SSD exposure) are the best positioned in our global semis coverage to benefit from these trends.

For DRAM, we assumed that an inferencing server has 20% more DRAM than an average server, and that a training server has 2.5X more DRAM than an average server as a starting point (or about 390 GB of DRAM in 2017, although there are models such as the DGX-1 that use 512 GB). We also assumed 16 GB of graphics DRAM for each GPU that is attached to an AI server (which are primarily in training). We assumed a low 20% CAGR in average content through 2025 (which is faster than we believe the DRAM industry could grow from shrinks alone). We assumed a mid-teens decline in ASP per Gb through 2025 (or relatively flat margins). We believe our starting assumptions are reasonable, as they imply that AI is currently a mid-single-digit percent of DRAM revenue. Most DRAM companies suggest AI is having an impact on sales but that it is still relatively small overall and difficult to quantify.

As higher memory bandwidth is required for better performance of the parallel computing using GPUs, DRAM makers have come out with a new type of memory called High Bandwidth Memory (HBM). Compared to the GDDR (Graphics Double Data Rate) memory currently mainly used for this purpose as it has a higher memory bandwidth compared to conventional DRAM, HBM has an even higher bandwidth, and also takes up less PCB space due to the stacking of DRAM dies through TSVs (through-silicon vias).

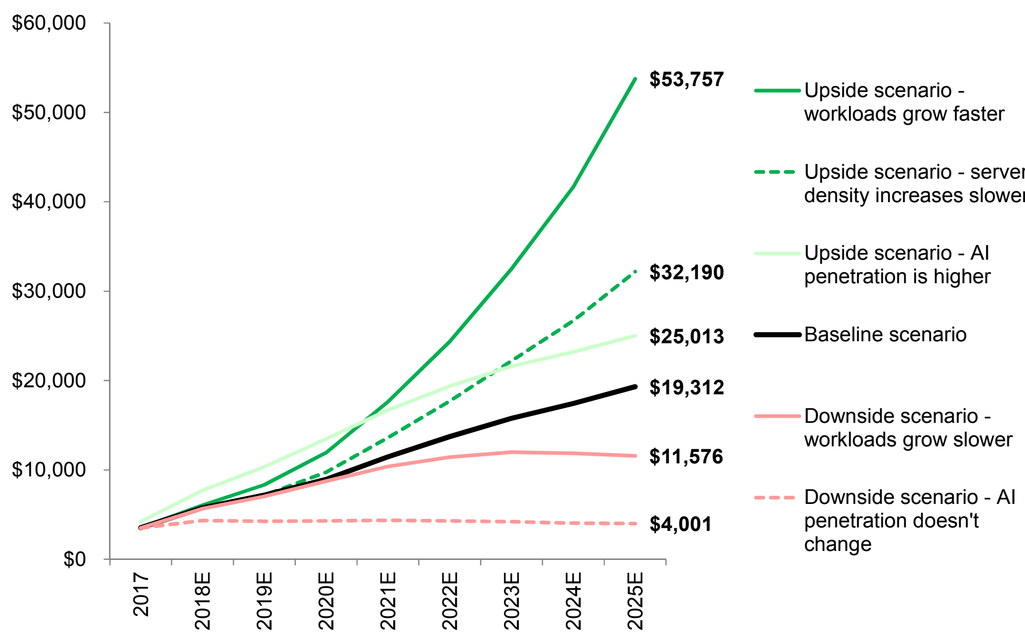
Both Samsung Electronics (SEC) and SK Hynix (Hynix) have made comments on HBM in their recent earnings calls, as the former stated that it started HBM2 supply in 2016 and is seeing a significant yoy growth in demand, while the latter said that it started volume production of HBM2 in 2H17 and is expecting market demand to more than double each year going forward. While we believe that the HBM sales share of total DRAM sales was still less than 1% in 2017, as the pricing is known to be at least 2X more expensive than regular DRAM, there could be a potential tailwind to AI memory ASP and margin if the portion within the DRAM industry becomes more significant going forward.

Exhibit 44: We expect the AI memory market to reach $19bn by 2025
AI Memory TAM



Source: Goldman Sachs Global Investment Research.

Exhibit 45: Our memory forecast could be as large as $53bn or as small as $4bn
Memory AI TAM scenario analysis ($mn)

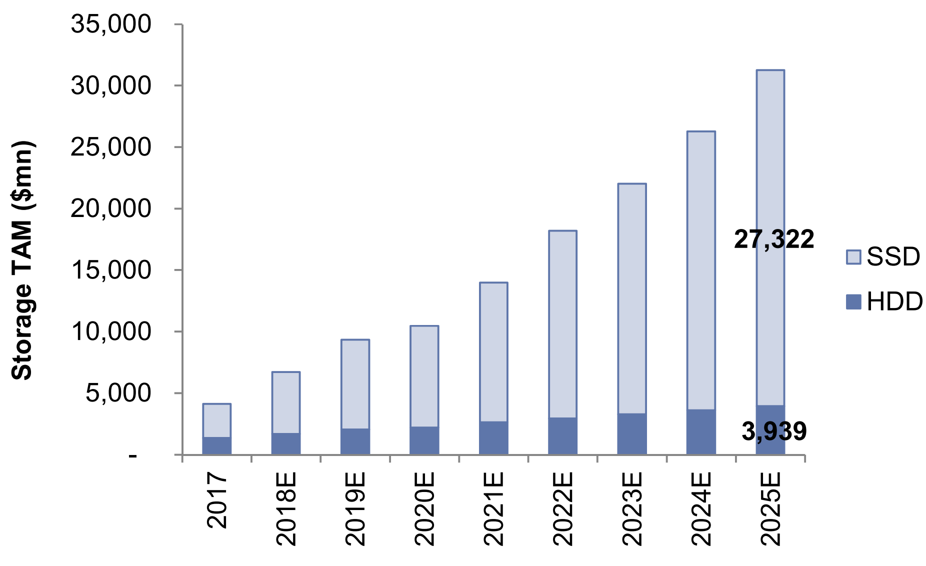


Source: Goldman Sachs Global Investment Research.

For storage, we looked at the installed base of traditional and cloud (public and private) bits (per Cisco) relative to the number of workloads to derive an average storage requirement per workload (or about 2.6 TB). We then made the following assumptions: 1) that an AI workload would use 15% more storage than the average for the industry; 2) that the amount of storage needed per workload increases at a mid-teens CAGR (consistent with the Cisco forecast); 3) that HDD ASPs decline at a mid-teens CAGR, and 4) that SSD ASPs decline at a high 20% range CAGR through 2025. Our $4.1bn storage forecast in 2017 assumes $1.3bn for HDDs and $2.7bn for SSDs. Our 2025 forecast consists of $3.9bn for HDDs and $27bn for SSDs. We assumed that 80% of bits are for capacity storage and 20% are for performance storage (which is more tilted to performance than the overall industry, which is 85%/15% at present), that all of the performance bits were on SSDs, and that by 2025 25% of capacity bits would be on SSDs with 75% on HDDs (as we assumed there is still more of a performance need for AI than the overall industry even in the "capacity" tier).

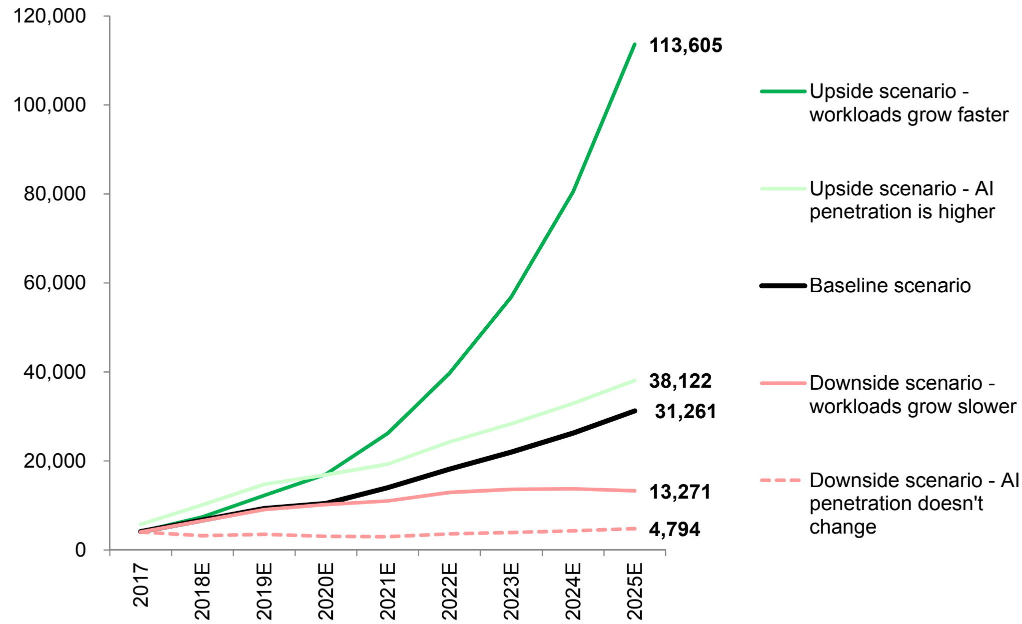
Beyond potential differences vs. our baseline assumptions (i.e. more or less capacity per server, different ASPs, etc.), one potential change to our forecast is with storage class memories (with XPoint being an example). By the later part of our 2025 outlook, these have the potential to become material. However, we believe technical and ecosystem improvements need to be made for these to become mainstream (such as improving write endurance in XPoint, and potentially lower costs). That said, we believe XPoint's focus on helping for applications like in-memory database analytics could position it well for certain types of AI workloads. Our forecast assumes that 15-20% of AI workloads will be for database analytics and IoT, and we believe that some portion of this could be filled by XPoint (or other storage class memories) in the middle to later parts of our forecast period.

Exhibit 46: We expect the AI storage market to reach $31bn by 2025
AI Storage TAM (SSD and HDD)



Source: Goldman Sachs Global Investment Research.

Exhibit 47: We see up to 263% upside and 85% downside to our storage forecast
Storage AI TAM scenario analysis ($mn)



Source: Goldman Sachs Global Investment Research.

Single stock ideas for Memory and Storage

Applying industry market share based on current IDC data (22% of DRAM and 5% of enterprise SSDs for Micron, and 48% of capacity HDDs and 12% of enterprise SSDs for Western Digital) to our baseline assumptions, we derived current and future datacenter AI revenue by company.

For Micron, this implies that $920mn (or 4%) of Micron's revenue is currently tied to datacenter AI, and that it will rise to about $2.4bn (or 10%) by 2020.

For WD, this implies that about $980mn (or 5%) of its revenue is currently tied to datacenter AI, and that it will rise to about $2bn (or 10%) by 2020.

We believe it's challenging to quantify the long-term EPS impact to memory and storage companies from AI. This is due to the fungibility over time of DRAM and NAND capacity (i.e. bits can be used for servers, PCs or smartphones), and in turn the fact that upside in demand for AI would likely have knock-on effects to pricing in other areas (for example mobile DRAM would benefit from increased allocation toward server and graphics DRAM, all else equal). We therefore believe that exploring changes to our ASP assumptions is the best way to evaluate upside demand from AI, especially as ASPs are the most sensitive model driver.

We are currently modeling normalization in DRAM and NAND pricing/margins in 2019/2020 in our Micron and Western Digital models to reflect historical cyclicality in the memory industry, and we believe this can occur even in a baseline AI growth market through 2020. However, to explore the potential impacts if AI is able to mitigate or even prevent a downturn, we examined scenarios in 2020E where margins in both DRAM and NAND are between 100 and 500 bps higher, and where WD's capacity HDD sales are between $50 and $250mn higher, and what the impact would be on 2020 EPS (Exhibits 48-49).

Exhibit 48: We see up to 72% upside to our FY20 Micron EPS estimate
Micron gross margin scenario analysis

	Changes in Micron's FY20 DRAM & NAND GM assumption					
Change in DRAM & NAND GM (bps)	*(500)*	*(250)*	*-*	*250*	*500*	*1,000*
FY20 EPS impact	-26.6%	-13.9%	0.0%	15.3%	32.2%	72.2%

Source: Goldman Sachs Global Investment Research.

Exhibit 49: We see up to 62% upside to our FY20 WD EPS estimate
Western digital gross margin scenario analysis

	Changes in WD's FY20 NAND GM & capacity HDD sales					
Change in NAND GM (bps)	*(500)*	*(250)*	*-*	*250*	*500*	*1,000*
Change in capacity HDD sales ($ mn)	*(100)*	*(50)*	*-*	*100*	*250*	*500*
FY20 EPS impact	-22.8%	-12.0%	0.0%	13.1%	28.1%	61.9%

Source: Goldman Sachs Global Investment Research.

Estimating the EPS impact for Samsung Electronics and for SK Hynix

Applying industry market share based on current IDC data and our market share assumptions (45% of DRAM and 43% of enterprise SSDs for SEC, and 28% of DRAM but assuming 3% of enterprise SSD in 2020 for Hynix) to our baseline assumptions, we derived current and future datacenter AI revenue for the Korean memory companies as well.

For SEC, this implies that $2.8bn (or around 1%) of SEC's revenue is currently tied to datacenter AI, and that it will rise to about $7.6bn (or 3%) by 2020.

For Hynix, this implies that about $990mn (or 4%) of its revenue is currently tied to datacenter AI, and that it will rise to about $2.8bn (or 8%) by 2020.

We conduct a scenario analysis to see what the impact to these companies' 2020 EPS would be when the same year DRAM and NAND operating margins differ from our base case by 1000 bps lower to 1000 bps higher. We use margin as the variable as we believe that higher/lower datacenter AI related memory mix will lead to higher/lower memory ASP and thus higher/lower memory margin, assuming shipment and cost are constant. As seen in the exhibits, the EPS impact for Hynix will be larger than SEC due to being a pure memory company while SEC has a more diversified earnings structure.

Exhibit 50: We see up to 31% upside to our Samsung Electronics 2020 EPS estimate
Samsung Electronics OPM scenario analysis

	Changes in SEC's 2020 DRAM and NAND OPM assumption						
Change in DRAM and NAND OPM (bps)	(1,000)	(500)	(250)	0	250	500	1,000
Impact on 2020E EPS (%)	**-21%**	**-11%**	**-6%**	**0%**	**7%**	**14%**	**31%**

Source: Goldman Sachs Global Investment Research.

Exhibit 51: We see up to 51% upside to our Hynix 2020 EPS estimate
Hynix OPM scenario analysis

	Changes in Hynix 2020 DRAM and NAND OPM assumption						
Change in DRAM and NAND OPM (bps)	(1,000)	(500)	(250)	0	250	500	1,000
Impact on 2020E EPS (%)	**-36%**	**-20%**	**-10%**	**0%**	**11%**	**23%**	**51%**

Source: Goldman Sachs Global Investment Research.

Micron:

We are Buy rated on MU shares. Our 12-month price target of $58 is based on 10X normalized EPS of $5.80.

Key risks relate to DRAM and NAND supply/demand, server/PC/handset unit growth, market share, margins, and the potential entry of China into the memory industry.

Western Digital:

We are Neutral rated on WDC shares. Our 12-month price target of $93 is based on 9X our normalized EPS estimate of $10.30 (including SBC).

Key risks relate to NAND supply/demand, HDD demand, margins (which is tied to HDD mix and NAND S/D), and leverage.

Samsung Electronics

Valuation: We have a Buy rating (on CL) and SOTP-based 12-month target price of W3.35mn for SEC common shares and a 12-month preferred share target price (based on our pref shares' liquidity and yield 2-factor model) of W2.55mn and Buy rating.

Risks: A major deterioration in memory supply/demand and/or a sharp contraction in smartphone margins.

SK Hynix

Valuation: We have a Buy rating and historical P/B vs. ROE-based 12-month target price of W105,000.

Risks: A sharp deterioration in DRAM/NAND supply/demand and/or an unexpected capacity increase.

FOUNDRY: $16BN TAM

We estimate that AI foundry market is $1.9bn market in 2018E with a growth CAGR of 38% in 2018-2025. TSMC, Samsung LSI, and GlobalFoundries are the main foundry suppliers for GPUs, FPGAs, and ASICs. We believe that these AI semiconductors are compute intensive and require leading edge technologies. In 2018, TSMC will likely become the clear leader of semiconductor manufacturing technology for the first time given our view that TSMC has a 1-2 year competitive advantage over key rivals Samsung LSI and Intel in mass production of 7nm process node and should pass Intel for the first time in transistor density level in 2018 and should at least maintain its dominant market share of 85% in 2018-2020, in our view. To quantify the foundry market from our AI semiconductor forecast, we assumed that AI semiconductors have 59% gross profit margin and 70% of COGS are wafer cost (i.e. foundry revenue) as is typical. We currently estimate that AI semiconductor to represent 1.7%, 3.3%, and 5.4% of total revenues at TSMC in 2018E-2020E, respectively.

Methodology:

We breakdown TSMC's AI TAM by GPU, CPU, FPGA and ASIC (AI only). For each category, we have following assumptions:

1) AI GPU: we assume foundry's revenue as % of AI GPU TAM is 27% and TSMC has 85% - 90% market share in 2018E-2020E;

2) AI CPU: we assume foundry's revenue as % of AI CPU TAM is 45% and TSMC has 1%-5% market share in 2018E-2020E;

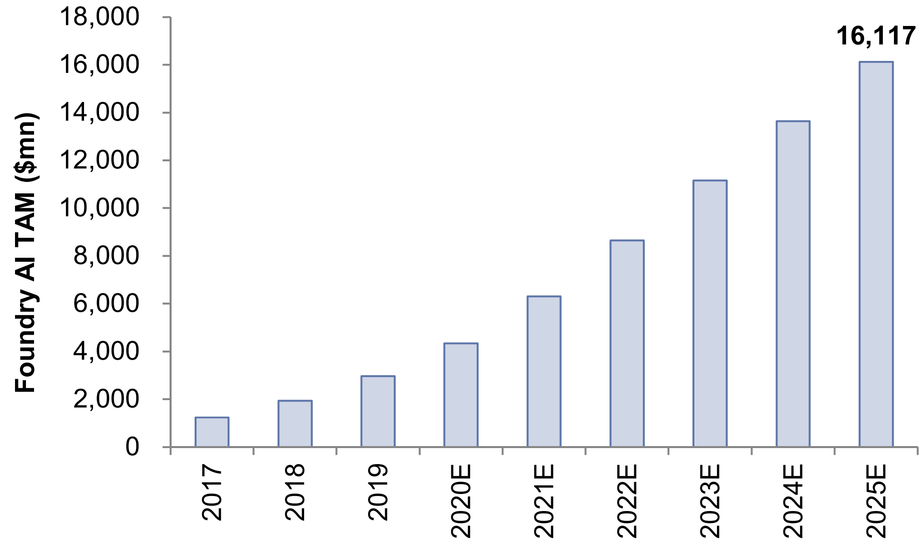
3) AI FPGA: we assume foundry's revenue as % of AI FPGA TAM is 21% and TSMC has 100% market share in Xilinx's FPGA while it has no share in Intel's through 2018E-2020E;

4) AI ASIC: we assume foundry's revenue as % of AI ASIC TAM is 42% and TSMC has 80% market share in 2018E-2020E.

Note that foundry's revenue as % of AI TAM is based on the assumption that 70% of the COGS is wafer cost.

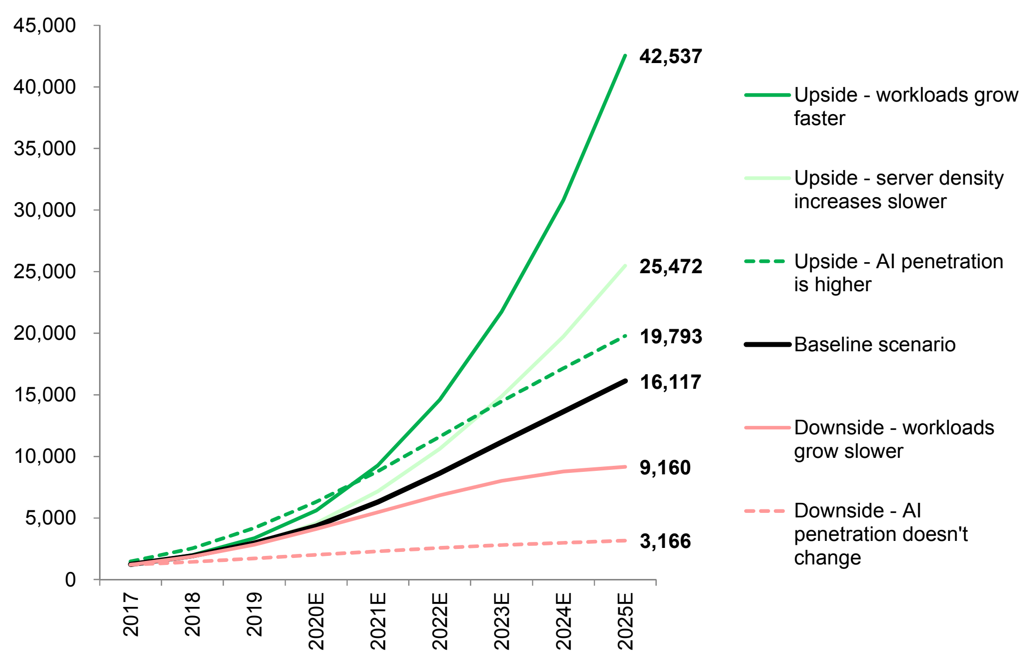
Therefore, the sum of the product of TSMC's market share and TAM of each application is our base case TSMC's AI revenue in 2018E-2020E. Then, we apply 2.2X leverage (TSMC's average operating leverage times financial leverage for the past three years) on the AI revenue, as it translates into the EPS contribution from AI.

Exhibit 52: We expect the AI foundry market to reach $16bn by 2025
AI Foundry TAM



Source: Goldman Sachs Global Investment Research.

Exhibit 53: We see up to 164% upside and 80% downside to our baseline foundry forecast
Foundry AI TAM scenario analysis ($mn)



Source: Goldman Sachs Global Investment Research.

Competitive landscape:

We believe TSMC has 1-2 year competitive advantage over key rivals Samsung LSI and Intel in mass production of 7nm process node and should pass Intel for the first time in transistor density level in 2018. We think AMD and Nvidia will design-in their first wave of 7nm products at TSMC: Intel makes mostly CPU and is an IDM, but TSMC makes dozens of leading-edge products for many customers. Thus, it is conceivable to us that Intel can optimize its process for its CPU, but not its FPGA and ASIC. Besides, AMD has been using GlobalFoundries as its main foundry supplier, but it announced on Aug 31st 2016 a 5 year amendment to its Wafer Supply Agreement with GlobalFoundries which provides AMD with flexibility in sourcing foundry services at 14nm and 7nm from additional foundries. Thus, we believe TSMC will also gain market share in AMD's 7nm business.

For more details, please refer to our report "TSMC (2330.TW): *The Asia Stock Collection: Taiwan's Tech Giant finds a spring in its step; up to Buy"* dated Jan 22nd 2018.

Scenario analysis:

Based on the above assumptions, our baseline scenario suggests that AI semiconductor to represent 1.7%, 3.3%, and 5.4% of total revenues at TSMC in 2018E-2020E, respectively. Our scenario analyses below show 1) the mix of AI revenue under a number of conditions and 2) potential EPS impacts to our base case assumption. Note that we reflect TSMC revenue estimates in US$ terms with the NT$/US$ exchange rate as of 3/7/18.

Exhibit 54: We think AI hardware could account for as much as 5% of TSMC's 2020 revenue
AI revenue contribution analysis

TSMC (US$mn)	2016	2017	2018E	2019E	2020E
TSMC total revenue (GHe)	29,389	32,124	37,221	42,786	45,340
AI Compute revenue	96	267	630	1,425	2,453
from GPU	*96*	*246*	*548*	*1,071*	*1,777*
from CPU	-	-	*12*	*41*	*78*
from FPGA	-	*10*	*32*	*87*	*212*
from ASIC	-	*11*	*39*	*227*	*385*
AI Compute as a % of TSMC revenue	0.3%	0.8%	1.7%	3.3%	5.4%

*NT dollar / US dollar exchange rate as of 3/7/18

Source: Goldman Sachs Global Investment Research.

Working on the basis our other assumptions are unchanged, we also conducted another 3 upside scenarios and 2 downside scenarios analysis as follows:

1) Upside scenario – workloads grow faster. AI semiconductor will represent 1.8%/3.8%/7.1% in 2018E-2020E and it has 0.2%-3.9% earnings upside to our baseline scenario.

Exhibit 55: TSMC's EPS could be 4% higher than GSe if workloads grow faster
TSMC scenario analysis

Upside - workloads grow faster ($mn)	2016	2017	2018E	2019E	2020E
TSMC implied total revenue (GHe)	29,293	32,124	37,252	43,014	46,149
AI Compute revenue	-	267	661	1,653	3,261
from GPU	-	*246*	*575*	*1,244*	*2,370*
from CPU	-	-	*12*	*45*	*94*
from FPGA	-	*10*	*33*	*101*	*283*
from ASIC	-	*11*	*41*	*264*	*513*
New AI Compute as a % of TSMC revenue	0.0%	0.8%	1.8%	3.8%	7.1%
Revenue change vs. GHe		**+0.0 pp**	**+0.1 pp**	**+0.5 pp**	**+1.8 pp**
EPS change vs. GHe		**+0.0 pp**	**+0.2 pp**	**+1.2 pp**	**+3.9 pp**

*NT dollar / US dollar exchange rate as of 3/7/18

Source: Goldman Sachs Global Investment Research.

2) Upside scenario – AI penetration is higher. AI semiconductor will represent 2.3%/4.7%/7.9% in 2018E-2020E and it has 1.3%-5.9% earnings upside to our baseline scenario.

Exhibit 56: TSMC's EPS could be 6% higher than GSe if AI penetration is higher
TSMC scenario analysis

Upside - AI penetration is higher (US$mn)	2016	2017	2018E	2019E	2020E
TSMC implied total revenue (GHe)	29,293	32,176	37,435	43,411	46,566
AI Compute revenue	-	318	844	2,051	3,679
from GPU	-	*293*	*733*	*1,543*	*2,674*
from CPU	-	-	*16*	*56*	*107*
from FPGA	-	*12*	*42*	*125*	*319*
from ASIC	-	*13*	*53*	*327*	*579*
New AI Compute as a % of TSMC revenue	0.0%	1.0%	2.3%	4.7%	7.9%
Revenue change vs. GHe		**+0.2 pp**	**+0.6 pp**	**+1.5 pp**	**+2.7 pp**
EPS change vs. GHe		**+0.4 pp**	**+1.3 pp**	**+3.2 pp**	**+5.9 pp**

*NT dollar / US dollar exchange rate as of 3/7/18

Source: Goldman Sachs Global Investment Research.

3) Upside scenario – density increases slower. AI semiconductor will represent 1.7%/3.3%/5.7% in 2018E-2020E and it has 0%-1.0% earnings upside to our baseline scenario.

Exhibit 57: TSMC's EPS could be 1% higher than GSe if server density increases at a slower rate
TSMC scenario analysis

Upside - server density increases slower (US$mn)	2016	2017	2018E	2019E	2020E
TSMC implied total revenue (GHe)	29,293	32,124	37,221	42,784	45,550
AI Compute revenue	-	267	630	1,423	2,663
from GPU	-	*246*	*548*	*1,071*	*1,935*
from CPU	-	-	*12*	*39*	*77*
from FPGA	-	*10*	*32*	*87*	*231*
from ASIC	-	*11*	*39*	*227*	*419*
New AI Compute as a % of TSMC revenue	0.0%	0.8%	1.7%	3.3%	5.8%
Revenue change vs. GHe		**+0.0 pp**	**-0.0 pp**	**-0.0 pp**	**+0.5 pp**
EPS change vs. GHe		**+0.0 pp**	**-0.0 pp**	**-0.0 pp**	**+1.0 pp**

*NT dollar / US dollar exchange rate as of 3/7/18

Source: Goldman Sachs Global Investment Research.

4) Downside scenario – AI penetration doesn't change. AI semiconductor will represent 1.3%/2.0%/2.7% in 2018E-2020E and it has 0.9%-6.2% earnings downside to our baseline scenario.

Exhibit 58: TSMC's EPS could be 6% lower than GSe if AI penetration doesn't change
TSMC scenario analysis

Downside - AI penetration doesn't change (US$mn)	2016	2017	2018E	2019E	2020E
TSMC implied total revenue (GHe)	29,293	32,122	37,066	42,204	44,063
AI Compute revenue	-	264	475	844	1,176
from GPU	-	*243*	*413*	*635*	*855*
from CPU	-	-	*9*	*23*	*34*
from FPGA	-	*10*	*24*	*51*	*102*
from ASIC	-	*11*	*30*	*135*	*185*
New AI Compute as a % of TSMC revenue	0.0%	0.8%	1.3%	2.0%	2.7%
Revenue change vs. GHe		**-0.0 pp**	**-0.4 pp**	**-1.4 pp**	**-2.8 pp**
EPS change vs. GHe		**-0.0 pp**	**-0.9 pp**	**-3.0 pp**	**-6.2 pp**

*NT dollar / US dollar exchange rate as of 3/7/18

Source: Goldman Sachs Global Investment Research.

5) Downside scenario – workloads grow slower. AI semiconductor will represent 1.7%/3.3%/5.3% in 2018E-2020E and it has 0.1%-0.3% earnings downside to our baseline scenario.

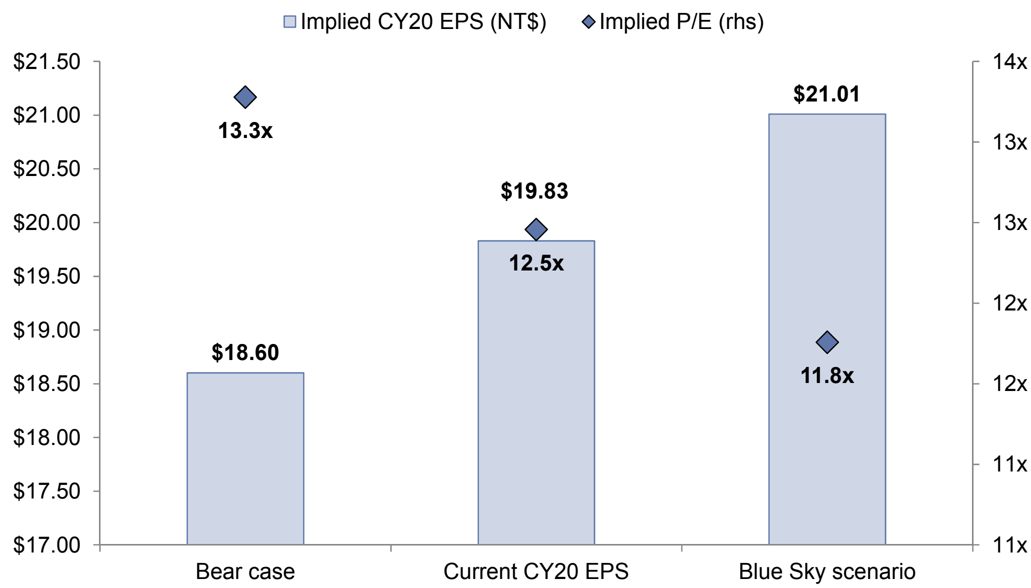
Exhibit 59: TSMC's EPS would be relatively unchanged if workloads grew at a slower rate
TSMC scenario analysis

Downside - workloads grow slower (US$mn)	2016	2017	2018E	2019E	2020E
TSMC implied total revenue (GHe)	29,293	32,122	37,211	42,755	45,278
AI Compute revenue	-	264	620	1,394	2,390
from GPU	-	*243*	*539*	*1,049*	*1,737*
from CPU	-	-	*11*	*38*	*69*
from FPGA	-	*10*	*31*	*85*	*207*
from ASIC	-	*11*	*39*	*222*	*376*
New AI Compute as a % of TSMC revenue	0.0%	0.8%	1.7%	3.3%	5.3%
Revenue change vs. GHe		**-0.0 pp**	**-0.0 pp**	**-0.1 pp**	**-0.1 pp**
EPS change vs. GHe		**-0.0 pp**	**-0.1 pp**	**-0.2 pp**	**-0.3 pp**

*NT dollar / US dollar exchange rate as of 3/7/18

Source: Goldman Sachs Global Investment Research.

Exhibit 60: Based on our scenario analysis TSMC would trade at 12x under the Blue Sky case
Pricing as of 3/7/18



Source: Goldman Sachs Global Investment Research.

TSMC (Buy): Our 12m TP of NT$287 is based on 17X NTM P/E. Key risks: intensified competition from Samsung and Intel at the advanced nodes; weaker-than-expected Bitcoin related wafer demand; slower-than-expected smartphone recovery.

Introduction to Horizon Robotics and DeePhi

Horizon Robotics or Hobot (unlisted) was founded by a team of experienced Artificial Intelligence (AI) scientists and Baidu/Facebook/Google alumni in July 2015. Based in Beijing, the Company is a leading global AI startup offering industry-leading embedded AI technology solutions with strong capabilities in AI algorithms, software, hardware and computing units. Dr. Kai Yu, Hobot's founder & CEO and former head of Baidu's AI R&D, is known for establishing and leading Baidu IDL (Institute of Deep Learning) and Baidu Autonomous Driving research.

Hobot offers integrated AI technology and system solutions to equip thousands of categories of devices, such as smart cameras and autonomous vehicles, with "brains" that enable intelligent capabilities from perception, interaction, and understanding to decision making. Hobot aims to deliver these solutions to device manufacturers, OEMs, Tier-1 suppliers and system integrators, with unparalleled performance and cost-/power-efficiency. This is only made possible by Hobot's comprehensive suite of proprietary algorithms, software and AI inference ASICs which encompass an ARM core and a BPU (Brain Processing Unit) with a proprietary architecture designed to run AI-based algorithms in an efficient manner.

Since inception, the Company has raised approximately US$100 million from a number of blue-chip financial investors including Morningside, Hillhouse, Sequoia, GSR, and Yuri Milner. On October 20, 2017, it announced an A-plus series funding of over US$100 million led by Intel who has been collaborating closely with Hobot in multiple areas. On December 20, 2017, Hobot launched its 1st-generation ASICs (Journey 1.0 and Sunrise 1.0), which are the first embedded AI processors ever developed by a Chinese company according to Xinhua News Agency. Journey 1.0 is designed for advanced driving assistance systems (ADAS), while Sunrise 1.0 for embedded vision recognition on smart cameras. In January 2018, Hobot has successfully demonstrated to global OEMs and Tier-1 auto suppliers a prototype of its 2nd-generation processors (expected to tape out in 2H18) tested live in two cars on Las Vegas streets. The prototype was capable of processing input from multiple cameras and recognizing more than 20 categories of visual objects in real time with a substantially high accuracy level.

DeePhi Tech, founded in 2016, is a startup that provides AI solutions based on their unique deep compression algorithms, Deep-learning Processing Unit (DPU) platform, and software environment. DeePhi has their own neural network compression algorithm including pruning and quantization that can greatly reduce the network model size and computation complexity with zero or little accuracy loss. DeePhi develops their proprietary DPU and implements it on Xilinx FPGAs as well as on their SoC chip. DeePhi's technologies in algorithm compression and DPU design has received the Best Paper Award from ICLR'16 and FPGA'17, the top academic conferences in deep learning and computer architecture respectively. So far, DeePhi has 120 employees with 90 engineers

DeePhi focuses on security surveillance, data center, and automotive applications. In late 2017, DeePhi released a series of products based on FPGA, including face recognition/video analytics modules for visual applications and PCI-E card solution for surveillance and NLP server markets. The FPGA based solution uses Xilinx Zynq UltraScale+ FPGA that can provide 2.7 TOPS neural network computation capacity with less than 11 Watts power consumption. Furthermore, Deephi believes that its proprietary algorithm compression could potentially boost AI performance by 3-7x times at the same TOPS rate. DeePhi's deep learning SoC, Tingtao, integrates quad-core ARM processor and dual-core DPU and will be ready for shipment in Q218. Tingtao is fabricated with TSMC 28nm process and can provide 4TOPS neural networks computation capacity with only 3W power consumption for a broad set of applications from edge to servers. In Q418, Deephi plans to release its next generation of DPU architecture with advanced features and high flexibility.

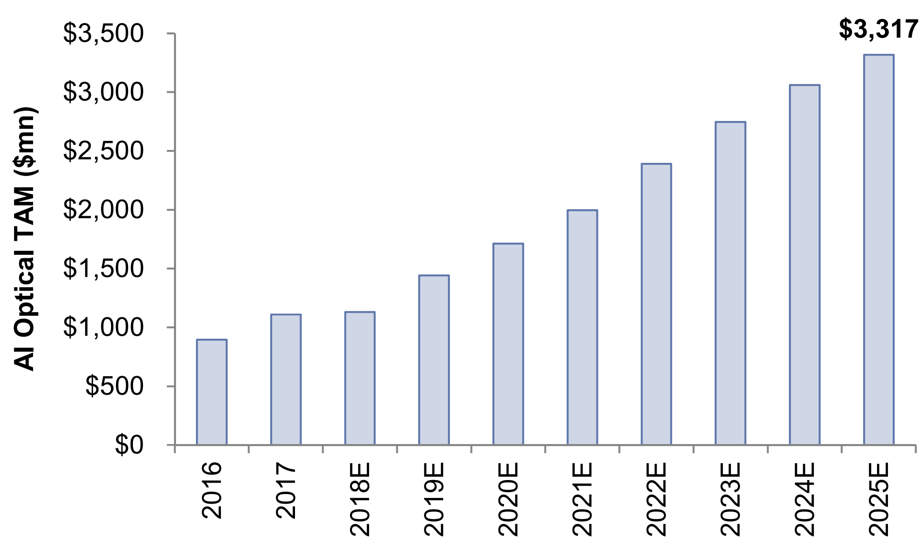
Founded only 2 years ago, Deephi has completed 3 rounds of financing with companies including Xilinx, Samsung, Ant Financial, and MediaTek.

OPTICAL: $3BN TAM

Optical components are deployed within data centers to transport traffic between servers, storage, and switches (and more broadly transport that traffic between data centers and onto service provider networks). While optical component companies are unlikely to directly participate in the functional changes associated with AI, there could be a derivative benefit as data center traffic increases. In particular, we expect AI computing demands to require higher capacity optical connections (such as 100G). As discussed in the section '**AI impacts on Networking and Storage**,' there remains some uncertainty as to how server architectures are structured.

We believe much of the generated traffic will remain inside (or intra-data center). As a result, we do believe optical component companies can benefit, but the impact may be somewhat tempered relative to other markets that supply more directly the AI ecosystem. Indeed, based on Cisco's February 2018 Global Cloud Index analysis, data center traffic is expected to grow at a 27% CAGR from 2016-2021. We estimate the data center optical components market can grow at a 7% CAGR from $2.7bn in 2017 to over $4.5bn by 2025. We assume optical unit demand is driven by server and switching port counts (and thus tied to data traffic growth). Specifically, we expect optical ports to grow at a 4% CAGR vs. server growth rate of 2%. Finally, we expect revenues per optical port to grow 2% on average per year due a higher mix of AI servers and higher capacity optical transceiver (i.e. 100G and above), offset by typical industry price declines (which run at about 10-15% per year).

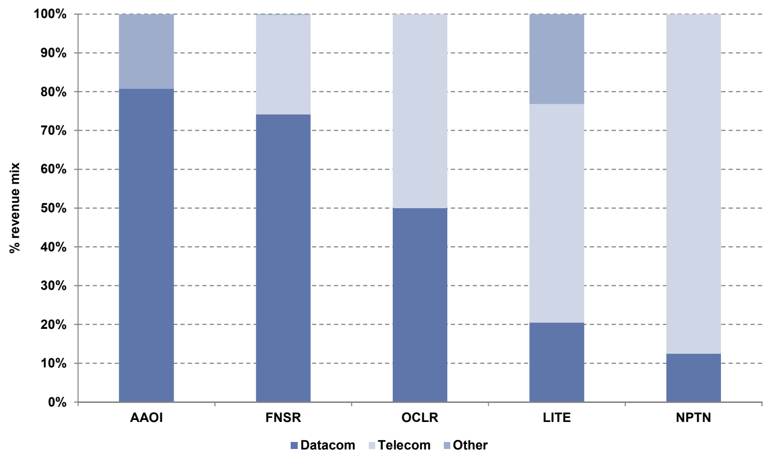
Exhibit 61: We expect the optical TAM to grow at a 7% CAGR
AI Optical TAM



Source: Goldman Sachs Global Investment Research.

The companies within our coverage of optical components (Finisar, Lumentum, Acacia) primarily sell optical transceivers and modules. In addition, there is a range of exposures split between datacom (data centers) and telecom (service provider networks). Considering Finisar and Applied Optoelectronics, both of these companies have higher datacom exposure, but much of this is for longer, inter-data center connections. Other companies that have intra-data center exposure include Innolight, Inphi, Luxtera, and Broadcom. From a technology standpoint, we will likely start to see 400G optical transceivers around 2019, which could be leveraged from high network traffic needs (and short distances) associated with AI workflows.

Exhibit 62: Optical component company revenue exposure to Datacom end markets
Optical company revenue exposure



Source: Company data, Goldman Sachs Global Investment Research.

THOUGHTS ON THE IMPACT OF AI ON NETWORKING AND STORAGE

We believe that both networking and storage systems demand are positively impacted by AI growth. However, the way that the impacts to demand manifest themselves is linked closely with exactly how the GPU computing infrastructure is deployed. In this section we detail how both pod and server architectures affect other infrastructure and we also break down Nvidia GPU shipments by those two deployment types to create a baseline idea for how infrastructure demand is moving today. We believe that AI will demand 25/100GE connectivity and we don't expect it to drive any further AI-specific technology enhancements. In terms of spending intensity, we would expect flash storage spend to be on a similar capacity/workload ratio regardless of pod or server based deployments. Much more work is required to estimate actual market impacts for networking and storage from AI, but we directionally would expect a moderate positive impact for networking and possible slightly less for storage simply because we believe our AFA models are already capturing some AI spend.

Pod-based deployments

We believe that most cloud AI and cutting edge AI is being deployed in what we refer to here as a 'supercomputer' architecture. By this, we mean large banks of discrete GPUs numbering in the hundreds and connected with PCIE like proprietary high speed bus architectures in large Pods. In this type of deployment there are 64+ TPU/GPU chips in a Pod that then require "feeding" from high speed storage for both training and inference. In our opinion, the incremental networking and storage requirement here is associated with this data feeding function. We believe our estimates for companies like Arista and NetApp probably already account for growth in AFAs and 100GE networking driven by AI deployments. However, we have not explicitly modeled for how much of current growth is represented by AI and how large this could get in the future. What we can determine for the supercomputer case is that Networking requirements will likely be materially lower per unit of compute than for a typical general purpose server configured workload due to the non-Ethernet nature of the chip to chip communications architectures being deployed within each pod. In terms of total storage required per unit of compute there is a relationship between computing throughput and total storage though we are still researching how best to quantify this.

Server deployments

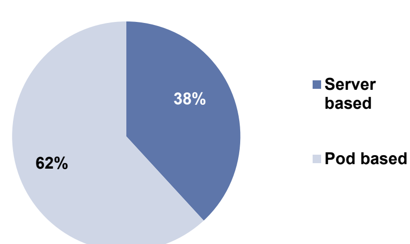
We also observe GPU capacity for both inference and training being deployed by enterprise and scientific organizations in a more traditional stacked server model. While this model doesn't yield optimal workload management and/or total system performance it is much easier for most organizations to deploy given the lack of pod-based standards that others can design to. Here the networking impacts are more similar to what we would expect in a classical general purpose computing deployment though we would assume a possible need for two 25GE ports per server to provide some bandwidth overhead as compared to a minimal design requirement of just one 25GE port per server. Storage calculations for a server oriented architecture should be similar to the pod-based case though we would expect a slightly or moderately lower storage need per GPU because of the materially lower computing efficiencies likely in the server deployed case.

The split between our two deployment scenarios

As a first step towards eventually quantifying the medium term incremental networking and storage markets associated with continued growth in AI computing capacity we estimate the current split between pod-based and server-based GPU sales. In Exhibit 63 below we estimate that about 62% of current GPU shipments are for pod-based installations and then the remaining 38% are for server-based deployment.

To arrive at this we assume that Nvidia's hyperscale revenue mix is ~60%, and HPC revenue mix is ~25%, with the rest being enterprise virtualization. For each end-market, we assume a mix for discrete units and DGX1 units, as well as a mix for server-based and pod-based units. We assume that 90% of hyperscale servers are discrete units, with the remaining 10% being DGX1 units. We assume that 90% of these hyperscale discrete units are pod-based, with the rest server-based. We also assume all DGX1 units are server-based. Next, we assume that 80% of HPC servers are discrete units, with the remaining 20% being DGX1 units. We also assume that 80% of these HPC discrete units are pod-based, with the rest server-based. Again, we assume all DGX1 units are server-based. Finally, ex-ing virtualization units, we calculate that the mix of server-based GPU shipments is 38%, vs. 62% for pod-based units.

Exhibit 63: We estimate that the majority of GPU deployments are Pod-based
Nvidia GPU server mix ex-virtualization (GSe)



Source: Goldman Sachs Global Investment Research.

Rating and Pricing Information

Intel Corp. (N/N, $50.74), Micron Technology Inc. (B/N, $55.22), Nvidia Corp. (B/N, $241.18), Samsung Electronics (B/N, W2,487,000), Samsung Electronics (Pref) (B/N, W2,080,000), SK Hynix Inc. (B/N, W83,300), TSMC (B/N, NT$250.50), TSMC (ADR) (B/N, $43.80), Western Digital Corp. (N/N, $97.22) and Xilinx Corp. (B/N, $75.69)

Financial Advisory Disclosures

Goldman Sachs and/or one of its affiliates is acting as a financial advisor in connection with an announced strategic matter involving the following company or one of its affiliates: Western Digital Corporation

Goldman Sachs and/or one of its affiliates is acting as a financial advisor in connection with an announced strategic matter involving the following company or one of its affiliates: Sk Hynix, Inc.

Disclosure Appendix

Reg AC

We, Toshiya Hari, Charles Long, Daiki Takayama, Donald Lu, Ph.D., Rod Hall, Mark Delaney, CFA, Doug Clark, CFA, Giuni Lee, Timothy Sweetnam, Balaji Krishnamurthy, CFA, Chris Jeon and Matthew Fahey, hereby certify that all of the views expressed in this report accurately reflect our personal views about the subject company or companies and its or their securities. We also certify that no part of our compensation was, is or will be, directly or indirectly, related to the specific recommendations or views expressed in this report.

Unless otherwise stated, the individuals listed on the cover page of this report are analysts in Goldman Sachs' Global Investment Research division.

GS Factor Profile

The Goldman Sachs Factor Profile provides investment context for a stock by comparing key attributes to the market (i.e. our coverage universe) and its sector peers. The four key attributes depicted are: Growth, Financial Returns, Multiple (e.g. valuation) and Integrated (a composite of Growth, Financial Returns and Multiple). Growth, Financial Returns and Multiple are calculated by using normalized ranks for specific metrics for each stock. The normalized ranks for the metrics are then averaged and converted into percentiles for the relevant attribute. The precise calculation of each metric may vary depending on the fiscal year, industry and region, but the standard approach is as follows:

Growth is based on a stock's forward-looking sales growth, EBITDA growth and EPS growth (for financial stocks, only EPS and sales growth), with a higher percentile indicating a higher growth company. **Financial Returns** is based on a stock's forward-looking ROE, ROCE and CROCI (for financial stocks, only ROE), with a higher percentile indicating a company with higher financial returns. **Multiple** is based on a stock's forward-looking P/E, P/B, price/dividend (P/D), EV/EBITDA, EV/FCF and EV/Debt Adjusted Cash Flow (DACF) (for financial stocks, only P/E, P/B and P/D), with a higher percentile indicating a stock trading at a higher multiple. The **Integrated** percentile is calculated as the average of the Growth percentile, Financial Returns percentile and (100% - Multiple percentile).

Financial Returns and Multiple use the Goldman Sachs analyst forecasts at the fiscal year-end at least three quarters in the future. Growth uses inputs for the fiscal year at least seven quarters in the future compared with the year at least three quarters in the future (on a per-share basis for all metrics).

For a more detailed description of how we calculate the GS Factor Profile, please contact your GS representative.

M&A Rank

Across our global coverage, we examine stocks using an M&A framework, considering both qualitative factors and quantitative factors (which may vary across sectors and regions) to incorporate the potential that certain companies could be acquired. We then assign a M&A rank as a means of scoring companies under our rated coverage from 1 to 3, with 1 representing high (30%-50%) probability of the company becoming an acquisition target, 2 representing medium (15%-30%) probability and 3 representing low (0%-15%) probability. For companies ranked 1 or 2, in line with our standard departmental guidelines we incorporate an M&A component into our target price. M&A rank of 3 is considered immaterial and therefore does not factor into our price target, and may or may not be discussed in research.

Quantum

Quantum is Goldman Sachs' proprietary database providing access to detailed financial statement histories, forecasts and ratios. It can be used for in-depth analysis of a single company, or to make comparisons between companies in different sectors and markets.

GS SUSTAIN

GS SUSTAIN is a global investment strategy focused on the generation of long-term alpha through identifying high quality industry leaders. The GS SUSTAIN 50 list includes leaders we believe to be well positioned to deliver long-term outperformance through superior returns on capital, sustainable competitive advantage and effective management of ESG risks vs. global industry peers. Candidates are selected largely on a combination of quantifiable analysis of these three aspects of corporate performance.

Disclosures

Coverage group(s) of stocks by primary analyst(s)

Toshiya Hari: America-Semiconductor Capital Equipment, America-Semiconductors. Daiki Takayama: Japan-Electronic Components, Korea Technology. Donald Lu, Ph.D.: A-share Telecoms and Technology, Greater China Telecoms and Technology. Rod Hall: America-Consumer Hardware & Mobility, America-IT Hardware, America-Networking. Mark Delaney, CFA: America-IT Supply Chain: Components, America-IT Supply Chain: Distributors, America-IT Supply Chain: Drives, America-IT Supply Chain: EMS, America-Semi Devices. Doug Clark, CFA: America-Consumer Hardware & Mobility, America-IT Hardware, America-Networking. Giuni Lee: Korea Technology.

A-share Telecoms and Technology: Accelink Technologies, China United Network Comm, Dahua Technology, Fiberhome Telecom Tech, Goodix, Hikvision Digital Technology, Hytera Communications Corp., SG Micro Corp., Wangsu Science & Technology Co., ZTE Corp. (A).

America-Consumer Hardware & Mobility: Apple Inc., Corning Inc., Garmin Ltd., GoPro Inc..

America-IT Hardware: Aerohive Networks Inc., CDW Corp., Hewlett Packard Enterprise Co., HP Inc., Motorola Solutions Inc., NetApp Inc., Nutanix Inc., Presidio Inc., Pure Storage Inc., Xerox Corp..

America-IT Supply Chain: Components: Amphenol Corp., Belden Inc., CommScope Holding, Sensata Technologies Holding, TE Connectivity Ltd..

America-IT Supply Chain: Distributors: Arrow Electronics Inc., Avnet Inc..

America-IT Supply Chain: Drives: Seagate Technology, Western Digital Corp..

America-IT Supply Chain: EMS: Flex, Jabil Circuit Inc..

America-Networking: Acacia Communications Inc., ADTRAN Inc., Arista Networks Inc., ARRIS International Plc, Ciena Corp., Cisco Systems Inc., F5 Networks Inc., Finisar Corp., Infinera Corp., Juniper Networks Inc., Lumentum Holdings.

America-Semi Devices: MACOM Technology Solutions Holding, Marvell Technology Group, Microchip Technology Inc., Micron Technology Inc., Microsemi Corp., ON Semiconductor Corp..

America-Semiconductor Capital Equipment: Applied Materials Inc., Entegris Inc., Keysight Technologies Inc., KLA-Tencor Corp., Lam Research Corp., Teradyne Inc., Versum Materials Inc..

America-Semiconductors: Advanced Micro Devices Inc., Analog Devices Inc., Broadcom Ltd., Integrated Device Technology Inc., Intel Corp., Maxim Integrated Products, Nvidia Corp., NXP Semiconductors NV, Qorvo Inc., Skyworks Solutions Inc., Texas Instruments Inc., Xilinx Corp..

Greater China Telecoms and Technology: China Communication Services, China Mobile (HK), China Mobile (HK) (ADR), China Telecom (ADR), China Telecom (H), China Unicom (ADS), China Unicom (H), Hua Hong Semiconductor Ltd., Mediatek, Parade Technologies Ltd., Silergy Corp., SMIC (ADR), SMIC (H), TSMC, TSMC (ADR), United Microelectronics Corp., United Microelectronics Corp. (ADR), ZTE Corp. (H).

Japan-Electronic Components: Alps Electric, Hirose Electric, Ibiden, IRISO Electronics, Japan Aviation Electronics Industry, Japan Display Inc., Kyocera, Mabuchi Motor, MinebeaMitsumi Inc., Murata Mfg., NGK Insulators, NGK Spark Plug, Nichicon, Nidec, Nippon Ceramic, Nitto Denko, Pacific Industrial, Shinko Electric Industries, Taiyo Yuden, TDK.

Korea Technology: LG Display, LG Electronics, Samsung Electro-Mechanics, Samsung Electronics, Samsung Electronics (Pref), Samsung SDI Co., Samsung SDS Co., SK Hynix Inc..

Company-specific regulatory disclosures

Compendium report: please see disclosures at http://www.gs.com/research/hedge.html. Disclosures applicable to the companies included in this compendium can be found in the latest relevant published research

Distribution of ratings/investment banking relationships

Goldman Sachs Investment Research global Equity coverage universe

	Rating Distribution			Investment Banking Relationships		
	Buy	Hold	Sell	Buy	Hold	Sell
Global	33%	54%	13%	63%	57%	52%

As of January 1, 2018, Goldman Sachs Global Investment Research had investment ratings on 2,867 equity securities. Goldman Sachs assigns stocks as Buys and Sells on various regional Investment Lists; stocks not so assigned are deemed Neutral. Such assignments equate to Buy, Hold and Sell for the purposes of the above disclosure required by the FINRA Rules. See 'Ratings, Coverage groups and views and related definitions' below. The Investment Banking Relationships chart reflects the percentage of subject companies within each rating category for whom Goldman Sachs has provided investment banking services within the previous twelve months.

Price target and rating history chart(s)

Compendium report: please see disclosures at http://www.gs.com/research/hedge.html. Disclosures applicable to the companies included in this compendium can be found in the latest relevant published research

Regulatory disclosures

Disclosures required by United States laws and regulations

See company-specific regulatory disclosures above for any of the following disclosures required as to companies referred to in this report: manager or co-manager in a pending transaction; 1% or other ownership; compensation for certain services; types of client relationships; managed/co-managed public offerings in prior periods; directorships; for equity securities, market making and/or specialist role. Goldman Sachs trades or may trade as a principal in debt securities (or in related derivatives) of issuers discussed in this report.

The following are additional required disclosures: **Ownership and material conflicts of interest:** Goldman Sachs policy prohibits its analysts, professionals reporting to analysts and members of their households from owning securities of any company in the analyst's area of coverage. **Analyst compensation:** Analysts are paid in part based on the profitability of Goldman Sachs, which includes investment banking revenues. **Analyst as officer or director:** Goldman Sachs policy generally prohibits its analysts, persons reporting to analysts or members of their households from serving as an officer, director or advisor of any company in the analyst's area of coverage. **Non-U.S. Analysts:** Non-U.S. analysts may not be associated persons of Goldman Sachs & Co. LLC and therefore may not be subject to FINRA Rule 2241 or FINRA Rule 2242 restrictions on communications with subject company, public appearances and trading securities held by the analysts.

Distribution of ratings: See the distribution of ratings disclosure above. **Price chart:** See the price chart, with changes of ratings and price targets in prior periods, above, or, if electronic format or if with respect to multiple companies which are the subject of this report, on the Goldman Sachs website at http://www.gs.com/research/hedge.html.

Additional disclosures required under the laws and regulations of jurisdictions other than the United States

The following disclosures are those required by the jurisdiction indicated, except to the extent already made above pursuant to United States laws and regulations. **Australia:** Goldman Sachs Australia Pty Ltd and its affiliates are not authorised deposit-taking institutions (as that term is defined in the Banking Act 1959 (Cth)) in Australia and do not provide banking services, nor carry on a banking business, in Australia. This research, and any access to it, is intended only for "wholesale clients" within the meaning of the Australian Corporations Act, unless otherwise agreed by Goldman Sachs. In producing research reports, members of the Global Investment Research Division of Goldman Sachs Australia may attend site visits and other meetings hosted by the companies and other entities which are the subject of its research reports. In some instances the costs of such site visits or meetings may be met in part or in whole by the issuers concerned if Goldman Sachs Australia considers it is appropriate and reasonable in the specific circumstances relating to the site visit or meeting. To the extent that the contents of this document contains any financial product advice, it is general advice only and has been prepared by Goldman Sachs without taking into account a client's objectives, financial situation or needs. A client should, before acting on any such advice, consider the appropriateness of the advice having regard to the client's own objectives, financial situation and needs. **Brazil:** Disclosure information in relation to CVM Instruction 483 is available at http://www.gs.com/worldwide/brazil/area/gir/index.html. Where applicable, the Brazil-registered analyst primarily responsible for the content of this research report, as defined in Article 16 of CVM Instruction 483, is the first author named at the beginning of this report, unless indicated otherwise at the end of the text. **Canada:** Goldman Sachs Canada Inc. is an affiliate of The Goldman Sachs Group Inc. and therefore is included in the company specific disclosures relating to Goldman Sachs (as defined above). Goldman Sachs Canada Inc. has approved of, and agreed to take responsibility for, this research report in Canada if and to the extent that Goldman Sachs Canada Inc. disseminates this research report to its clients. **Hong Kong:** Further information on the securities of covered companies referred to in this research may be obtained on request from Goldman Sachs (Asia) L.L.C. **India:** Further information on the subject company or companies referred to in this research may be obtained from Goldman Sachs (India) Securities Private Limited, Research Analyst - SEBI Registration Number INH000001493, 951-A, Rational House, Appasaheb Marathe Marg, Prabhadevi, Mumbai 400 025, India, Corporate Identity Number U74140MH2006FTC160634, Phone +91 22 6616 9000, Fax +91 22 6616 9001. Goldman Sachs may beneficially own 1% or more of the securities (as such term is defined in clause 2 (h) the Indian Securities Contracts (Regulation) Act,

1956) of the subject company or companies referred to in this research report. **Japan:** See below. **Korea:** Further information on the subject company or companies referred to in this research may be obtained from Goldman Sachs (Asia) L.L.C., Seoul Branch. **New Zealand:** Goldman Sachs New Zealand Limited and its affiliates are neither "registered banks" nor "deposit takers" (as defined in the Reserve Bank of New Zealand Act 1989) in New Zealand. This research, and any access to it, is intended for "wholesale clients" (as defined in the Financial Advisers Act 2008) unless otherwise agreed by Goldman Sachs. **Russia:** Research reports distributed in the Russian Federation are not advertising as defined in the Russian legislation, but are information and analysis not having product promotion as their main purpose and do not provide appraisal within the meaning of the Russian legislation on appraisal activity. **Singapore:** Further information on the covered companies referred to in this research may be obtained from Goldman Sachs (Singapore) Pte. (Company Number: 198602165W). **Taiwan:** This material is for reference only and must not be reprinted without permission. Investors should carefully consider their own investment risk. Investment results are the responsibility of the individual investor. **United Kingdom:** Persons who would be categorized as retail clients in the United Kingdom, as such term is defined in the rules of the Financial Conduct Authority, should read this research in conjunction with prior Goldman Sachs research on the covered companies referred to herein and should refer to the risk warnings that have been sent to them by Goldman Sachs International. A copy of these risks warnings, and a glossary of certain financial terms used in this report, are available from Goldman Sachs International on request.

European Union: Disclosure information in relation to Article 4 (1) (d) and Article 6 (2) of the European Commission Directive 2003/125/EC is available at http://www.gs.com/disclosures/europeanpolicy.html which states the European Policy for Managing Conflicts of Interest in Connection with Investment Research.

Japan: Goldman Sachs Japan Co., Ltd. is a Financial Instrument Dealer registered with the Kanto Financial Bureau under registration number Kinsho 69, and a member of Japan Securities Dealers Association, Financial Futures Association of Japan and Type II Financial Instruments Firms Association. Sales and purchase of equities are subject to commission pre-determined with clients plus consumption tax. See company-specific disclosures as to any applicable disclosures required by Japanese stock exchanges, the Japanese Securities Dealers Association or the Japanese Securities Finance Company.

Ratings, coverage groups and views and related definitions

Buy (B), Neutral (N), Sell (S) -Analysts recommend stocks as Buys or Sells for inclusion on various regional Investment Lists. Being assigned a Buy or Sell on an Investment List is determined by a stock's total return potential relative to its coverage. Any stock not assigned as a Buy or a Sell on an Investment List with an active rating (i.e., a stock that is not Rating Suspended, Not Rated, Coverage Suspended or Not Covered), is deemed Neutral. Each regional Investment Review Committee manages various regional Investment Lists to a global guideline of 25%-35% of stocks as Buy and 10%-15% of stocks as Sell; however, the distribution of Buys and Sells in any particular analyst's coverage group may vary as determined by the regional Investment Review Committee. Additionally, each Investment Review Committee manages Regional Conviction lists, which represent investment recommendations focused on the size of the total return potential and/or the likelihood of the realization of the return across their respective areas of coverage. The addition or removal of stocks from such Conviction lists do not represent a change in the analysts' investment rating for such stocks.

Total return potential represents the upside or downside differential between the current share price and the price target, including all paid or anticipated dividends, expected during the time horizon associated with the price target. Price targets are required for all covered stocks. The total return potential, price target and associated time horizon are stated in each report adding or reiterating an Investment List membership.

Coverage groups and views: A list of all stocks in each coverage group is available by primary analyst, stock and coverage group at http://www.gs.com/research/hedge.html. The analyst assigns one of the following coverage views which represents the analyst's investment outlook on the coverage group relative to the group's historical fundamentals and/or valuation. **Attractive (A).** The investment outlook over the following 12 months is favorable relative to the coverage group's historical fundamentals and/or valuation. **Neutral (N).** The investment outlook over the following 12 months is neutral relative to the coverage group's historical fundamentals and/or valuation. **Cautious (C).** The investment outlook over the following 12 months is unfavorable relative to the coverage group's historical fundamentals and/or valuation.

Not Rated (NR). The investment rating and target price have been removed pursuant to Goldman Sachs policy when Goldman Sachs is acting in an advisory capacity in a merger or strategic transaction involving this company and in certain other circumstances. **Rating Suspended (RS).** Goldman Sachs Research has suspended the investment rating and price target for this stock, because there is not a sufficient fundamental basis for determining, or there are legal, regulatory or policy constraints around publishing, an investment rating or target. The previous investment rating and price target, if any, are no longer in effect for this stock and should not be relied upon. **Coverage Suspended (CS).** Goldman Sachs has suspended coverage of this company. **Not Covered (NC).** Goldman Sachs does not cover this company. **Not Available or Not Applicable (NA).** The information is not available for display or is not applicable. **Not Meaningful (NM).** The information is not meaningful and is therefore excluded.

Global product; distributing entities

The Global Investment Research Division of Goldman Sachs produces and distributes research products for clients of Goldman Sachs on a global basis. Analysts based in Goldman Sachs offices around the world produce equity research on industries and companies, and research on macroeconomics, currencies, commodities and portfolio strategy. This research is disseminated in Australia by Goldman Sachs Australia Pty Ltd (ABN 21 006 797 897); in Brazil by Goldman Sachs do Brasil Corretora de Títulos e Valores Mobiliários S.A.; Ombudsman Goldman Sachs Brazil: 0800 727 5764 and / or ouvidoriagoldmansachs@gs.com. Available Weekdays (except holidays), from 9am to 6pm. Ouvidoria Goldman Sachs Brasil: 0800 727 5764 e/ou ouvidoriagoldmansachs@gs.com. Horário de funcionamento: segunda-feira à sexta-feira (exceto feriados), das 9h às 18h; in Canada by either Goldman Sachs Canada Inc. or Goldman Sachs & Co. LLC; in Hong Kong by Goldman Sachs (Asia) L.L.C.; in India by Goldman Sachs (India) Securities Private Ltd.; in Japan by Goldman Sachs Japan Co., Ltd.; in the Republic of Korea by Goldman Sachs (Asia) L.L.C., Seoul Branch; in New Zealand by Goldman Sachs New Zealand Limited; in Russia by OOO Goldman Sachs; in Singapore by Goldman Sachs (Singapore) Pte. (Company Number: 198602165W); and in the United States of America by Goldman Sachs & Co. LLC. Goldman Sachs International has approved this research in connection with its distribution in the United Kingdom and European Union.

European Union: Goldman Sachs International authorised by the Prudential Regulation Authority and regulated by the Financial Conduct Authority and the Prudential Regulation Authority, has approved this research in connection with its distribution in the European Union and United Kingdom; Goldman Sachs AG and Goldman Sachs International Zweigniederlassung Frankfurt, regulated by the Bundesanstalt für Finanzdienstleistungsaufsicht, may also distribute research in Germany.

General disclosures

This research is for our clients only. Other than disclosures relating to Goldman Sachs, this research is based on current public information that we consider reliable, but we do not represent it is accurate or complete, and it should not be relied on as such. The information, opinions, estimates and forecasts contained herein are as of the date hereof and are subject to change without prior notification. We seek to update our research as appropriate, but various regulations may prevent us from doing so. Other than certain industry reports published on a periodic basis, the large majority of reports are published at irregular intervals as appropriate in the analyst's judgment.

Goldman Sachs conducts a global full-service, integrated investment banking, investment management, and brokerage business. We have investment banking and other business relationships with a substantial percentage of the companies covered by our Global Investment Research Division. Goldman Sachs & Co. LLC, the United States broker dealer, is a member of SIPC (http://www.sipc.org).

Our salespeople, traders, and other professionals may provide oral or written market commentary or trading strategies to our clients and principal trading desks that reflect opinions that are contrary to the opinions expressed in this research. Our asset management area, principal trading desks and investing businesses may make investment decisions that are inconsistent with the recommendations or views expressed in this research.

The analysts named in this report may have from time to time discussed with our clients, including Goldman Sachs salespersons and traders, or may discuss in this report, trading strategies that reference catalysts or events that may have a near-term impact on the market price of the equity securities discussed in this report, which impact may be directionally counter to the analyst's published price target expectations for such stocks. Any such trading strategies are distinct from and do not affect the analyst's fundamental equity rating for such stocks, which rating reflects a stock's return potential relative to its coverage group as described herein.

We and our affiliates, officers, directors, and employees, excluding equity and credit analysts, will from time to time have long or short positions in, act as principal in, and buy or sell, the securities or derivatives, if any, referred to in this research.

The views attributed to third party presenters at Goldman Sachs arranged conferences, including individuals from other parts of Goldman Sachs, do not necessarily reflect those of Global Investment Research and are not an official view of Goldman Sachs.

Any third party referenced herein, including any salespeople, traders and other professionals or members of their household, may have positions in the products mentioned that are inconsistent with the views expressed by analysts named in this report.

This research is not an offer to sell or the solicitation of an offer to buy any security in any jurisdiction where such an offer or solicitation would be illegal. It does not constitute a personal recommendation or take into account the particular investment objectives, financial situations, or needs of individual clients. Clients should consider whether any advice or recommendation in this research is suitable for their particular circumstances and, if appropriate, seek professional advice, including tax advice. The price and value of investments referred to in this research and the income from them may fluctuate. Past performance is not a guide to future performance, future returns are not guaranteed, and a loss of original capital may occur. Fluctuations in exchange rates could have adverse effects on the value or price of, or income derived from, certain investments.

Certain transactions, including those involving futures, options, and other derivatives, give rise to substantial risk and are not suitable for all investors. Investors should review current options disclosure documents which are available from Goldman Sachs sales representatives or at http://www.theocc.com/about/publications/character-risks.jsp. Transaction costs may be significant in option strategies calling for multiple purchase and sales of options such as spreads. Supporting documentation will be supplied upon request.

Differing Levels of Service provided by Global Investment Research: The level and types of services provided to you by the Global Investment Research division of GS may vary as compared to that provided to internal and other external clients of GS, depending on various factors including your individual preferences as to the frequency and manner of receiving communication, your risk profile and investment focus and perspective (e.g., marketwide, sector specific, long term, short term), the size and scope of your overall client relationship with GS, and legal and regulatory constraints. As an example, certain clients may request to receive notifications when research on specific securities is published, and certain clients may request that specific data underlying analysts' fundamental analysis available on our internal client websites be delivered to them electronically through data feeds or otherwise. No change to an analyst's fundamental research views (e.g., ratings, price targets, or material changes to earnings estimates for equity securities), will be communicated to any client prior to inclusion of such information in a research report broadly disseminated through electronic publication to our internal client websites or through other means, as necessary, to all clients who are entitled to receive such reports.

All research reports are disseminated and available to all clients simultaneously through electronic publication to our internal client websites. Not all research content is redistributed to our clients or available to third-party aggregators, nor is Goldman Sachs responsible for the redistribution of our research by third party aggregators. For research, models or other data related to one or more securities, markets or asset classes (including related services) that may be available to you, please contact your GS representative or go to http://360.gs.com.

Disclosure information is also available at http://www.gs.com/research/hedge.html or from Research Compliance, 200 West Street, New York, NY 10282.

Mindcraft: Our Thematic Deep Dives

Innovation & Disruption

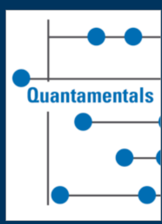
Virtual Reality | Drones | Factory of the Future | Blockchain | Precision Farming | Advanced Materials | Space | Artificial Intelligence | The Store of the Future

Rethinking Mobility | Unfolding OLED | Gaming Consoles

Insights & Policy

Top of Mind | Fortnightly Thoughts | US Infrastructure | Healthcare's Holy Grail | Trade Policy

Quantamentals



Commodity Corner

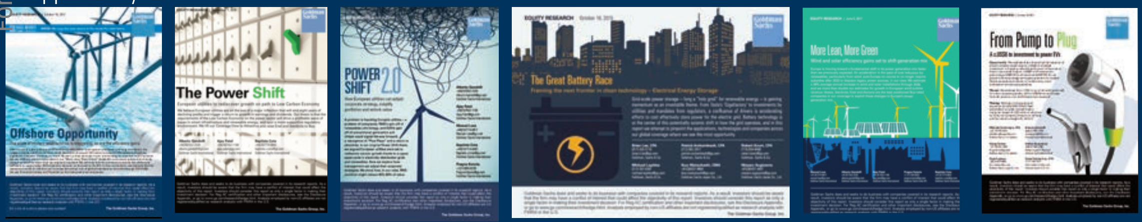
Top Oil & Gas Projects | Shale Innovation | Copper Top Projects | Big Oil Rises

Music's Return to Growth

Bear Necessities

The Low Carbon Economy

Offshore Opportunity | New Era for Euro Utilities | Power Shift 2.0 | The Great Battery Race | More Lean, More Green | Pump to Plug



Consumer Currents

Supermarket Shift: Expect Measured Disruption

Rising Asia

Apple Suppliers' Dilemma | The Asia Stock Collection | Rise of China FinTech | China AI

GS SUSTAIN

The PM's Guide To The ESG Revolution | GS SUSTAIN 50

Capex Upturn

Capex Tracker

... and more in **The Mindcraft Library**