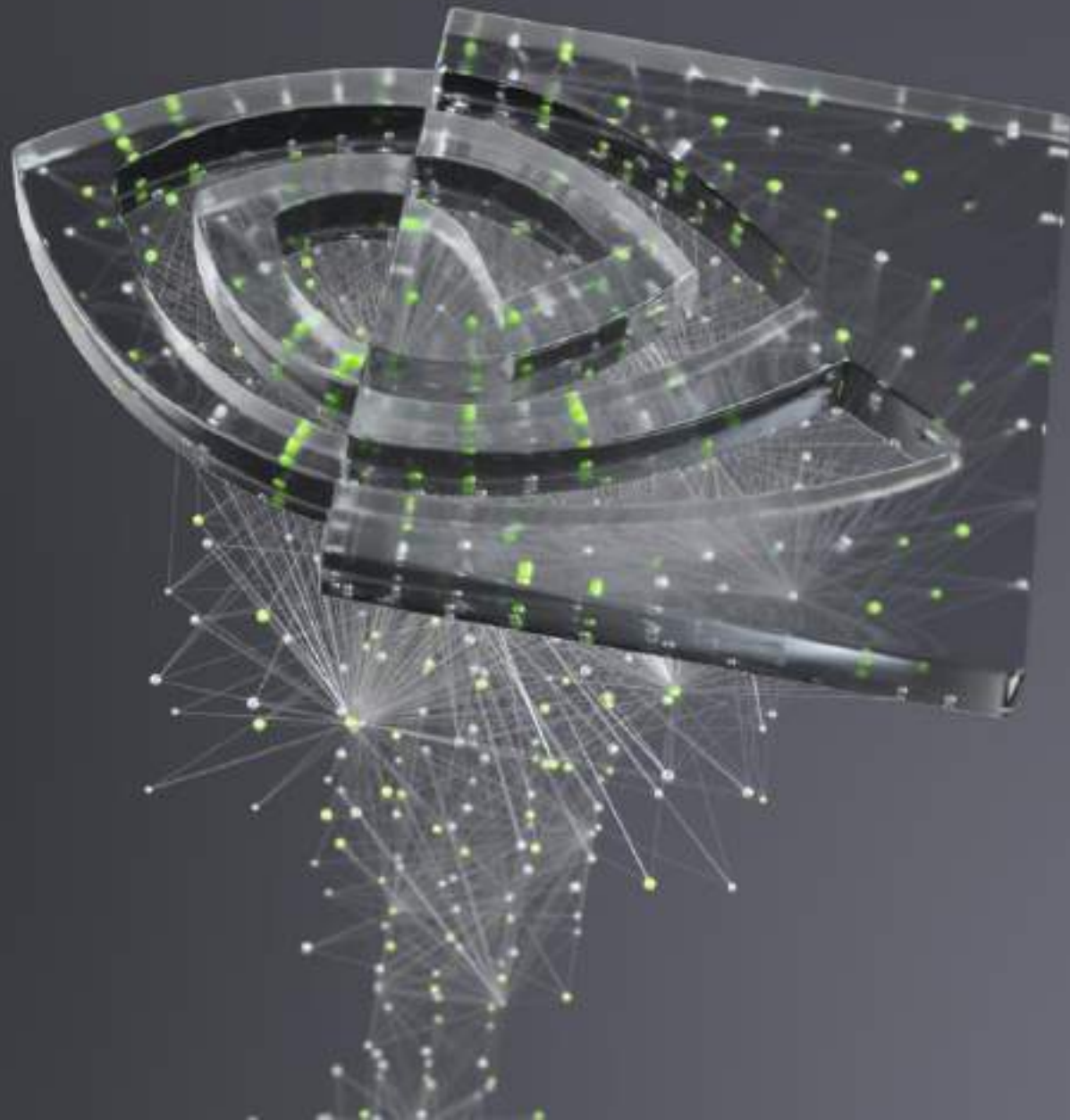




NVIDIA ENTERPRISE

Emilio Tamez

Julio 2020



“ES OFICIAL: NVIDIA YA NO ES SOLO UNA EMPRESA DE JUEGOS”

DATA CENTER KNOWLEDGE

Los data centers modernos ejecutan cargas de trabajo de una escala y una complejidad increíbles, impulsados por las aplicaciones de IA y machine learning, que aprovechan millones de usuarios en simultáneo. Estas aplicaciones son tan grandes que ningún servidor, por poderoso que sea, puede procesarlas.

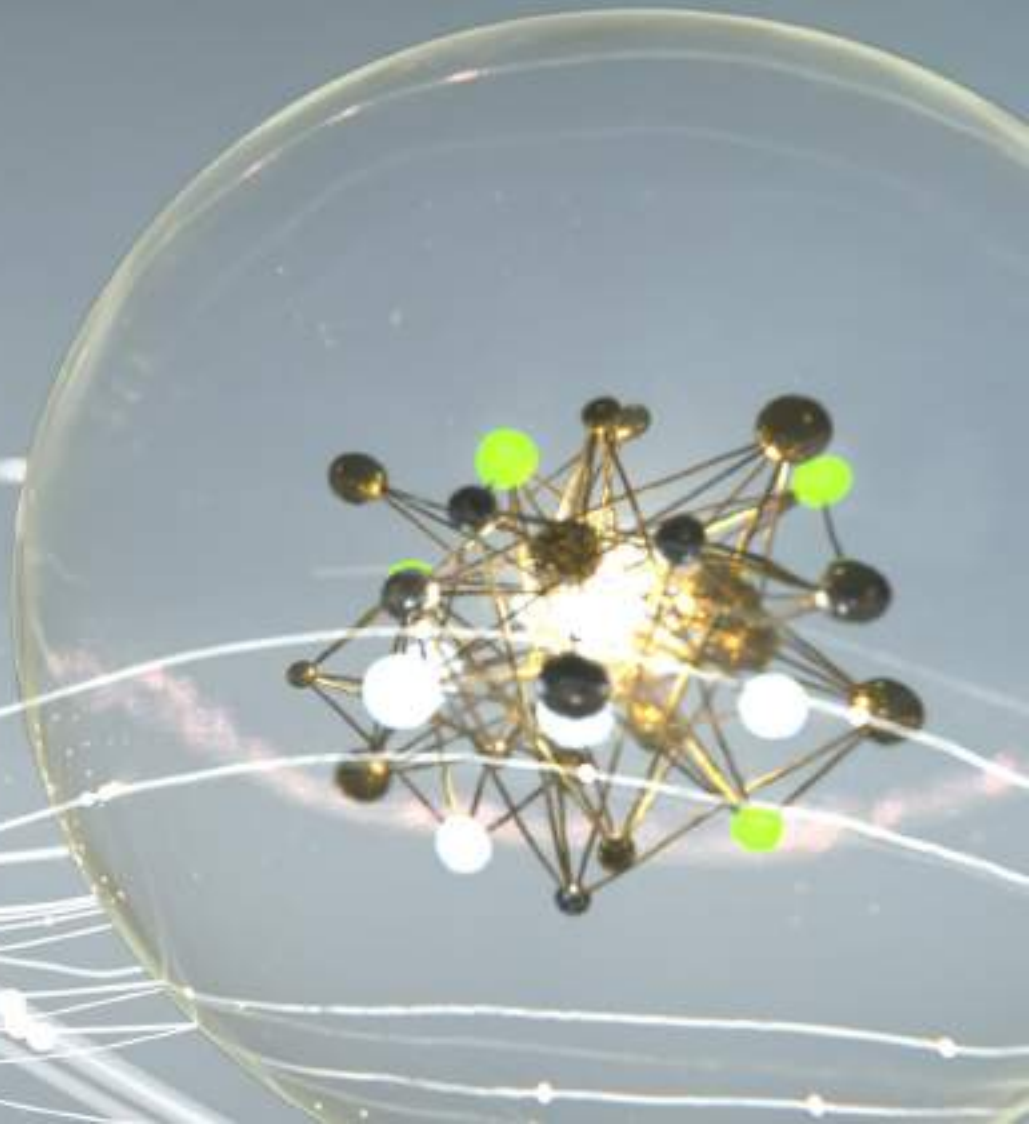
Con Mellanox, podemos aplicar innovaciones por completo, desde el procesamiento, las redes y el almacenamiento hasta seguridad, en todo el data center. Y en todo el conjunto, los desde chips hasta los servidores, los conmutadores de red y el software sofisticado.



LA NUEVA UNIDAD DE PROCESAMIENTO PARA EL DATA CENTER MODERNO

Solíamos pensar en un servidor con CPU como la unidad básica de procesamiento. Sin embargo, para satisfacer las demandas de las cargas de trabajo de IA y machine learning de la actualidad, debemos optimizar todo el data center de punta a punta. Las redes, el almacenamiento, el procesamiento y la seguridad.

La nueva unidad de procesamiento es el data center en sí mismo.



“NVIDIA ahora disfruta de un ecosistema robusto y autosustentable de software, universidades, empresas emergentes y socios que le han permitido convertirse en el amo de su propio universo recién creado.”

FORBES

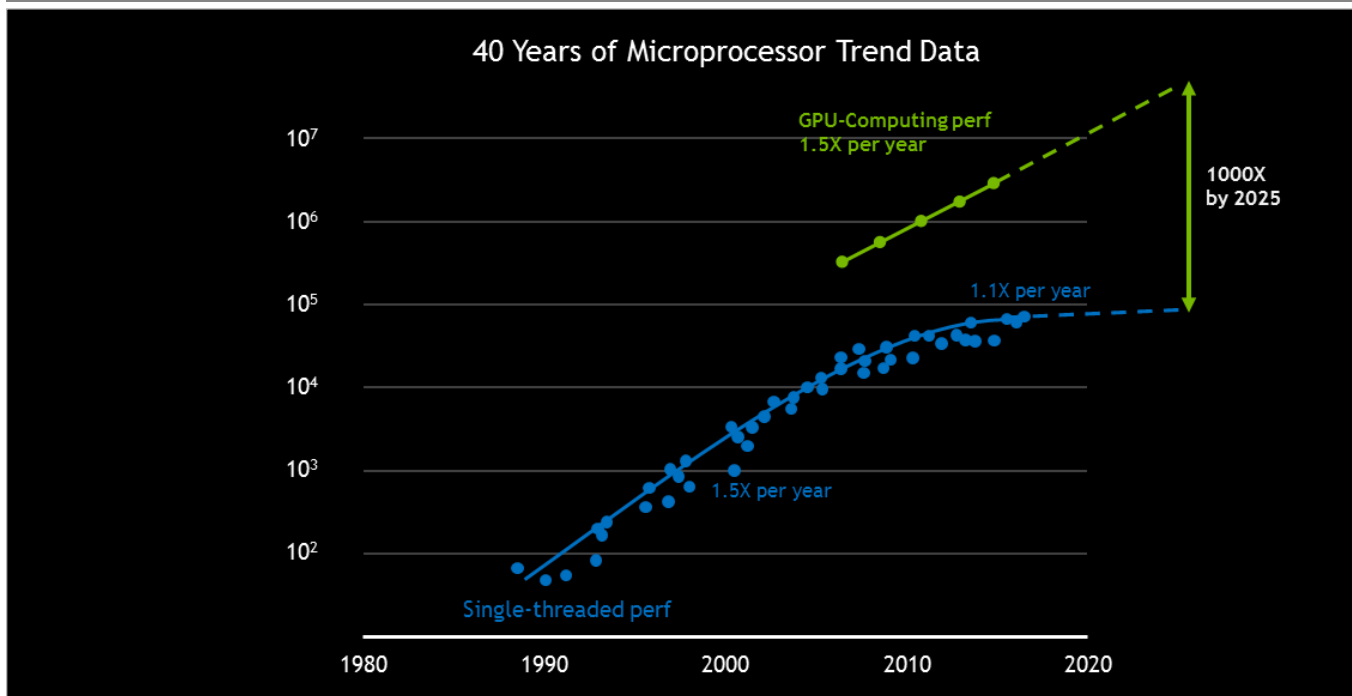
1.8M
Developers



GPU COMPUTING AT THE HEART OF AI

New Advancements Leapfrog Moore's Law

Performance Beyond Moore's Law



Big Bang of Modern AI



NVIDIA PRODUCTS & MARKETS

Strength Across Multiple Domains



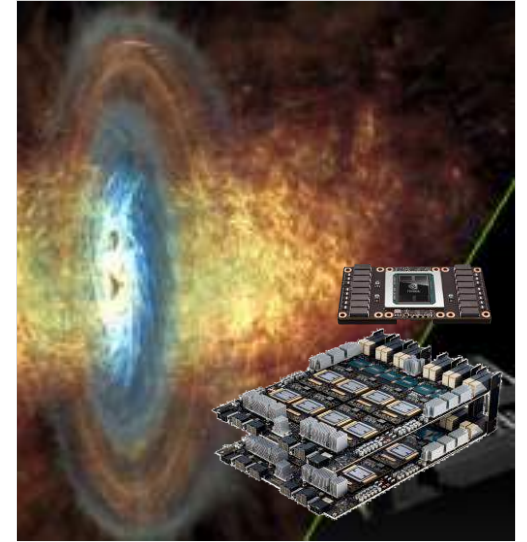
Robotics, Auto
Jetson



Consumer, Gaming
GeForce



Professional Visualization
Quadro



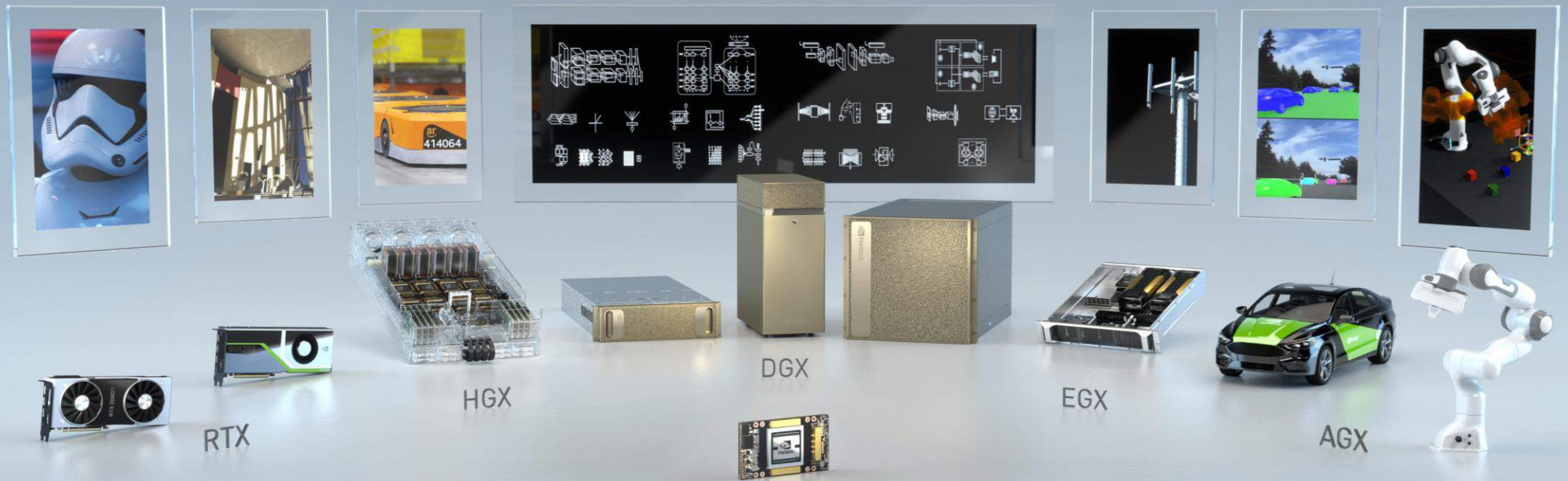
HPC, Datacenter, Cloud
Tesla

“Hemos recorrido un largo camino desde la invención de la GPU para acelerar los juegos hasta la reinención de la GPU para que sea el coprocesador más diverso y poderoso que hayamos visto”.

LA PRÓXIMA PLATAFORMA

NVIDIA fue pionera en el procesamiento acelerado para hacer frente a los desafíos que las computadoras comunes no pueden resolver. Hacemos computadoras para que los da Vincis y Einsteins de nuestro tiempo puedan ver y crear el futuro.

El procesamiento acelerado requiere algo más que un chip potente. Logramos aceleraciones increíbles a través de la invención del conjunto completo, desde el chip y los sistemas hasta los algoritmos y las aplicaciones que ejecutan.



HPC & AI TRANSFORMS INDUSTRIES



Healthcare



Industrial



Consumer
Internet



Advertising



Automotive



Media and
Entertainment



Retail



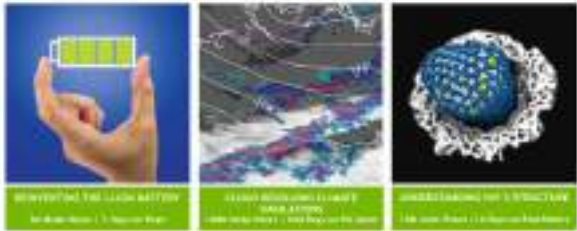
Financial

ENTERPRISE NVIDIA SOLUTIONS

HPC

High Performance Computing

GRAND CHALLENGES REQUIRE MASSIVE COMPUTING



Artificial Intelligence

AI TO TRANSFORM EVERY INDUSTRY



PLATFORM BUILT FOR AI

Accelerating Every Framework And Fueling Innovation



NVIDIA GRID (Virtualization GPU)



UNLOCKING NEW USE CASES FOR GRID

- New NVIDIA GRID Features
- + 2 GB profile to support "low normal" for knowledge workers
- + Linux VDI for developers
- + Multiple Workloads: up to two VMs on one or four HD workloads





THE LEARNING MACHINES

Using massive amounts of data to recognize photos and speech, deep-learning computers are taking a big step towards true artificial intelligence.

BY NICOLA JONES

Three years ago, researchers at the secretive Google X lab in Mountain View, California, extracted some 10 million still images from YouTube videos and fed them into Google Brain—a network of 1,000 computers programmed to scan up the world much as a human toddler does. After three days looking for recurring patterns, Google Brain decided, all on its own, that there were certain repeating categories it could identify: human faces, human bodies and... cats.

Google Brain's discovery that the Internet is full of cat videos provoked a flurry of letters from journalists. But it was also a landmark in the resurgence of deep learning, a three-decade-old technique in which massive amounts of data and processing power

help computers to crack some problems that humans solve almost intuitively, from recognizing faces to understanding language.

Deep learning itself is a revival of an even older idea for computing: neural networks. These systems, loosely inspired by the densely interconnected neurons of the brain, mimic human learning by changing the strength of simulated neural connections on the basis of experience. Google Brain, with about 1 million simulated neurons and 1 billion simulated connections, was ten times larger than any deep neural network before it. Project founder Andrew Ng, now director of the Artificial Intelligence Laboratory at Stanford University in California, has gone on to make deep-learning systems ten times larger again.

Such advances made for exciting times in

TESLA REVOLUTIONIZES DEEP LEARNING

GOOGLE BRAIN APPLICATION

	BEFORE TESLA	AFTER TESLA
Cost	\$5,000K	\$200K
Servers	1,000 Servers	16 Tesla Servers
Energy	600 KW	4 KW
Performance	1x	6x



Images Per Sec: 4.7

“EL NUEVO SISTEMA DGX A100 DE NVIDIA OFRECE 5 PETAFLUPS DE RENDIMIENTO EN UN SOLO NODO”

TWEAKTOWN

NVIDIA DGX™ es un sistema diseñado internamente, desde cero, y específicamente para la inteligencia artificial y la ciencia de datos de vanguardia.

NVIDIA DGX A100 consta de 8 GPU A100 y 9 unidades de procesamiento de datos Mellanox CX6 InfiniBand. Un solo sistema DGX tiene el rendimiento equivalente de 150 servidores de alta gama que cuestan más de un millón de dólares. La DGX A100 se vende por \$199,000. Está en plena producción y se envía a todo el mundo.



TODAY'S AI DATA CENTER

- ▶ 50 DGX-1 systems for AI training
- ▶ 600 CPU systems for AI inference
- ▶ \$11M
- ▶ 25 racks
- ▶ 630 kW



DGX A100 DATA CENTER

- ▶ 5 DGX A100 systems for AI training and inference
- ▶ \$1M
- ▶ 1 rack
- ▶ 28 kW



\$1M 28 kW

$1/10^{\text{th}}$
COST

$1/20^{\text{th}}$
POWER

THE EXPANDING UNIVERSE OF MODERN AI

THE BIG BANG

Big Data
Algorithms
GPU

RESEARCH



FRAMEWORKS



AI-25-18 PLATFORM



START-UPS



INDUSTRY LEADERS



4,000+ AI START-UPS
\$33B IN FUNDING

Source: Crunchbase & Pitchbook

AI-ACCELERATED CYBER DEFENSE

Our daily life, economic vitality, and national security depend on a stable, safe and resilient cyberspace. But attacks on IT systems are becoming more complex and relentless, resulting in loss of information and money and disruptions to essential services. Accenture's dedicated cyber security lab uses NVIDIA GPUs, CUDA libraries, and machine learning to accelerate the analysis and visualization of 200M-300M alerts daily so analysts can take timely action.

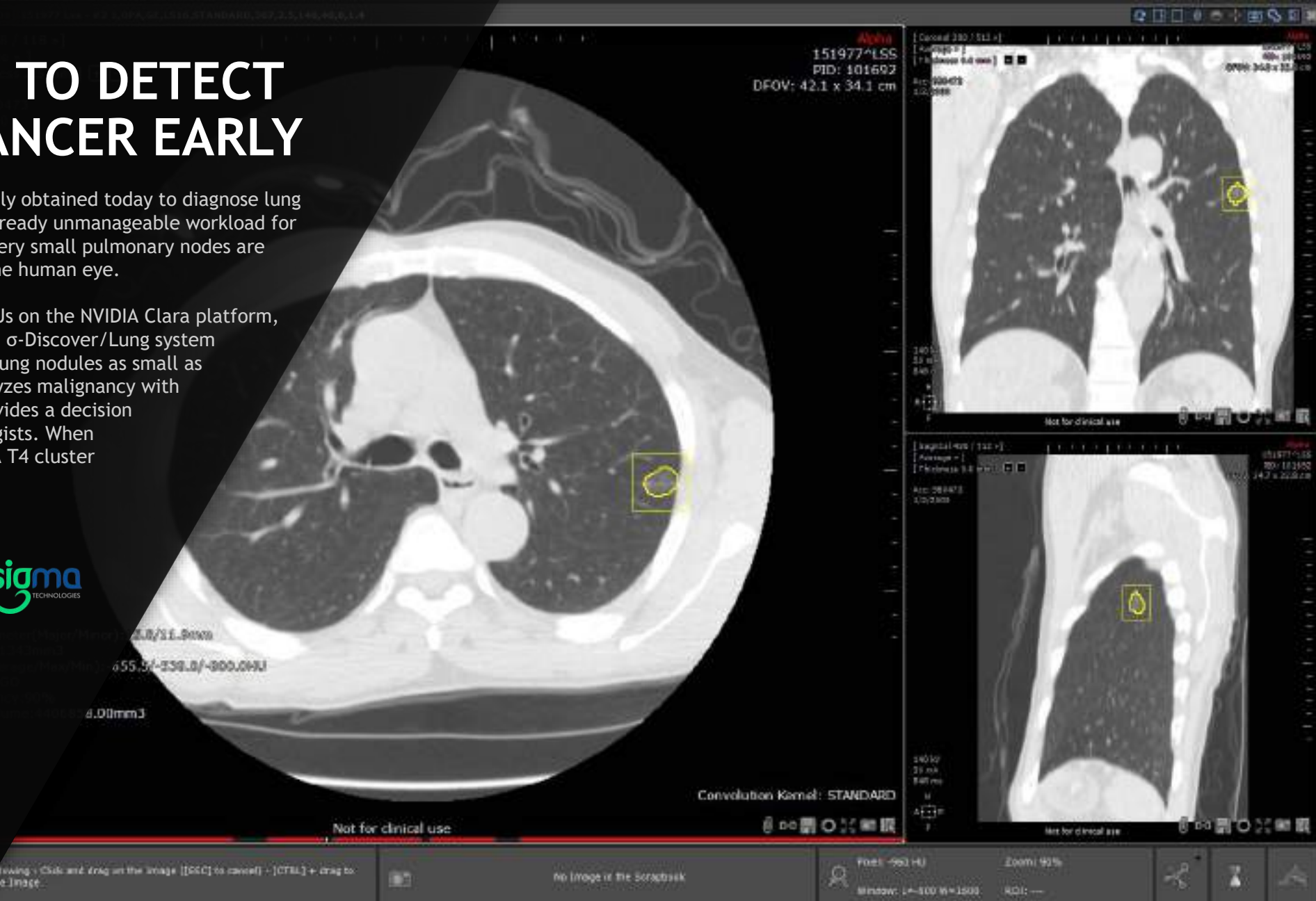
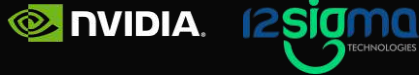


VIRUS
DETECTED

USING AI TO DETECT LUNG CANCER EARLY

CT scans are increasingly obtained today to diagnose lung cancer, adding to an already unmanageable workload for radiologists. Further, very small pulmonary nodules are difficult to spot with the human eye.

Powered by NVIDIA GPUs on the NVIDIA Clara platform, 12 Sigma Technologies' σ -Discover/Lung system automatically detects lung nodules as small as .01% of an image, analyzes malignancy with >90% accuracy and provides a decision support tool to radiologists. When optimized on an NVIDIA T4 cluster the system runs up to 18x faster.



A NEW ERA FOR VISUAL EFFECTS

Turing-based Quadro RTX delivers photoreal graphics that creators didn't expect for another 5-10 years.

Quadro RTX GPUs can now accelerate photoreal rendering for large industries that previously only used CPU server farms: film, animation, architecture, product design, and others.

NVIDIA has reinvented computer graphics, again.



A NEW WAVE OF AI BUSINESS APPS

Many companies sponsor televised events to promote their brands, yet ROI can take weeks, or even months, to measure. SAP Brand Impact, powered by NVIDIA deep learning, measures brand attributes in near real-time with superhuman accuracy. **With deep neural networks trained on the NVIDIA DGX-1, and with the TensorRT inference engine, SAP improves performance by 40X, reduces hourly costs by 32X, and delivers immediate, accurate, and auditable results.**



AI IS ON TRACK TO SAFEGUARD RAILWAY INTEGRITY

To maintain the integrity of its 3,232 km of tracks, the Swiss Federal Railways (SBB) runs diagnostics trains to photograph and monitor tracks in real-time. But traditional data processing methods produce false positives/negatives. To remedy this, SBB and CSEM (Swiss Research and Development Center) launched the *Railcheck* project which applies deep learning, powered by the NVIDIA DGX Station, to improve the automatic detection and classification of faults.

 NVIDIA. :: 

 SBB CFF FFS



SMARTER INSPECTION SERVICES

In business, ensuring equipment uptime and meeting safety and regulatory compliance is non-negotiable. Using deep neural networks developed on NVIDIA DGX-1 in the data center that can easily extend to NVIDIA DGX Station in the field, Avitas Systems delivers inspection services using robotic-based autonomous inspection and advanced data analytics. In addition to safeguarding workers, Avitas Systems AI solutions can reduce inspection costs by 25% and reduce maintenance downtime by 15%.



The robots can handle the heat and use infrared cameras and chemical and other sensing technologies to inspect assets under dangerous conditions and keep production running. Image credit: Avitas Systems

PUMPING AI INTO THE OIL & GAS INDUSTRY

NVIDIA and Baker Hughes (BHGE) are using AI and GPU-accelerated computing to help companies distill oceans of data and reduce the cost of finding, extracting, processing and delivering oil. BHGE's AI Factory and NVIDIA's end-to-end AI supercomputing solutions—from the NVIDIA DGX-1 in data centers, to the NVIDIA DGX Station deskside or at remote locations, to NVIDIA Jetson at the edge—can unlock insights from data that was previously as hidden as the oil underground.



REAL-TIME FRAUD DETECTION

Recently, PayPal was looking to deploy a new fraud detection system. The team working on it set a high bar: this system had to operate worldwide 24/7, and work in real-time to protect customer transactions from potential fraud. In spec'ing the system, it became evident that CPU-only servers couldn't meet these requirements.

Using NVIDIA T4 GPUs, PayPal delivered a new level of service, using GPU inference to improve real-time fraud detection by 10% while lowering server capacity by nearly 8x.



LOSS PREVENTION

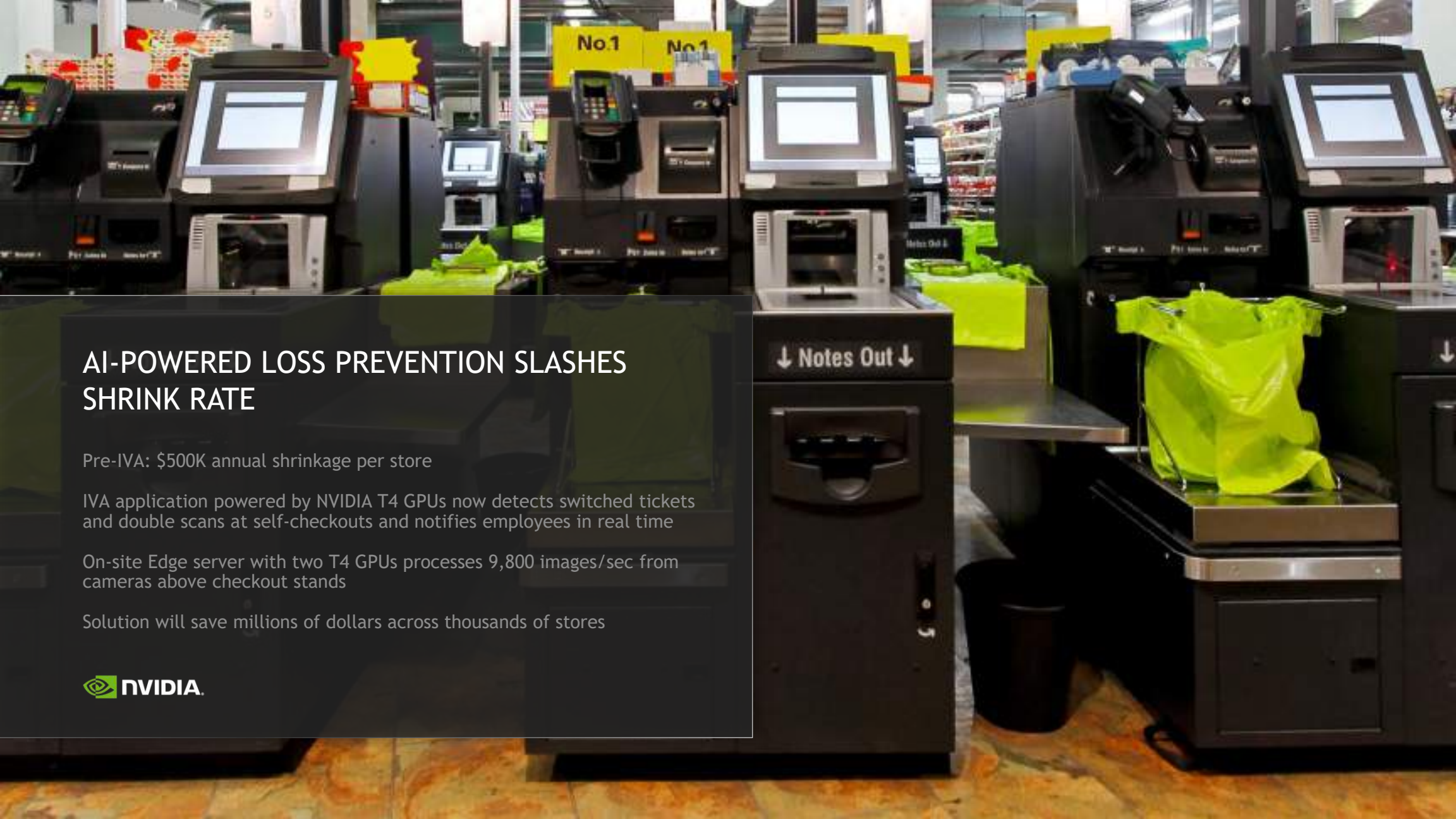
Using existing cameras, a retailer can install highly effective computer vision algorithms to detect shopper traffic patterns and prevent loss.

In the US, Loss Prevention is a \$50B problem impacting all retailers. At the same time, investment in Loss Prevention staff is flat of shrinking.

While the average cost of shoplifting incidents is doubling to \$798, 30% of inventory shrinkage is an inside-job.

Using computer vision can identify theft, shrinkage, and shoplifting incidents. This new technology can invigorate a longstanding problem for retail.





AI-POWERED LOSS PREVENTION SLASHES SHRINK RATE

Pre-IVA: \$500K annual shrinkage per store

IVA application powered by NVIDIA T4 GPUs now detects switched tickets and double scans at self-checkouts and notifies employees in real time

On-site Edge server with two T4 GPUs processes 9,800 images/sec from cameras above checkout stands

Solution will save millions of dollars across thousands of stores



OPTIMIZING QUALITY INSPECTIONS WITH AI

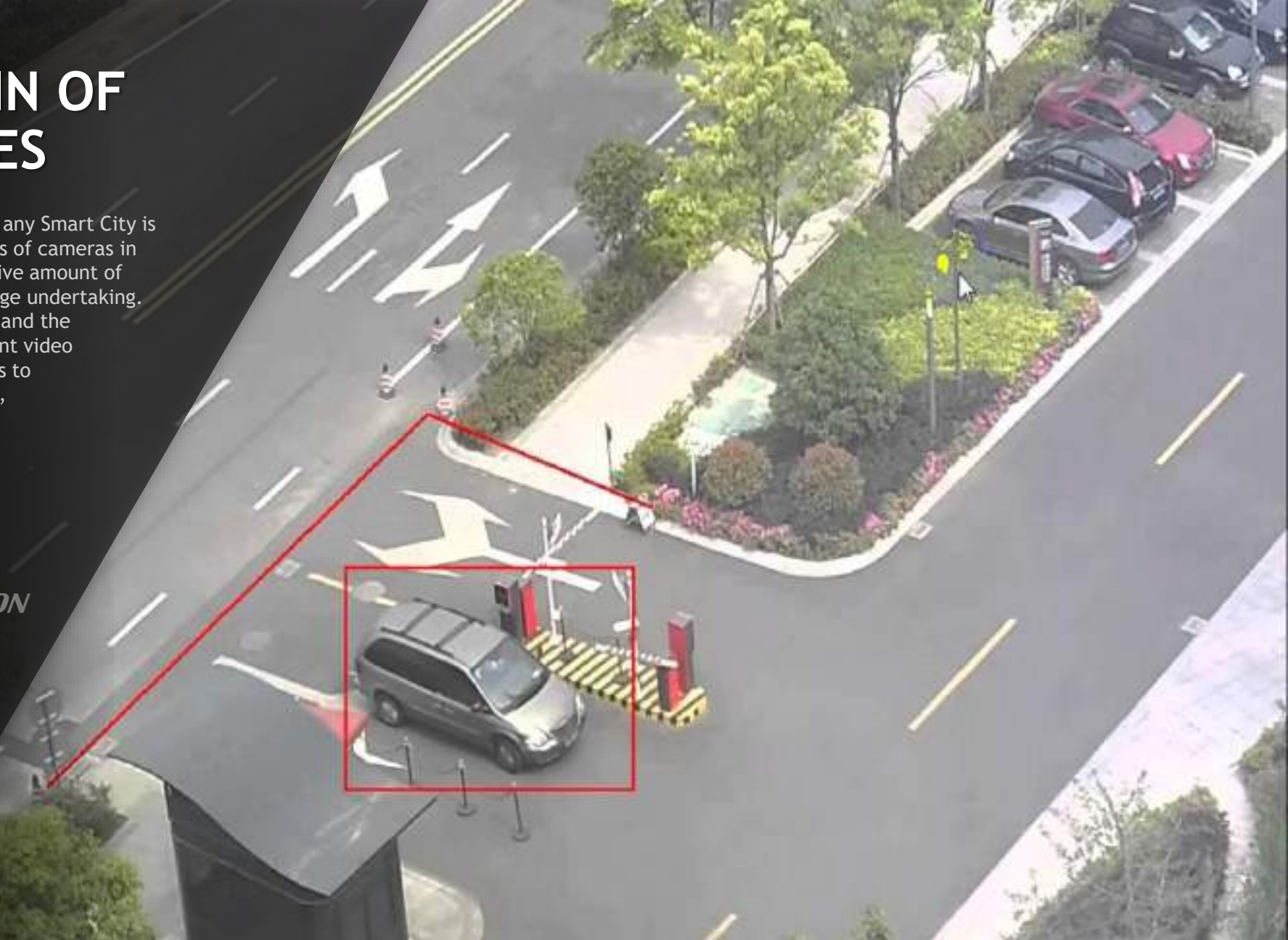
Due to the increasingly sophisticated design of its cars and the high-quality standards at Audi, the company inspects all components —doors, engine hoods, fenders, etc.— in its press shop.

In addition to visual inspection by Audi employees, several small deep learning-based cameras trained on NVIDIA GPUs installed directly in the presses detect the finest cracks in sheet metal with the utmost precision in a matter of seconds.



THE AI BRAIN OF SMART CITIES

A key technology component of any Smart City is video surveillance. With millions of cameras in large cities, analyzing the massive amount of video content generated is a huge undertaking. Hikvision is using deep learning and the NVIDIA DGX-1 to speed intelligent video content analysis for Smart Cities to safeguard citizens and property, manage traffic, and increase criminal investigation efficiencies.





GPU VIRTUALIZATION FOR ANY WORKLOAD

NVIDIA delivers GPU virtualization for both graphics and compute workloads





WORKPLACE TRENDS 2020

91%

Indicated they have recently implemented 'work from home' arrangements

54%

Indicated that poor technology and/or infrastructure for remote working is the biggest barrier to effective remote working



IDC predictions for G2000 companies by 2021

60%

will utilize employee experience (EX) as a key differentiator

60%

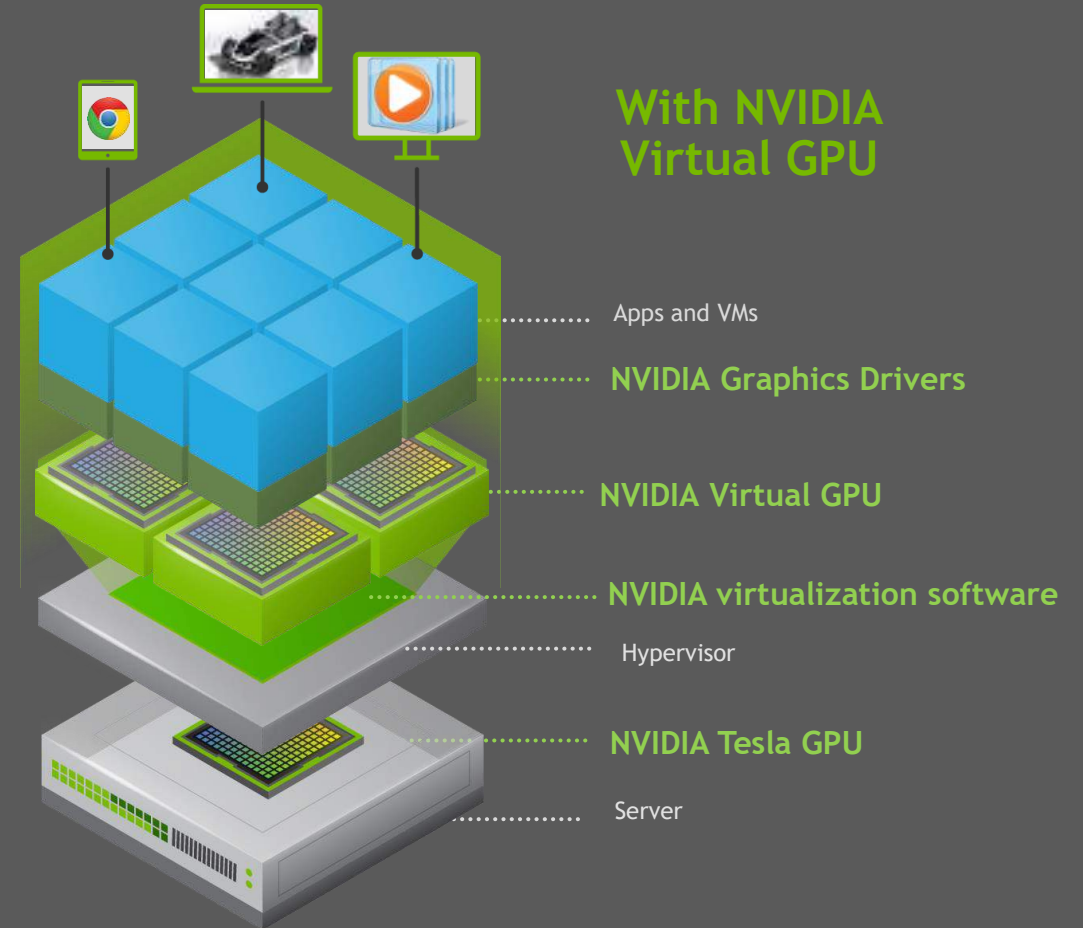
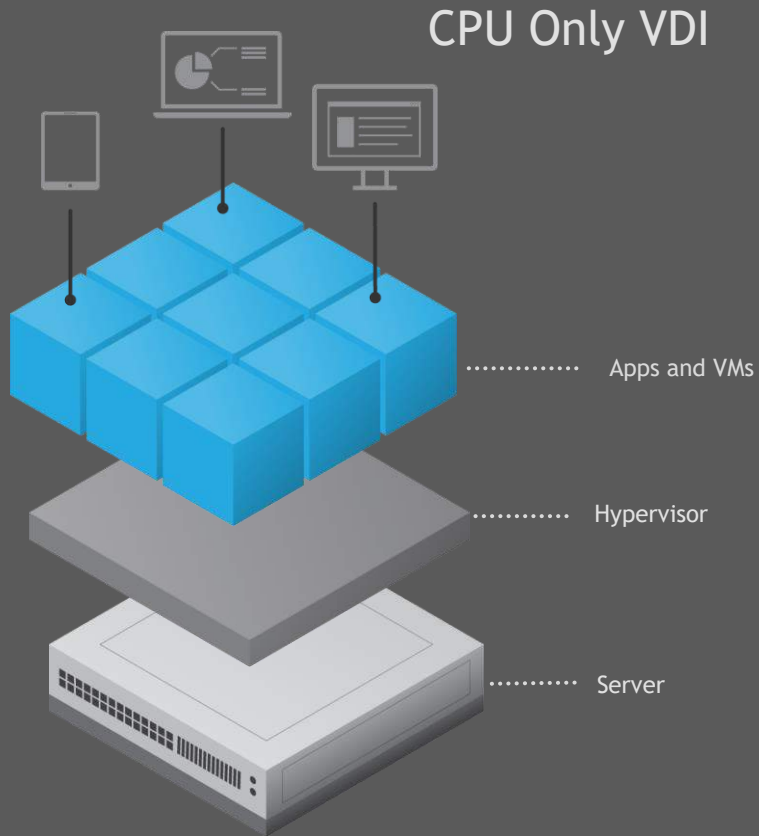
will have adopted a future workspace model

65%

will offer remote work options for accessing enterprise and collaboration apps

NVIDIA VIRTUAL GPU

GPU Accelerated Experience to every Virtual Desktop with NVIDIA HW and SW



ACCELERATE PRODUCTIVITY

For Every User, Any App



Knowledge Worker

Providing business users the highest level of experience for all their apps on any device



Creative & Technical Professional

Uncompromised experience for professional graphics users allowing them to design on the go



Scalability

Performance

IMMERSIVE VIRTUAL WORKSPACES



Knowledge/Business Worker

NVIDIA GRID Virtual PC

NVIDIA GRID Virtual Apps



Creative/Technical Professional

NVIDIA Quadro Virtual
Data Center Workstation

VCS - Virtual Compute Server

QUADRO VIRTUAL WORKSTATIONS

Bringing the Power of Quadro to Professional Workflows Anywhere



AEC



Education



Energy



Financial Services



Government



Healthcare



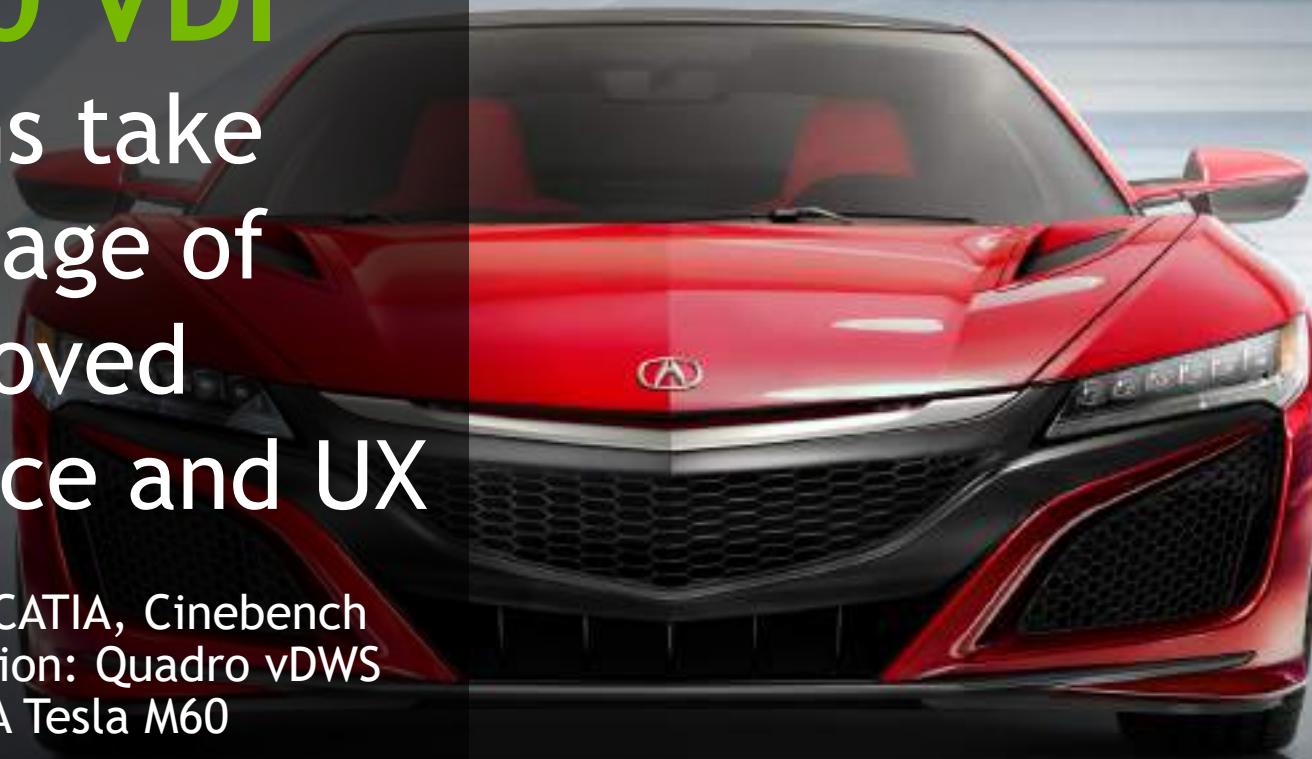
Manufacturing



Media & Entertainment

More than
4,000 VDI
systems take
advantage of
improved
performance and UX

Key Applications: CATIA, Cinebench
Graphics Acceleration: Quadro vDWS
GPU: NVIDIA Tesla M60



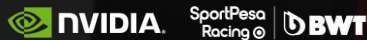
HONDA

WINNING THE RACE TO DIGITAL INNOVATION WITH NVIDIA QUADRO VDWS

SportPesa Racing Point F1 Team, are a group of motor racing specialists who design, manufacture, and compete in the extreme world of Formula 1.

When they are not racing, they are heads down perfecting their racing machines using NVIDIA Quadro vDWS to accelerate CAD applications and stay ahead of the competition.

[READ STORY](#)



VIRTUAL LABS IMPROVE STUDENT EXPERIENCE

Georgia Tech has long required that its students own computers—typically laptops—so that it can better serve their educational needs. However, when it came to accessing the graphics-intensive applications required by many science and engineering courses, the student laptops weren't capable of running these applications due to hardware requirements or licensing restrictions. So students could only use a limited number of computer labs where they were installed. By creating Vlab, students can now access all required applications on their own devices, where ever they are.

**Georgia
Tech**



NVIDIA GPUs EVERYWHERE

140+ Servers from more than 30 system vendors

Hypervisor Infrastructure



Industry Standard Servers



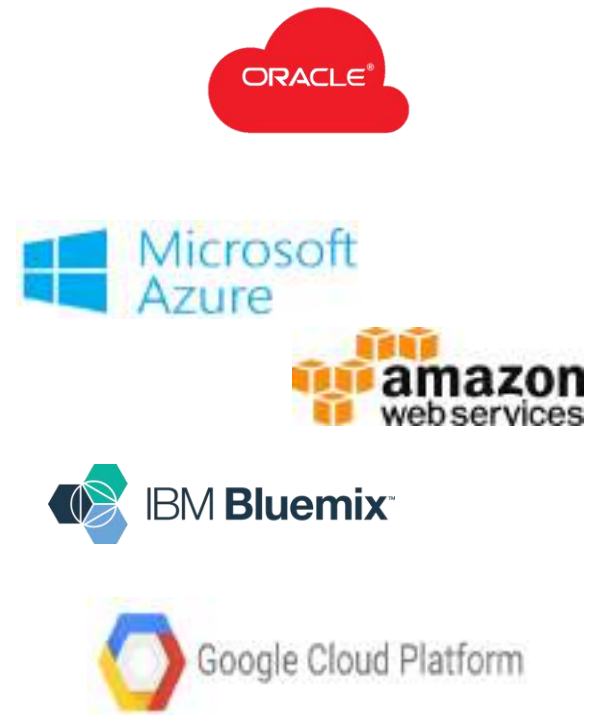
Hyper-Converged Infrastructure



Blade Servers




Public Cloud



NVIDIA PLATFORM

Comprehensive End to End Hardware and Software Stack

HPC			AI			VISUALIZATION		NVIDIA GPU CLOUD 
NVIDIA AI	DRIVE	METROPOLIS	CLARA	ISAAC	RTX	vGPU		
WORKSTATIONS			SERVERS			CLOUD		
CUDA & GPU COMPUTING ARCHITECTURE								



XAVIER



TITAN



QUADRO



TESLA

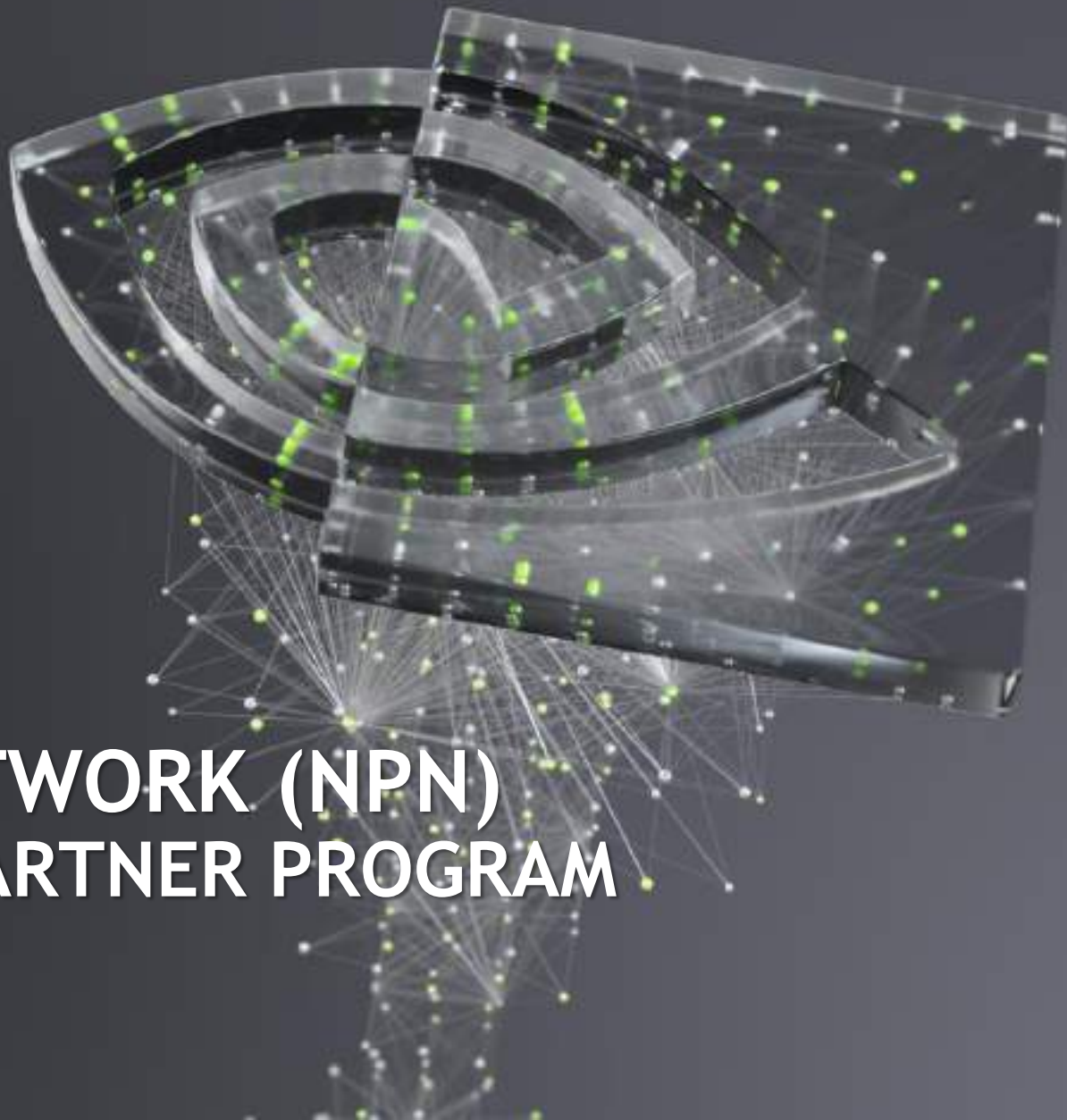


DGX



NVIDIA PARTNER NETWORK (NPN) SOLUTION PROVIDER PARTNER PROGRAM

July 2020

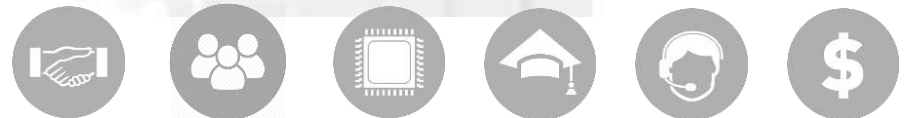


PARTNERING WITH NVIDIA

Accelerating Profitability and Success

NVIDIA PARTNER NETWORK (NPN) PARTNER PROGRAM

NVIDIA's award-winning NPN Partner Program offers a comprehensive portfolio of benefits, incentives and resources, aligned with a partner's level of commitment and deep expertise in NVIDIA GPU computing, to increase profitability, expand into to markets, and empower customer success.





NPN BENEFITS AND REQUIREMENTS

NPN SOLUTION PROVIDER

Program Benefits*

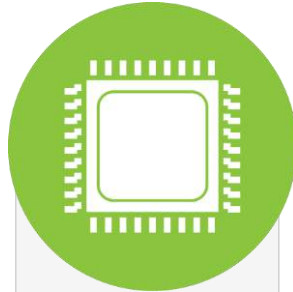


Tiered membership levels**

Partner management team

Business planning & review meetings

Business Partnership



Award-winning technologies, solutions and brands

Easy access to products and technology information

Products & Technology



Access to PBMs and sales support

Incentives and promotions

Opportunity Registration (Compute & Virtualization)

Demo Program (NFR)

Sales



Partner marketing resources

MDF

Marketing tools

Partner Portal

Partner Locator & Where To Buy

Marketing



Self-paced online sales and technical training

Competency curriculums

Industry Vertical Training

Training & Enablement



Partner technical team resources

Online support tools

Priority technical support

Support

* This is a condensed overview of the NPN Program Benefits. A comprehensive list is outlined in the NPN Program Guide.

** New partners will enter the program at the Registered level

FY21 NPN PARTNER* COMPENSATION

Excluding Product Margin

Discounts / Rebates	Elite	Preferred	Registered
Opportunity Registration <ul style="list-style-type: none"> • DGX Products • vGPU Products 	Yes 15% 21%	Yes 15% 21%	Yes 15% 21%
NEW: Fixed Rebate on all POS**	1%**	-	-
Change/New: QPB + New Annual Target**	1%	-	-
QPB Stretch for Above Target Achievement**	2%	-	-
Quarterly Incentives (retained margin)	Yes	Yes	-
Edu & Inception Discount Eligibility	Yes	Yes	Yes
MDF	Yes	Yes	-

* Solution Provider and Solutions Integration Partners must be in good standing and compliant in the NPN Partner Program.

** Maximum Rebate Payout Applies.

NPN SOLUTION PROVIDER

Program Requirements*



Master Agreement & Competency Exhibits

Territory Authorization

Business Planning and Review Meetings

Business Partnership



Competency POS Threshold Attainment

Provide Sales Out Data

Sales



NVIDIA Brand Compliance

Drive Marketing Activities

Host Local Events

Marketing



NVIDIA Sales Professional Training

NVIDIA Technical Professional Training

Training & Enablement



Provide front-line customer support

Drive RMA Process

Support

* This is a condensed overview of the NPN Program Requirements. A comprehensive list is outlined in the NPN Program Guide and Competency Documents.

SP COMPUTE DGX COMPETENCY THRESHOLDS

Curriculum Requirements and Annual Sales Schedule

Compute Competency Program Requirements		Elite	Preferred	Registered
Compute DGX Products	NVIDIA Trained Sales Professionals*	4	2	2
	NVIDIA Trained Technical Professionals*	2	1	1
	DGX Hands-on Installation Training	2	2	2
	Complete DLI Courses	2	N/A	N/A

Compute DGX					
Schedule	Geo	Country	Elite POS USD	Preferred POS USD	Registered POS USD
1	APAC	China	\$2.5M	\$625K	\$0
	Japan	Japan			
	NALA	US			
2	APAC	Korea	\$1.5M	\$375K	\$0
	EMEA	Germany			
3	APAC	Hong Kong, India, Singapore, Taiwan	\$1M	\$250K	\$0
	EMEA	Belgium, France, Sweden, UK, Switzerland			
	NALA	Canada			
4	Rest of World	All Other Countries not listed above	\$500K	\$125K	\$0

*For partners to be training compliant, Elite partners are required to have 6 (six) unique individuals and Preferred and Registered partners are required to have 3 (three) unique individuals complete the training curriculum.



PROGRAMS AND ENABLEMENT

NPN TRAINING & ENABLEMENT

SP Competencies and Curriculums - Time Investment

Time to complete curriculums (hours : minutes)*

COMPENTENCIES	CURRICULUM(S)	SALES PROFESSIONAL CURRICULUM*	TECHNICAL PROFESSIONAL CURRICULUM*
Compute	Compute	5:22	6:49
Compute DGX**	Compute + Compute DGX	5:22 2:53	6:49 3:01
Visualization	Visualization	2:15	5:05
Virtualization	Virtualization	1:50	3:52



www.nvidia.com/npnportal

* Requirements vary by program type, curriculum and level within the program.

* Unique individuals are required to complete each curriculum.

** Compute DGX 2020 Curriculum is required. Compute 2020 Curriculum is suggested, but not required. Compute 2021 will be required when it launches in the fall of this year.

INDUSTRY ENABLEMENT

Training for Knowledge, Resources for Guidance & Engagement

Content and concise training designed to help you identify target customers, initiate conversations and qualify sales opportunities for a range of industry focused use cases.

Industry Solutions highlighted on the NPN Partner Portal:

- Healthcare
- Higher Education and Research
- Financial Services and Insurance
- Manufacturing and Industrial
- Media and Entertainment
- Energy/Oil and Gas
- Retail
- Additional resources in progress



www.nvidia.com/npnportal - GPU Genius Training

OPPORTUNITY REGISTRATION

Overview*

- ▶ **Eligibility:** NPN Solution Providers, meeting competency requirements and in good standing

PRODUCT	DISCOUNT	MINIMUM
DGX	15%	No minimum
vGPU software	21%	\$10K USD (list price)
DLI Workshop, 1-day (\$10k SKU only)	10%	No minimum

- ▶ **Products:**

- **Key Criteria:**

- Must be a new opportunity, not previously registered. Can be a new customer, new department or a new solution at existing location
- Partner presales efforts must be in place, min deal size met within auth. territory

- ▶ **Timing:**

- ▶ Opportunity Registrations are open for 90 days
 - ▶ One Extension allowed for additional 90 days (180 days total), partner must be highly active
 - ▶ Do not let the Opp expire or terminate! They cannot be resurrected.

- ▶ **Important:**

- ▶ Be sure to respond to system emails prompting next steps and deadlines
- ▶ Stay engaged with partners

DEMO PROGRAM OVERVIEW

Not for Resale (NFR) Product Discounts to Drive Revenue



Continuous Innovation

NPN Demo Program provides partners with discounted pricing on specific hardware and software to support the sales process. NFR demo products are intended for a partner's in-house development, customer demonstration or proof-of-concept.

Program eligibility is based on the program type and level of membership.



PARTNER RESOURCES

NPN PARTNER PORTAL

Online access to a full suite of enablement tools & resources to accelerate sales success

Sales Tools

- Product datasheets & whitepapers
- Customer Presentations
- Demo Program (NFR)

Marketing Resources

- Product Launch Kits
- Marketing Campaign Tools & Resources
- Webinars, News and Events

GPU Genius Training

- Comprehensive Course Catalog
- Competency Curriculums
- Industry Vertical Resources

Support

- Technical Support Tools
- Knowledge Database
- Partner Support Contacts



www.nvidia.com/npnportal

NPN Partner Portal* - Your “One-Stop” Shop

* NPN Partner Portal provides role-based permissions.

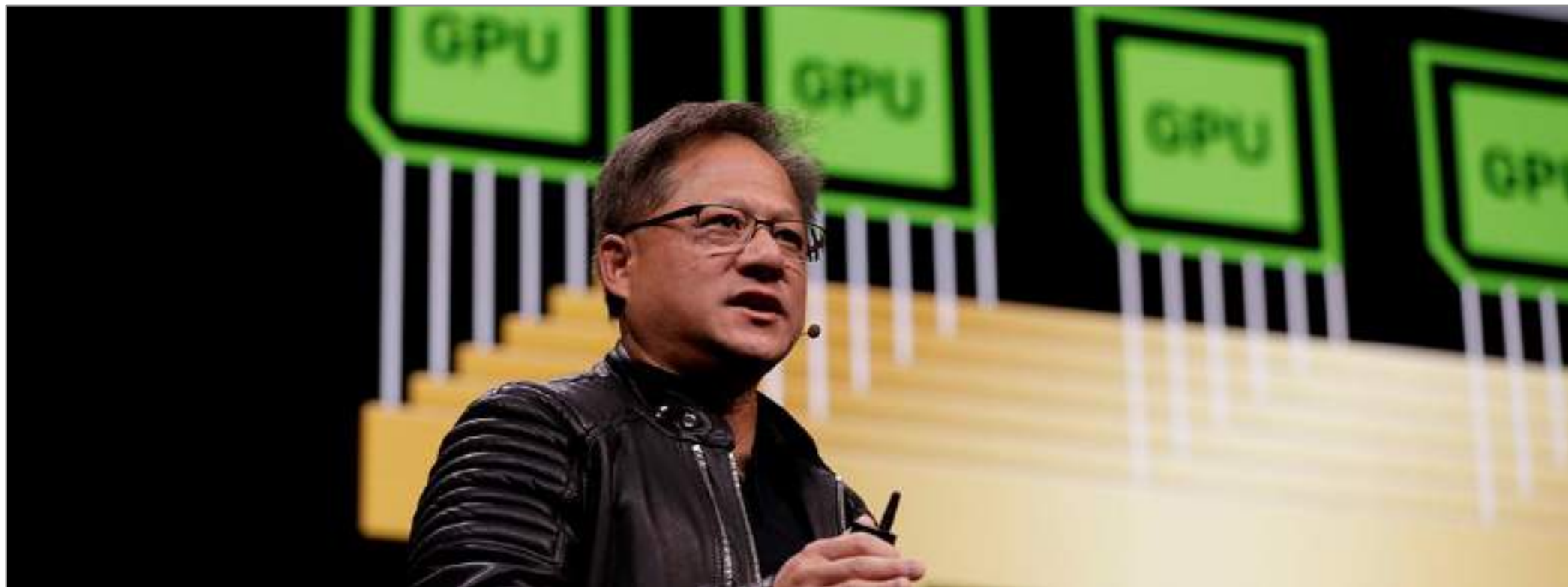


DEEP
LEARNING
INSTITUTE

Training developers, data scientists, researchers and IT professionals how to solve their most challenging problems

GPU TECHNOLOGY CONFERENCE

GTC is the premier AI Conference providing training, insights, and direct access to experts on the hottest topics in computing today.





NVIDIA



[/NVIDIA.LA](#)



[/NVIDIALA](#)



[/nvidia-latinoamerica](#)

Sigue NVIDIA