NVIDIA ACCELERATED COMPUTING GROWTH

50% GROWTH OF NVIDIA DEVELOPERS

800K DEVELOPERS
2018 2019
+50%

1.2M DEVELOPERS

8M CUDA DOWNLOADS
2018 2019
+40%

50% GROWTH IN TOP500

#1 World, US — ORNL Summit
#1 Europe — CSCS PizDaint
#1 Japan — AIST ABGI
22 of Top 25 Energy-Efficient

600+ CUDA APPS

CRYPTOMAC: Cryo
MICROVOLITION: Nanoscale
PARABRICKS: Materials
GRIMM: Chemistry
GRIMM: Biology
FUSION: Geosciences
CHIMERA: Neuroimaging
WRF: Weather

MORE PERF

25X 40X
2018 2019
AMBER CHARMM LAMMPS MILC NAMD Quantum ESPRESSO S3ID
NVIDIA CUDA-X
GPU-ACCELERATED COMPUTING LIBRARIES
NVIDIA CUDA-X
GPU-ACCELERATED COMPUTING LIBRARIES
CHAPTER 1: GRAPHICS
BROAD SUPPORT FOR NVIDIA RTX
RTX COMES TO 9M 3D CREATORS IN 2019

1 MILLION
Architects

3 MILLION
Designers

3 MILLION
3D Artists

2 MILLION
M&E Pros
<table>
<thead>
<tr>
<th># of Nodes</th>
<th>Total Render Time</th>
<th>Cost of Power [5 yrs]</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU - Dual Skylake</td>
<td>25</td>
<td>38 Hours</td>
<td>$70k</td>
</tr>
<tr>
<td>RTX Server - 4x RTX 8000</td>
<td>1</td>
<td>6 Hours</td>
<td>$10k</td>
</tr>
</tbody>
</table>

*Pixar artists already rely on NVIDIA ray tracing, and RTX more than doubles the performance they will see. We’re excited to use RTX on our upcoming films.*

— STEVE MAY, CTO, Pixar
ANNONCING NVIDIA OMNIVERSE
OPEN COLLABORATION PLATFORM TO SIMPLIFY STUDIO WORKFLOWS FOR REAL-TIME GRAPHICS
ANNOUNCING RTX SERVER
Datacenter Graphics Server Design

40 Turing GPUs in 8U
Virtualize graphics apps up to 320 CCU
Optimized end-to-end stack for rendering, remote workstation, and cloud gaming
ANNOUNCING RTX SERVER POD
Modular Designs for Enterprise & Cloud Edge Datacenters

Pods scale to 32 RTX servers
1,280 GPUs in 10 racks
High-speed storage connected with MLNX IB
Up to 10,000 concurrent users per RTX Pod
CHAPTER 2: AI & HPC
DATA SCIENCE – A NEW PILLAR OF DISCOVERY

DATA → DATA ANALYTICS → FEATURES → NLU, CV, ML, DL → PREDICTIVE MODEL → INFERENCE → PREDICTION
ETL, Pandas, Spark, Graph, TensorFlow, PyTorch, MXNet, Scikit-Learn, XGBoost, TensorFlow Serving, ONNX, SageMaker NEO, CSV, PARQ, HDFS
DATA SCIENCE – A NEW PILLAR OF DISCOVERY

DATA → DATA ANALYTICS

FEATURES → PREDICTIVE MODEL

AI → NLU, CV

ML → DL

PREDICTION → INFEERENCE

cuIO, cuDF, cuGraph → cuDNN, cuML

TensorRT, TRTIS
NVIDIA CUDA-X AI ECOSYSTEM

FrameWorks
- Chainer
- Caffe
- PyTorch
- TensorFlow

Cloud ML Services
- Accenture
- AWS SageMaker
- Baidu Brain
- Deloitte
- Intel
- IBM
- Microsoft
- NVIDIA
- NVIDIA Tensorflow
- ONNX
- Servo

Deployment
- NVIDIA Tensorflow Serving

CUDA-X-AI
- DA
- Graph
- ML
- DL Train
- DL Inference

CUDA

Announcing Tensor Core Mixed-Precision Automatically Supported in TensorFlow, PyTorch, MXNet

<table>
<thead>
<tr>
<th>Workstation</th>
<th>Server</th>
<th>Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dell, HP, Lenovo</td>
<td>NVIDIA, Google, AWS, Microsoft, IBM, Intel</td>
<td>AWS, Google Cloud, Azure, Alibaba Cloud, Tencent Cloud</td>
</tr>
</tbody>
</table>
NVIDIA CUDA-X AI ECOSYSTEM

FRAMEWORKS
- PyTorch
- TensorFlow
- Keras
- PyTorch
- TensorFlow

CLOUD ML SERVICES
- AWS SageMaker
- Google Cloud AI
- Microsoft Azure Machine Learning

DEPLOYMENT
- ONNX
- TensorFlow Serving

CUDA-X-AI
- DA
- GRAPH
- ML
- DL TRAIN
- DL INFERENCE

CUDA
- Workstation
- Server
- Cloud

Announcing RAPIDS with GPU VM images and Kubeflow

Google Cloud
We are excited to announce that we are partnering with NVIDIA to bring RAPIDS to Accenture’s new Applied Intelligence Platform enabling our joint venture with Kansas Electric Power, K4Digital, to transform business process and create new opportunities.

– MIKE SUTCLIFF, Group Chief Executive, Accenture Digital

accenture
NVIDIA CUDA-X AI ECOSYSTEM

FRAMEWORKS
- Chainer
- Caffe
- PyTorch
- TensorFlow

CLOUD ML SERVICES
- Accenture
- AWS
- Google Cloud
- Microsoft
- NVIDIA

DEPLOYMENT
- Tensorflow Serving

CUDA-X AI

CUDA
- Workstation
- Server
- Cloud
What are different types of lighting for a living room?

Ambient, task and accent
ANNOUNCING WORLD’S LEADING TECH COMPANIES ADOPT CUDA-X AI TO ACCELERATE MODEL DEPLOYMENT

6X TensorRT Downloads

- 50K
- 300K

2017
2018

Voice Search
Image Search
Recommendations
Home Assistant
News Feed
Translation
eCommerce
ANNOUNCING CLARA AI TOOLKIT

BUILD, MANAGE AND DEPLOY AI APPLICATIONS FOR RADIOLOGY

- AI-Assisted Annotation – Hours to Minutes
- 10x Less Training Data Needed
- 13 Pre-Trained Models
- Reference Training and Deployment Pipelines

Available at developer.nvidia.com/clara
CLARA AI – BUILT FOR RADIOLOGY

MGH
3D Annotation from Hours to Minutes

NIH
Trained State-of-the-Art Model in Weeks

OSU
Clinical Model Deployment in Less Than 24 hours

DKFZ
Clara AI Annotation Integrated in MITK Open Source Viewer
DATA SCIENCE IS THE NEW HPC CHALLENGE
Charter Spectrum
25 million customers in 40+ states in US

500K access points
1TB | day
A WORKSTATION FOR DATA SCIENTISTS

POWERED BY NVIDIA GPU AND CUDA-X AI

Dual Quadro RTX 8000 with 96 GB Memory
Pre-installed for CUDA-X Accelerated Data Science — RAPIDS, TensorFlow, PyTorch, Caffe, Anaconda Distribution 10X Faster
A WORKSTATION FOR DATA SCIENTISTS

POWERED BY NVIDIA GPU AND CUDA-X AI

Dual Quadro RTX 8000 with 96 GB Memory
Pre-installed for CUDA-X Accelerated Data Science —
RAPIDS, TensorFlow, PyTorch, Caffe, Anaconda Distribution
10X Faster
ANNOUNCING WORLD’S TOP COMPUTER MAKERS OFFER WORKSTATIONS OPTIMIZED FOR DATA SCIENCE

POWERED BY NVIDIA GPU AND CUDA-X AI

Dual Quadro RTX 8000 with 96 GB Memory
Pre-installed for CUDA-X Accelerated Data Science — RAPIDS, TensorFlow, PyTorch, Caffe, Anaconda Distribution
10X Faster
SUPERCOMPUTER vs. HYPERSCALE

Supercomputer | Capability Machine | Scale-up Architecture

![Supercomputer Diagram](image1)

![Hyperscale Diagram](image2)

- **SUPERCOMPUTER**
  - Memory (PB): 1T PF, 1B PF
  - CPU: 1T PF, 1B PF

- **HYPERSCALE**
  - Memory (PB): 1K, 1M, 1B
  - CPU: 1K, 1M, 1B
DATA SCIENCE – THE NEW HPC CHALLENGE

NVIDIA DGX-2
AI Supercomputer Appliance
16x V100 | 2 PF | 512GB HBM2
8x MLNX IB
DATA SCIENCE – THE NEW HPC CHALLENGE

NVIDIA DGX-2
AI Supercomputer Appliance
16x V100 | 2 PF | 512GB HBM2
8x MLNX IB
DATA SCIENCE – THE NEW HPC CHALLENGE

**NVIDIA DGX-2**
AI Supercomputer Appliance
16x V100  |  2 PF  |  512GB HBM2
8x MLNX IB

**Data Science Server**
4x T4  |  260 TF FP16  |  64GB GDDR6
MLNX or BRCM EN
DATACENTER BECOMES A COMPUTE ENGINE
NVIDIA DGX POD — ZERO TO DATA SCIENCE IN 1 DAY
ANNOUNCING WORLD'S MOST POPULAR ENTERPRISE SERVERS OPTIMIZED FOR DATA SCIENCE

Powered by NVIDIA T4 and New CUDA-X AI
NGC Certified
ANNOUNCING WORLD’S MOST POPULAR ENTERPRISE SERVERS OPTIMIZED FOR DATA SCIENCE

Powered by NVIDIA T4 and New CUDA-X AI
NGC Certified
### Acceleration of Data Science Clusters

<table>
<thead>
<tr>
<th>Activity</th>
<th>10X CPU Nodes</th>
<th>10X T4 Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>End to End</td>
<td>35</td>
<td>3</td>
</tr>
<tr>
<td>Training</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>Data Prep</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

**Minutes (smaller is better)**

### Deep Learning Scaling

- **Images per Second (ResNet-50)**
- **Number of Servers**

The graph shows the scaling of deep learning processing, indicating an increase in images per second as the number of servers increases.
THE WORLD’S LEADING BRANDS USE AMAZON EC2 P3 INSTANCES WITH NVIDIA V100 GPUS
INTRODUCING AMAZON EC2 G4 INSTANCES WITH NVIDIA T4 TENSOR CORE GPUs

AWS

MACHINE LEARNING INFERENCE
GRAPHICS INTENSIVE APPLICATIONS
VIDEO TRANSCODING
CHAPTER 3: ROBOTICS
JETSON POWERING AUTONOMOUS MACHINES

WAREHOUSE

DELIVERY

AGRICULTURE

RETAIL

INDUSTRIAL
ANNOUNCING JETSON NANO

$99 NVIDIA CUDA-X AI COMPUTER
CUDA-X acceleration stack
High-resolution sensor support
Runs all CUDA-X AI models
ANNOUNCING ISAAC OPEN SDK

Isaac Robot Engine - Modular robot framework
Isaac Sim - Virtual robotics laboratory
Isaac Gym – Reinforcement learning simulator
Isaac Robot Apps - Kay, Carter and Link

Available at developer.nvidia.com/isaac-sdk
THE DRIVE INITIATIVE

DGX Saturn V

Constellation

Xavier

DRIVE AV

DRIVE IX

KITT Resim

Image of a car interior with a display showing a person's face.
DRIVE OPEN PLATFORM ECOSYSTEM

CARS
- TOYOTA
- Mercedes Benz
- VOLVO
- NIO
- ISUZU

TRUCKS
- UBER
- N搬

MOBILITY SERVICES
- BOSCH
- ZF
- NVIDIA
- Nuro
- Metrika

TIER 1s
- Continental
- Valeo
- Aptiv

AI ISVs
- SoundHound Inc.
- Waymo
- Zoox
- Zoox

MAPPING
- Cogniota
- Sony
- Velodyne

SIMULATION
- NVIDIA

SENSORS
- FLIR
ANNOUNCING DRIVE AP2X RELEASE 9.0

HIGH FUNCTION L2+ AUTOPilot SYSTEM
- On-ramp to off-ramp
- Surround perception for auto lane-change
- Localization to all major HD maps
- Real-time mapping with camera and radar
- Integrated AR and VR confidence visualization
- Integrated UI and voice command
ANNOUNCING DRIVE AV — PATH PLANNING WITH SAFETY FORCE FIELD

Path Planning + Prediction + Safety Force Field
Designed for AV safety and drive comfort
Mathematically verifiable, validated in simulation
Open platform

PATH PERCEPTION & PLANNING
ANNOUNCING DRIVE AV — PATH PLANNING WITH SAFETY FORCE FIELD

Path Planning + Prediction + Safety Force Field
Designed for AV safety and drive comfort
Mathematically verifiable, validated in simulation
Open platform
Path Planning + Prediction + Safety Force Field
Designed for AV safety and drive comfort
Mathematically verifiable, validated in simulation
Open platform

ANNOUNCING DRIVE AV — PATH PLANNING WITH SAFETY FORCE FIELD

PATH PERCEPTION & PLANNING

Emergency Braking

SAFETY FORCE FIELD

PREDICTION
ANNOUNCING DRIVE AV — PATH PLANNING WITH SAFETY FORCE FIELD

Path Planning + Prediction + Safety Force Field
Designed for AV safety and drive comfort
Mathematically verifiable, validated in simulation
Open platform

PATH PERCEPTION & PLANNING

Emergency Braking

Intersection Handling

SAFETY FORCE FIELD
ANNOUNCING DRIVE AV — PATH PLANNING WITH SAFETY FORCE FIELD

Path Planning + Prediction + Safety Force Field
Designed for AV safety and drive comfort
Mathematically verifiable, validated in simulation
Open platform

PATH PERCEPTION & PLANNING

Emergency Braking

Intelligent Steering

Intersection Handling

SAFETY FORCE FIELD
ANNOUNCING DRIVE AV — PATH PLANNING WITH SAFETY FORCE FIELD

Path Planning + Prediction + Safety Force Field
Designed for AV safety and drive comfort
Mathematically verifiable, validated in simulation
Open platform

PATH PERCEPTION & PLANNING

Emergency Braking
Intelligent Steering
Intersection Handling
Congested Traffic

SAFETY FORCE FIELD
ANNOUNCING DRIVE CONSTELLATION
AVAILABLE NOW

Virtual AV Test Fleet

Bit-accurate, hardware-in-the-loop simulator
Test corner and rare conditions
Simulate previous failure scenarios
Cloud-based workflow
Open platform
ANNOUNCING DRIVE CONSTELLATION
AVAILABLE NOW
Virtual AV Test Fleet

Bit-accurate, hardware-in-the-loop simulator
Test corner and rare conditions
Simulate previous failure scenarios
Cloud-based workflow
Open platform
ANNOUNCING DRIVE CONSTELLATION
AVAILABLE NOW

Virtual AV Test Fleet

Bit-accurate, hardware-in-the-loop simulator
Test corner and rare conditions
Simulate previous failure scenarios
Cloud-based workflow
Open platform
ANNOUNCING NVIDIA AND TRI-AD PARTNER TO CREATE FUTURE OF AUTONOMOUS VEHICLES
ACCELERATED COMPUTING – THE PATH FORWARD “PRADA”
RTX | OMNIVERSE | 5G GFN ALLIANCES