NVIDIA ACCELERATED COMPUTING GROWTH

50% GROWTH OF NVIDIA DEVELOPERS

- 800K developers in 2019 (1.2M developers 20% growth compared to 2018)
- 8M CUDA downloads in 2019 (13M CUDA downloads 40% growth compared to 2018)

50% GROWTH IN TOP500

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

600+ CUDA APPS

- Cryptomining
- Frenetic
- DHF
- MitoMol
- Microscopy
- PARABRICS
- WRF
- Quantum Espresso

MORE PERF

- 25X AMBER
- 40X CORONA, LAMMPS, MILC
- Quantum Espresso, SF3D

Timeline Graphs:

- Y-axis: Year (2010-2018)
- X-axis: Growth %

- Growth rates for CUDA developers and downloads.
NVIDIA CUDA-X
GPU-ACCELERATED COMPUTING LIBRARIES
NVIDIA CUDA-X
GPU-ACCELERATED COMPUTING LIBRARIES

Programmable Domains
SD, TTS, GID, TTA, SOP, COI, COD, COI, GIA, GID, COI

RTX, HPC, AI, DR, IS, CL, ME

CUDA-X

CUDA

RTX, DGX, HGX, AGX
NVIDIA CUDA-X
GPU-ACCELERATED COMPUTING LIBRARIES

PRogrammable
Acceleration
Domains
Architecture
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
COMPLEX RENDERING PIPELINE

CONCEPT | MODELING | TEXTURE | RIGGING
ANIMATION | LIGHTING | RENDER
ANNOUNCING 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 → AI, NLU, CV, ML, DL

PREDICTIVE MODEL → INFEERENCE

→ PREDICTION
DATA SCIENCE – A NEW PILLAR OF DISCOVERY

DATA
CSV, PARQ, HDFS

DATA ANALYTICS
ETL, Pandas, Spark, Graph

FEATURES
NLU, CV, ML, DL

PREDICTIVE MODEL
TensorFlow, PyTorch, ONNX, Scikit-Learn, XGBoost

INFERANCE
TensorFlow Serving, DNN, SageMaker NEO

PREDICTION
DATA SCIENCE – A NEW PILLAR OF DISCOVERY

DATA ANALYTICS
- cuIO
- cuDF
- cuGraph

FEATURES
- AI
- NLU
- CV
- ML
- DL

PREDICTIVE MODEL
- cuDNN
- cuML

INFERENC
- TensorRT
- TRTIS

PREDICTION
NVIDIA CUDA-X AI ECOSYSTEM

FRAMWORKS

CLOUD ML SERVICES

DEPLOYMENT

CUDA-X.AI

CUDA

Workstation

Server

Cloud

Announcing Tensor Core Mixed-Precision
Automatically Supported in TensorFlow, PyTorch, MXNet
NVIDIA CUDA-X AI ECOSYSTEM

FRAMESWORKS
- Chainer
- MXNet
- PyTorch
- TensorFlow
- Caffe

CLOUD ML SERVICES
- Amazon SageMaker
- Azure Machine Learning
- Google Cloud ML
- IBM Watson

DEPLOYMENT
- ONNX
- TensorRT
- Kubernetes
- Argo

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

CUDA

Workstation
- Dell
- HP
- Lenovo

Server
- Intel
- AMD
- ARM

Cloud
- Amazon
- Google
- Microsoft

Announcing RAPIDS Now Integrated with Databricks Unified Analytics Platform
NVIDIA CUDA-X AI ECOSYSTEM

Serving Amazon SageMaker Neo

Announcing RAPIDS with GCP VM images and Kubeflow

CUDA-X-AI

CUDA

DEPLOYMENT

FRAMEWORKS

CLOUD ML SERVICES

DEPLOYMENT

ANNEX

CUDA

Workstation

Server

Cloud

Dell

HP

Lenovo

Graphcore

Inception

Microsoft

NVIDIA

Oracle

Google Cloud

IBM

Amazon

NVIDIA

Neo

TensorFlow

ML

DL TRAIN

DL INFERENCE

Graph

Workstation

Server

Cloud

ANNEX

CUDA

Google Cloud

IBM

Amazon

NVIDIA

Neo

TensorFlow

DL TRAIN

DL INFERENCE

ANNEX

CUDA

Dell

HP

Lenovo

Graphcore

Inception

Microsoft

NVIDIA

Oracle

Google Cloud

IBM

Amazon

NVIDIA

Neo

TensorFlow

ANNEX

CUDA

Graphcore

Inception

Microsoft

NVIDIA

Oracle

Google Cloud

IBM

Amazon

NVIDIA

Neo

TensorFlow
NVIDIA CUDA-X AI ECOSYSTEM

FRAMEWORKS
- TensorRT
- PyTorch
- TensorFlow

CLOUD ML SERVICES
- Amazon SageMaker
- Microsoft Azure Machine Learning
- Google Cloud ML

DEPLOYMENT
- ONNX
- RAPIDS
- TensorFlow Serving

Microsoft

Announcing RAPIDS integrated into Azure Machine Learning Service
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
Announcing NVIDIA TensorRT Integrated into Microsoft ONNX Runtime
What are different types of lighting for a living room?

- Ambient
- Task
- Accent
ANNOUNCING WORLD'S LEADING TECH COMPANIES ADOPT CUDA-X AI TO ACCELERATE MODEL DEPLOYMENT

6X TensorRT Downloads

- 50K
- 300K

Year
- 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 vs. HYPERSCALE

Supercomputer | Capability Machine | Scale-up Architecture

Hyperscale | Capacity Machine | Scale-out Architecture

Graph showing comparison between supercomputer and hyperscale in terms of processing power or scale-up/scale-out architecture.
DATA SCIENCE – THE NEW HPC CHALLENGE
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 6DDR6
MLNX or BRCM EN
DATA CENTER BECOMES A COMPUTE ENGINE
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

Cisco  Dell EMC  Fujitsu  Hewlett Packard Enterprise

inspur  Lenovo  Sugon
### Acceleration of Data Science Clusters

<table>
<thead>
<tr>
<th>Task</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>

### Deep Learning Scaling

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

- 10X CPU Nodes curve starts lower and grows faster than 10X T4 Nodes curve.
MATT GARMAN
Vice President of Compute Services
Amazon Web Services
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

MACHINE LEARNING INFERENCE
GRAPHICS INTENSIVE APPLICATIONS
VIDEO TRANSCODING
CHAPTER 3: ROBOTICS
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 - Kaya, Carter and Link

Available at developer.nvidia.com/isaac-sdk
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
DRIVE AV — PERCEPTION

Objects

Free Space

Lanes & Road Markings

Distance
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
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
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

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
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
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”