DEVNET Create

Design For Scale-
GPU Based Infrastructure for Autonomous Vehicles

Manish Harsh | Global Developer Relations,

Enterprise Automotive, NVIDIA

AGENDA

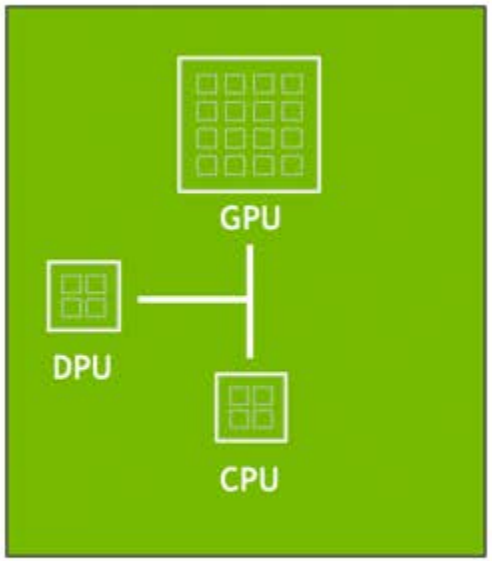
Autonomous Vehicle Development

- GPU Platform Overview
- Fundamentals of AV Stack Architecture
- AV Workflow and Platform: Training and Inference
- Architecture to Scale
- NVIDIA DRIVE – End to End Infra for AV development

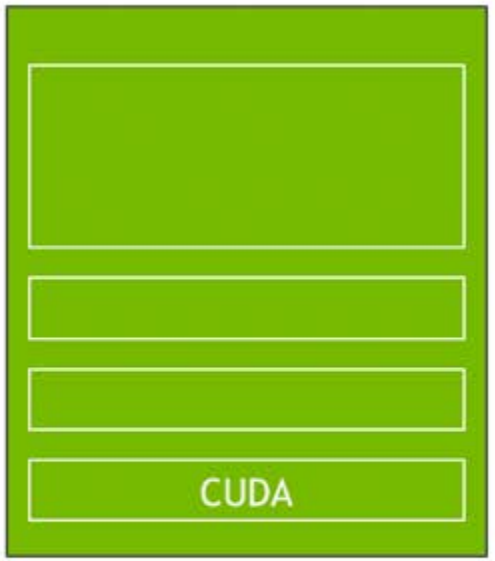
Conclusion

- Key Pointers
- Developer Engagement Platforms
- NGC – NVIDIA Software Hub

FUNDAMENTALS OF NVIDIA PLATFORM

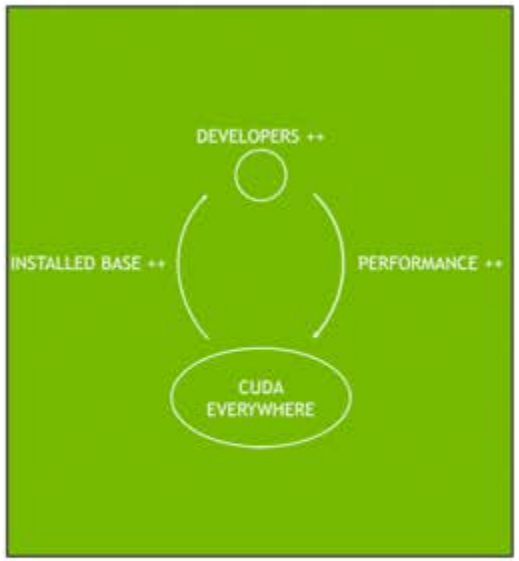


X-FACTOR SPEED UP



FULL STACK

DATA-CENTER SCALE



ONE ARCHITECTURE

AV Development Workflows

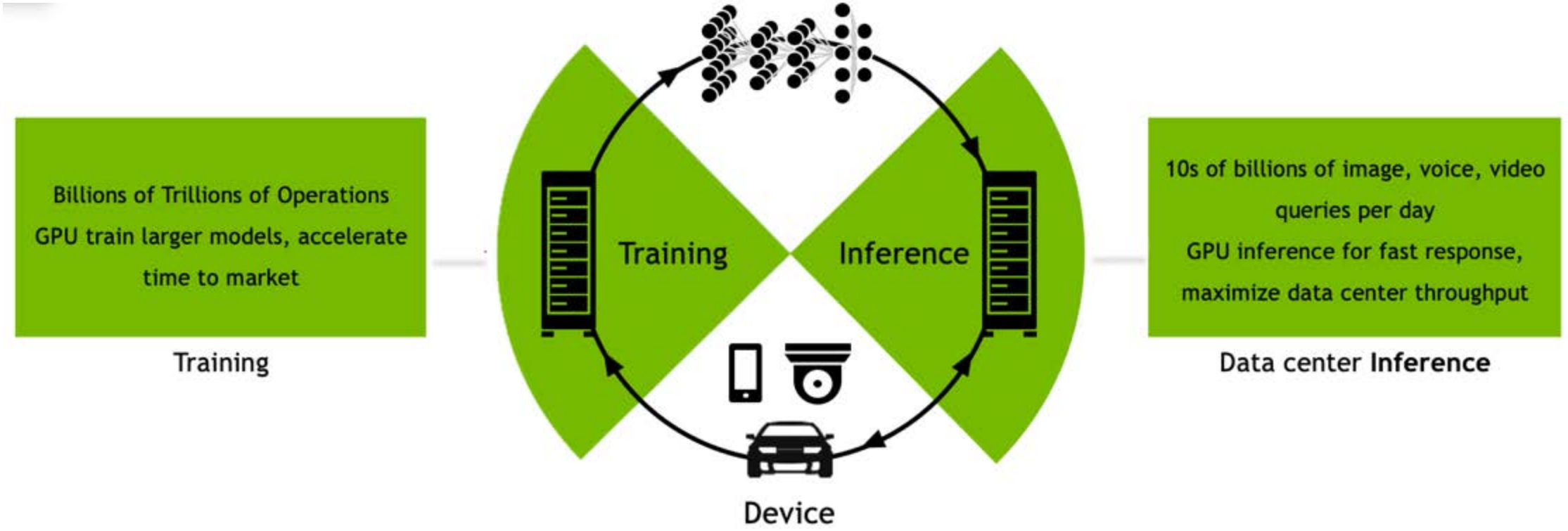
Auto AV Training and Inference

NGC FOR PERFORMANCE

Architecture to Scale

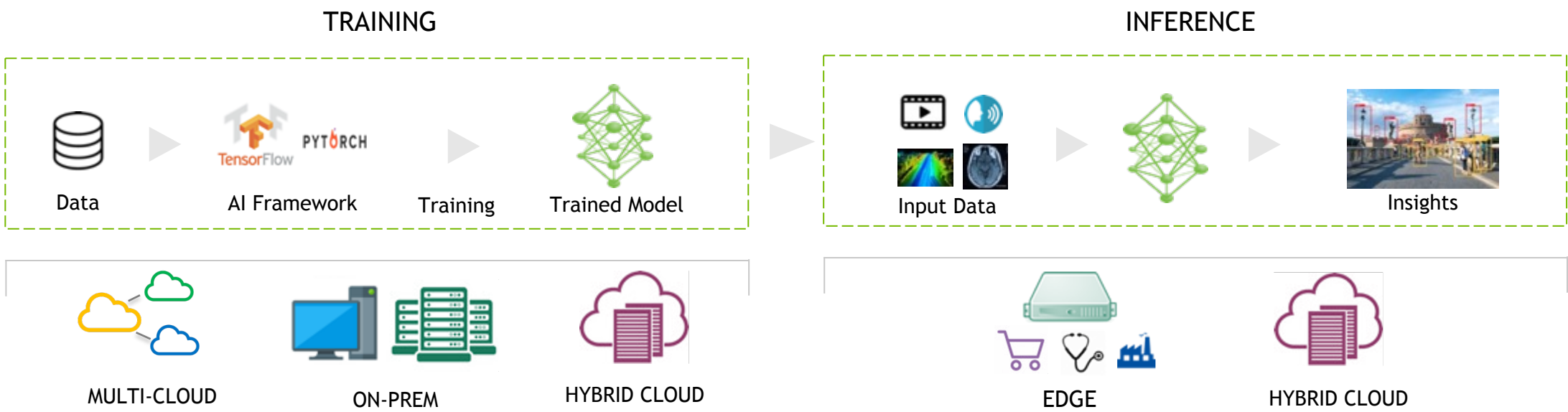
AUTONOMOUS VEHICLES

Training and Inference at Scale



AI WORKFLOW

Training and Inference



CAMBRIAN EXPLOSIAN OF AI MODELS

Convolutional Networks

Encoder/Decoder

ReLu

BatchNorm

Concat

Dropout

Pooling

Recurrent Networks

LSTM

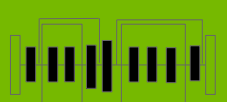
GRU

WaveNet

Beam Search

Transformer

BERT



Megatron

Generative Adversarial Networks

3D-GAN

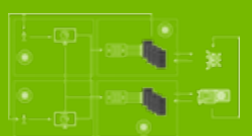
MedGAN

Conditional GAN

Coupled GAN

Speech Enhancement GAN

Reinforcement Learning



DQN

Simulation

DDPG

New Species

Capsule Nets

Mixture of Experts

Wide and Deep

NVIDIA DRIVE INFRASTRUCTURE

HW & SW Infrastructure | Applications, Tools, Services

Applications / Solutions

DATASET | TRAINING | REPLAY | SIMULATION

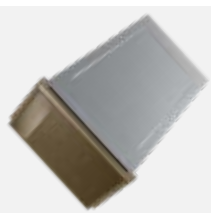
Datacenter Workflow & Management Software

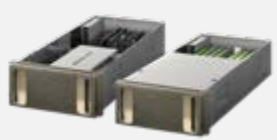
DATASET GENERATION | MODEL GENERATION | TESTING / VALIDATION

DATA MANAGEMENT | WORKFLOW MANAGEMENT

DATA CENTER BACKBONE

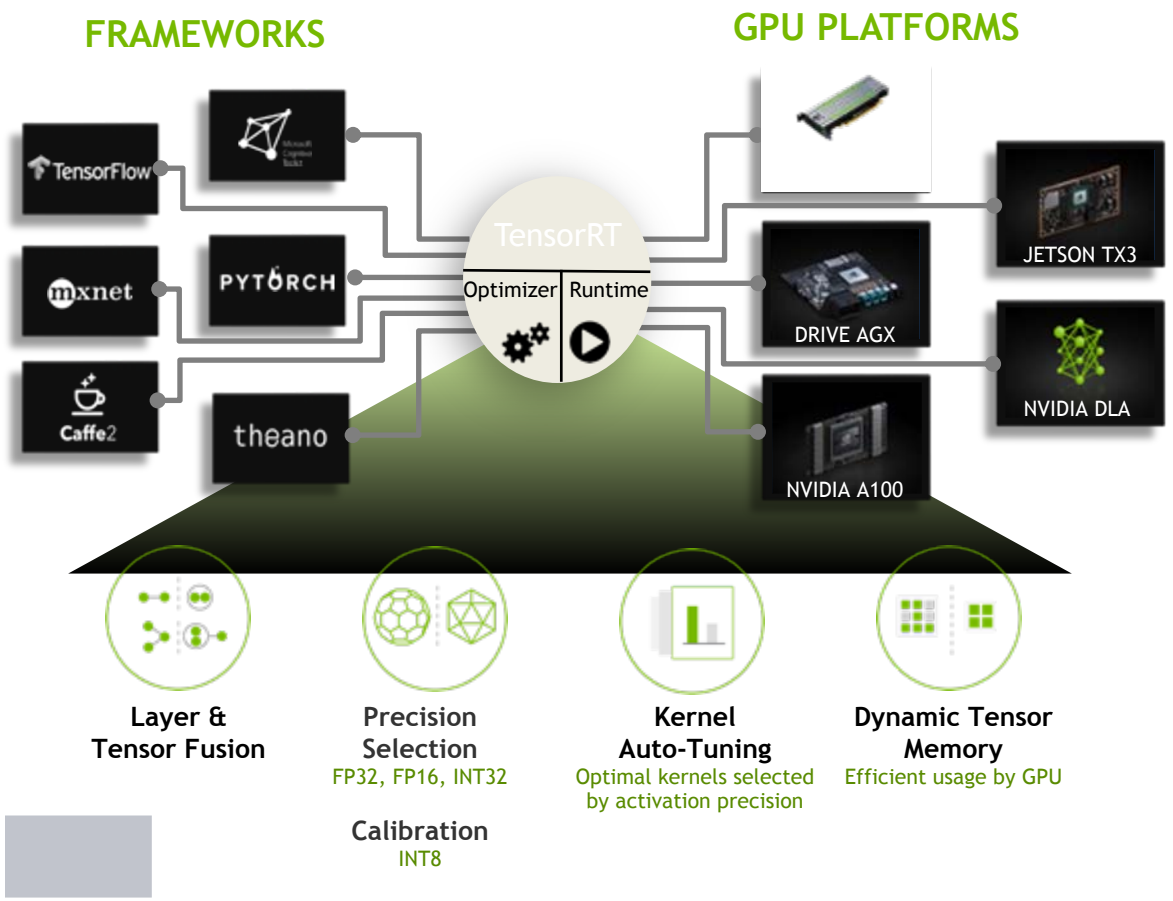
Hardware Appliances (SaturnV @ NVIDIA)

DGX 

CONSTELLATION 

NVIDIA TENSORRT INFERENCE PLATFORM

- AV Stack Pillars: Frameworks and Infrastructure



Automotive

Drive

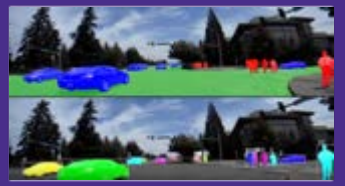
Embedded

Jetson

Data Center

DRIVE SOFTWARE MODELS

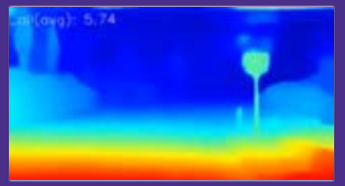
Perception | Mapping | Planning | Driver Monitoring


Obstacles


Distance


Time to Collision (RNN)


Free Space


Structure from Motion


Lidar


Paths


Signs


Map


High Beam


Parking


Camera Blindness


Intersection


Traffic Lights


Gestures/Pose


Gaze


Prediction (RNN)

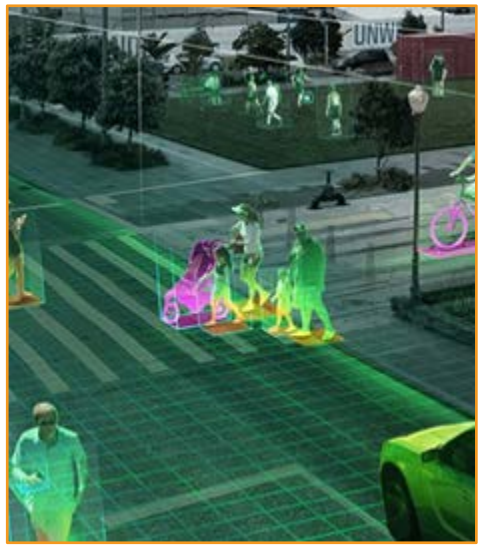

Radar

DEVNET Create

NVIDIA DRIVE SOFTWARE

DRIVE OS FOR SAFETY

The First Functional Safety (FuSa) Operating System for Accelerated Computing and Artificial Intelligence



Automotive Industry
Safety Standards

Scalable
from L2 to L5

Comprehensive
Security Model

Industry Partners
And Experience

https://docs.nvidia.com/drive/drive_os_5.1.6.1L/nvvib_docs/index.html

NVIDIA DRIVE

E2E AV flow to enable large scale AV development & testing

DRIVE AV

Perception | Mapping | Planning

1,000's Engineers
20+ DNNs, 50 Parallel Experiments

DRIVEWORKS

Sensor Abstraction | Vehicle IO,Tools | Calibrate
Image / Point Cloud | Recorder | Ego Motion

1,000's Engineers HW & SW
20 million lines of code

Sensor
Data+Logs

ACTIVE LEARNING
TARGETED LEARNING

TRANSFER LEARNING
FEDERATED LEARNING

INGESTION | CURATION | LABELING | TRAINING | REPLAY | SIMULATION

300k hours, 300M frames
10 PB/wk data collected

1,500 labelers
50M+ labeled images

100's DGX
Saturn V

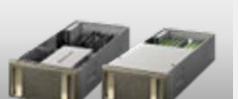
1M+ virtual miles driven
200+ DRIVE Constellations

DRIVE INFRA

Data Management | Workflow Management

Data Center Backbone

1,000's of engineering hours develop and run Deep Ops & Dev Ops SW on Saturn V



ENABLING PORTABILITY WITH NGC CONTAINERS

NGC Deep Learning Containers

Extensive

- Diverse range of workloads and industry specific use cases

Optimized

- DL containers updated monthly
- Packed with latest features and superior performance

Secure & Reliable

- Scanned for vulnerabilities and crypto
- Tested on workstations, servers, & cloud instances

Scalable

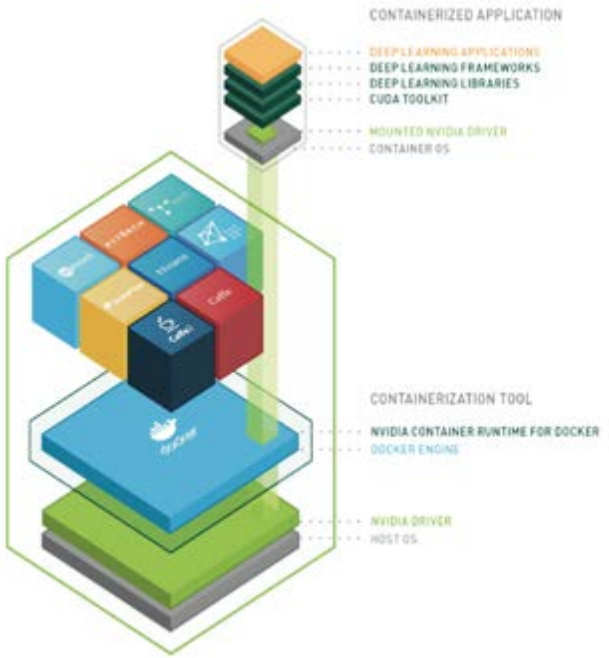
- Supports multi-GPU & multi-node systems

Designed for Enterprise & HPC

- Supports Docker, Singularity & other runtimes

Run Anywhere

- Bare metal, VMs, Kubernetes
- x86, ARM, POWER
- Multi-cloud, on-prem, hybrid, edge



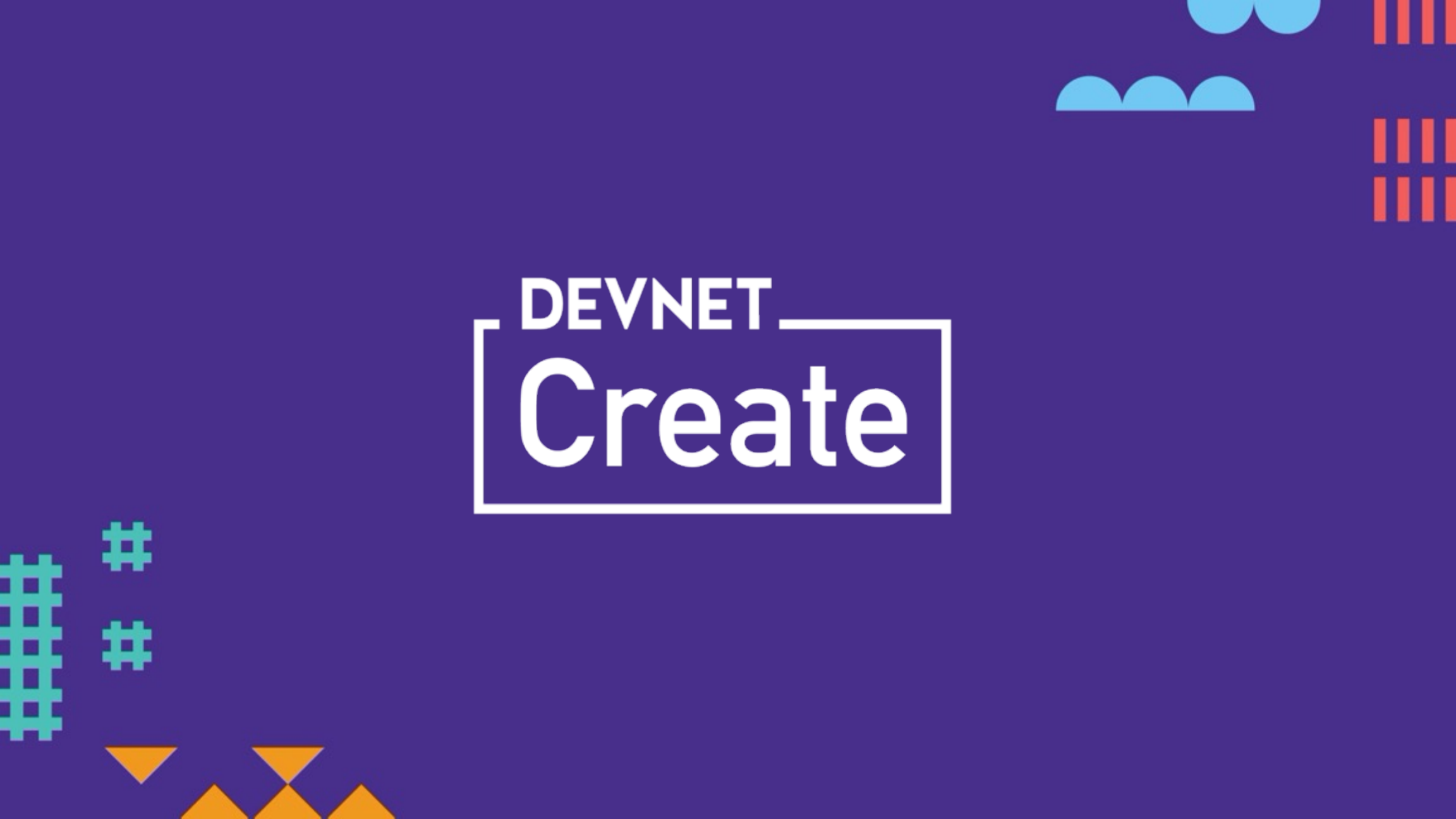
Learn more about NGC Containers

CONCLUSION

- Key Pointers
  - NVIDIA AV Full Stack Scalable solution
    - AV Infrastructure addressing Training, Replay and Inference at Scale
    - AV stack developed on Standardized Frameworks
    - Customizable AV Workflow management tools
  - DRIVE AGX Platform for production deployment
    - HW Platform: DRIVE Xavier / Orin + Auto-grade discrete GPU
    - SW Platform: DRIVE OS, DRIVEWORKS, DRIVE AV, DRIVE IX
    - Safety certified HW and SW platform

DEVELOPER ENGAGEMENT PLATFORMS

Information, downloads, special programs, code samples, and bug submission	developer.nvidia.com
Containers for cloud and workstation environments	ngc.nvidia.com
Insights & help from other developers and NVIDIA technical staff	devtalk.nvidia.com
Technical documentation	docs.nvidia.com
Deep Learning Institute: workshops & self-paced courses	courses.nvidia.com
In depth technical how to blogs	devblogs.nvidia.com
Developer focused news and articles	news.developer.nvidia.com
Webinars	nvidia.com/webinar-portal
GTC on-demand content	gputechconf.com


DEVNET
Create