

GTC2018 Japan



FROM THE CAR TO THE CLOUD

自動運転からサポート・クラウドまでのトータル・システム

Toru Baji NVIDIA Technology Adviser, GPU Evangelist

馬路 徹 NVIDIA 技術顧問、GPUエバンジェリスト

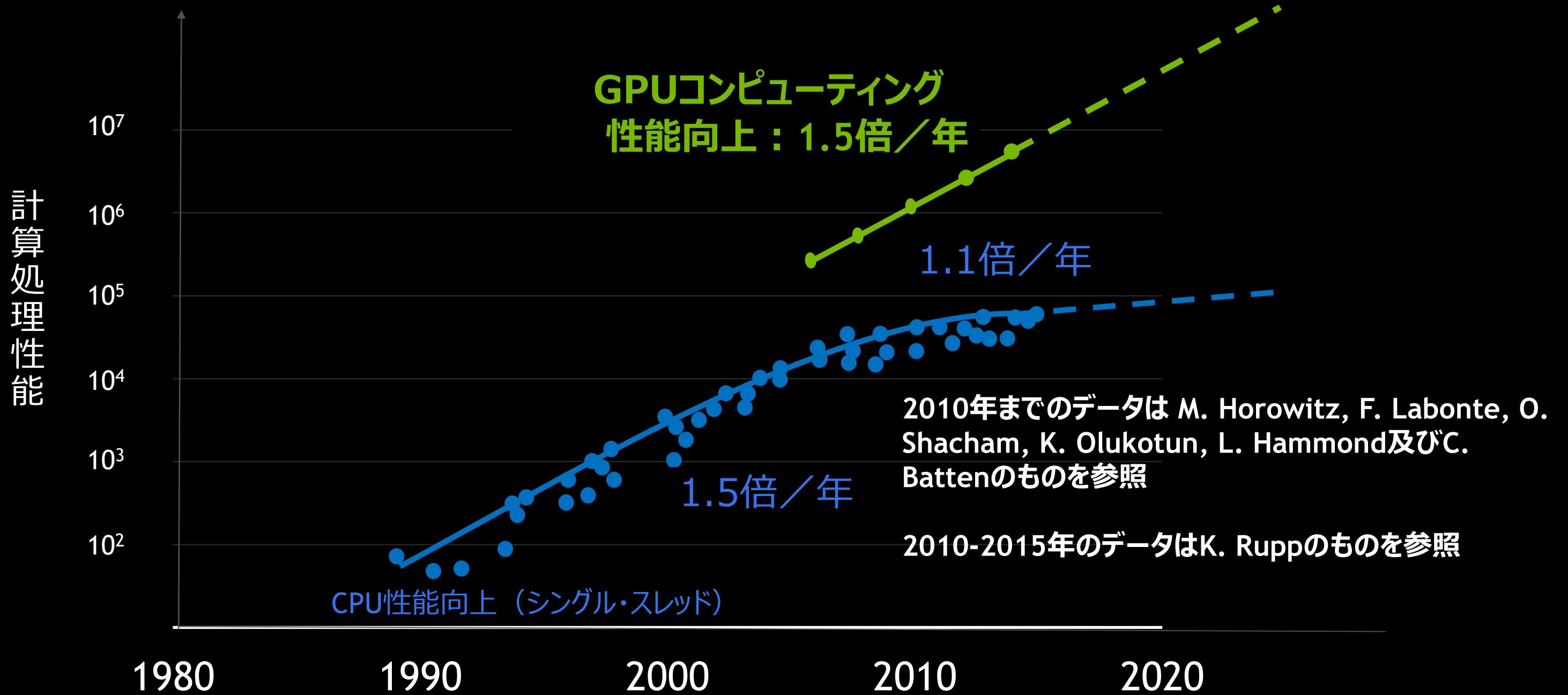
Sept. 13, 2018

講演目次

- 1. FROM THE CARS: NVIDIA自動運転プラットフォームの最新状況**
 - 1) NVIDIA 高性能、高効率GPUの広範な応用
 - 2) ビッグGPUと同一のGPUアーキテクチャをもつ自動運転用SoC
 - 3) レベル3 - レベル5（完全自動運転）用SoC Xavier及びDRIVEプラットフォーム
- 2. TO THE CLOUD: 学習、シミュレーション、検証までのEnd to Endソリューション**
 - 1) 自動運転車開発に必須なEnd-to-Endソリューションとは
 - 2) 自動運転用AIモデルの学習
 - 3) 自動運転用シミュレーションとテスト
 - 4) 自動運転システム開発及び量産インフラストラクチャ

講演目次

- 1. FROM THE CARS: NVIDIA自動運転プラットフォームの最新状況**
 - 1) NVIDIA 高性能、高効率GPUの広範な応用
 - 2) ビッグGPUと同一のGPUアーキテクチャをもつ自動運転用SoC
 - 3) レベル3 - レベル5（完全自動運転）用SoC Xavier及びDRIVEプラットフォーム
- 2. TO THE CLOUD: 学習、シミュレーション、検証までのEnd to Endソリューション**
 - 1) 自動運転車開発に必須なEnd-to-Endソリューションとは
 - 2) 自動運転用AIモデルの学習
 - 3) 自動運転用シミュレーションとテスト
 - 4) 自動運転システム開発及び量産インフラストラクチャ



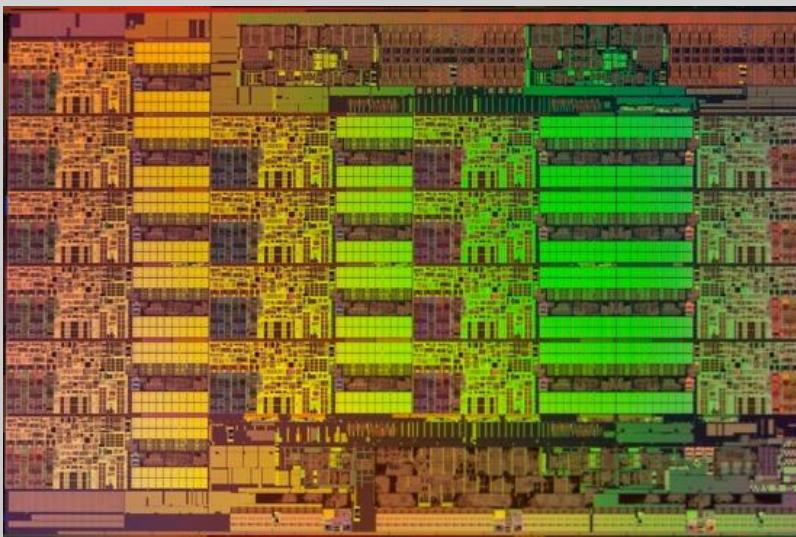
CPU VS GPU : 大きな電力効率の差

CPU

126 pJ/flop (SP)

Optimized for Latency

Deep Cache Hierarchy



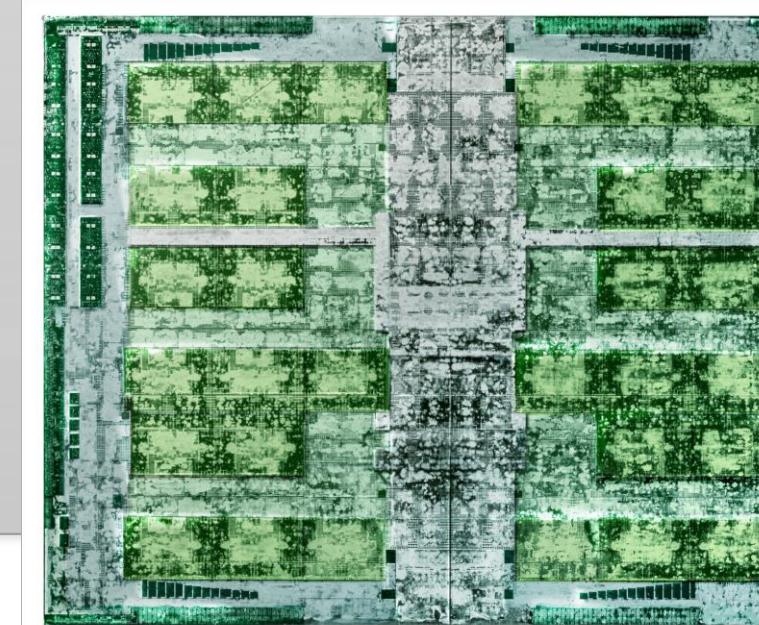
Broadwell E5 v4
14 nm

GPU

28 pJ/flop (SP)

Optimized for Throughput

Explicit Management
of On-chip Memory



Pascal P100
16 nm

世界最速のスーパーコンピューター 7 台のうち、5 台が NVIDIA GPU を搭載

フランクフルト - 国際スーパーコンピューティング会議 (ISC) – 2018 年 6 月 25 日

- 1) 性能 : TOP500 リストにおいて増大した演算性能の 56% は NVIDIA の GPU がもたらしたもの
- 2) 電力効率 : GREEN500 リストに掲載された、エネルギー効率の上位 20 システムのうち、17 のシステムに GPU が搭載
- 3) AI 性能 : NVIDIA の GPU を採用する TOP500 Lists の 1 位・3 位・5 位のスーパーコンピューターのディープラーニングにおける演算性能の合計値は、TOP500 残りの 497 基のスーパーコンピューターの全ディープラーニング演算性能を上回る

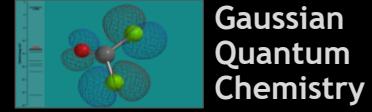
Ranking	Rmax (TFLOPS)	Super Computer/Country	Processor
1	122,300	Summit / USA	IBM Power9 22C, NVIDIA Volta GV100 GPU
2	93,015	Sunway TaihuLight / China	Sunway SW26010 260C
3	71,610	Sierra / USA	Power9 22C, NVIDIA Volta GV100 GPU
4	61,445	Tianhe-2A	Xeon E5-269v2 12C
5	19,880	AI Bridging Cloud Infrastructure (ABCI) / Japan	PRIMERGY CX2550 M4, Xeon Gold 6148 20C, NVIDIA Tesla V100 SXM2 (GV100 GPU)
6	19,590	Piz Daint / Switzerland	Cray XC50, Xeon E5, NVIDIA Tesla P100 GPU
7	17,590	Titan / USA	Cray XK7, Opteron 6274, NVIDIA K20x GPU
8	17,173	Sequoia / USA	BlueGene/Q, Power BQC 16C
9	14,137	Trinity / USA	Cray XC40, Xeon Phi 7250
10	14,015	Cori / USA	Cray XC40, Xeon Phi 7250

NVIDIA GPUの 今日の 広範な応用

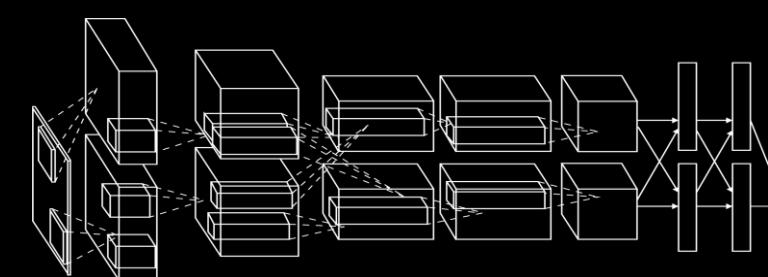
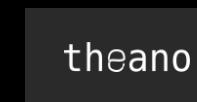
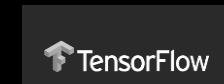
コンピュータ・グラフィックス



科学計算



AI/ディープラーニング



データ解析・データベース

K-Means Clustering

Gradient Boosting

Generalized Linear Model

Support Vector Machine

ANALYTICS



DATABASES



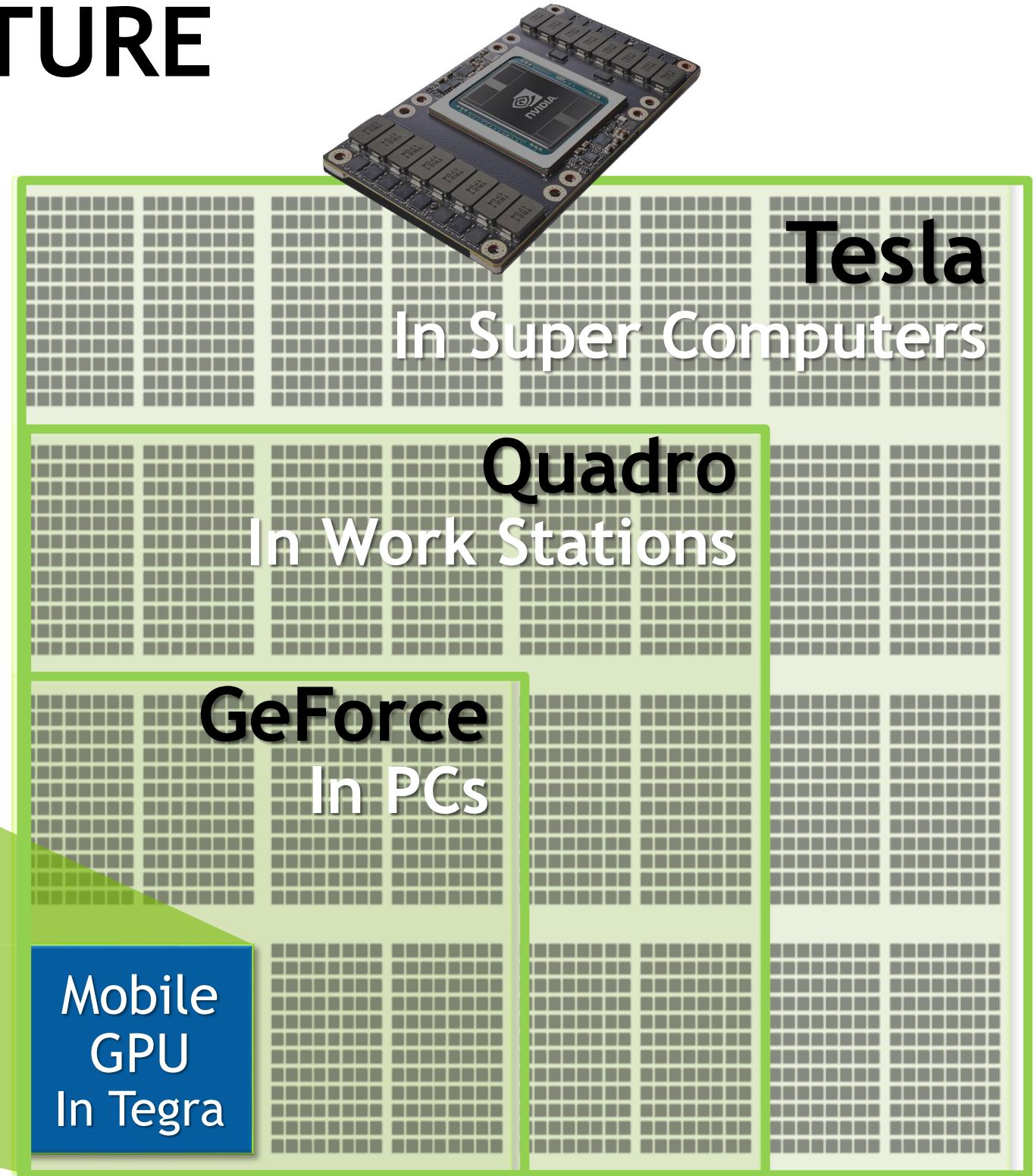
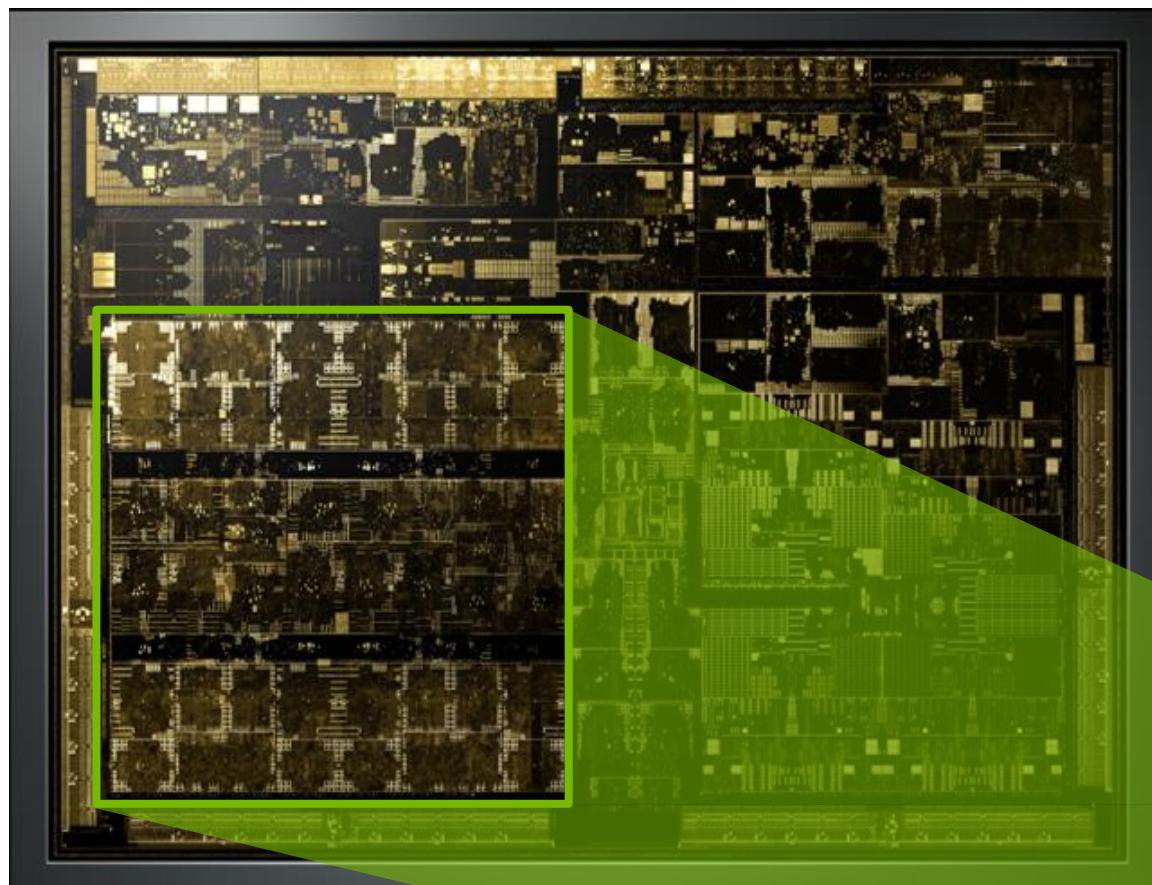
CUDA: Compute Unified
Device Architecture

NVIDIA CUDA (超並列計算プラットフォーム及びプログラミング・モデル)

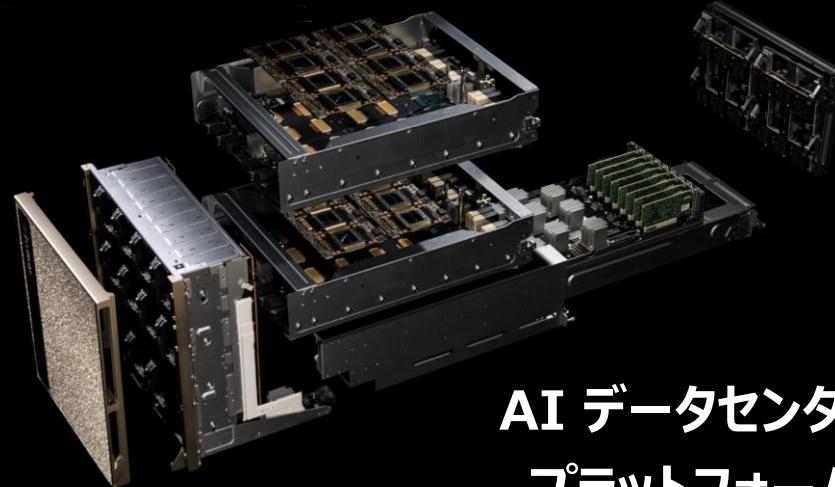
NVIDIA ONE-ARCHITECTURE

スーパーコンピュータから車載SOCまで

自動運転車載 AI プロセッサ Xavier



一貫した自動運転用アーキテクチャ



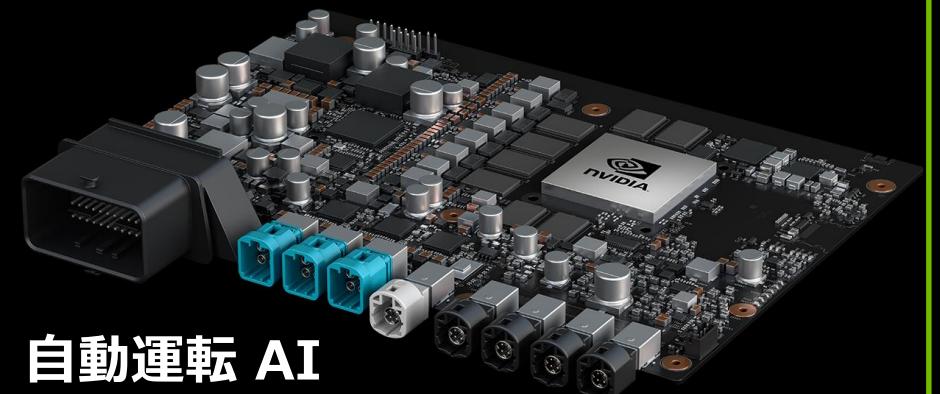
AI データセンター
プラットフォーム



データセンタGPU



車載 AI プロセッサ



自動運転 AI
プラットフォーム

ONE ARCHITECTURE

NVIDIA DRIVE 機能安全アーキテクチャ



System Operates Safely Even when Faults Detected
Holistic System – Process & Methods, Processor Design, Software, Algorithms, System Design, Validation
ISO 26262 ASIL-D Safety Level | Partnership with BlackBerry QNX and TTTech | New AutoSIM Virtual Reality 3D Simulator

XAVIER 機能安全にも万全を尽くした自動運転用AIプロセッサ

Diverse Engines

- Computation with GPU/CPU
- DL with GPU/DLA
- CV with GPU/PVA
and more

Dual Execution

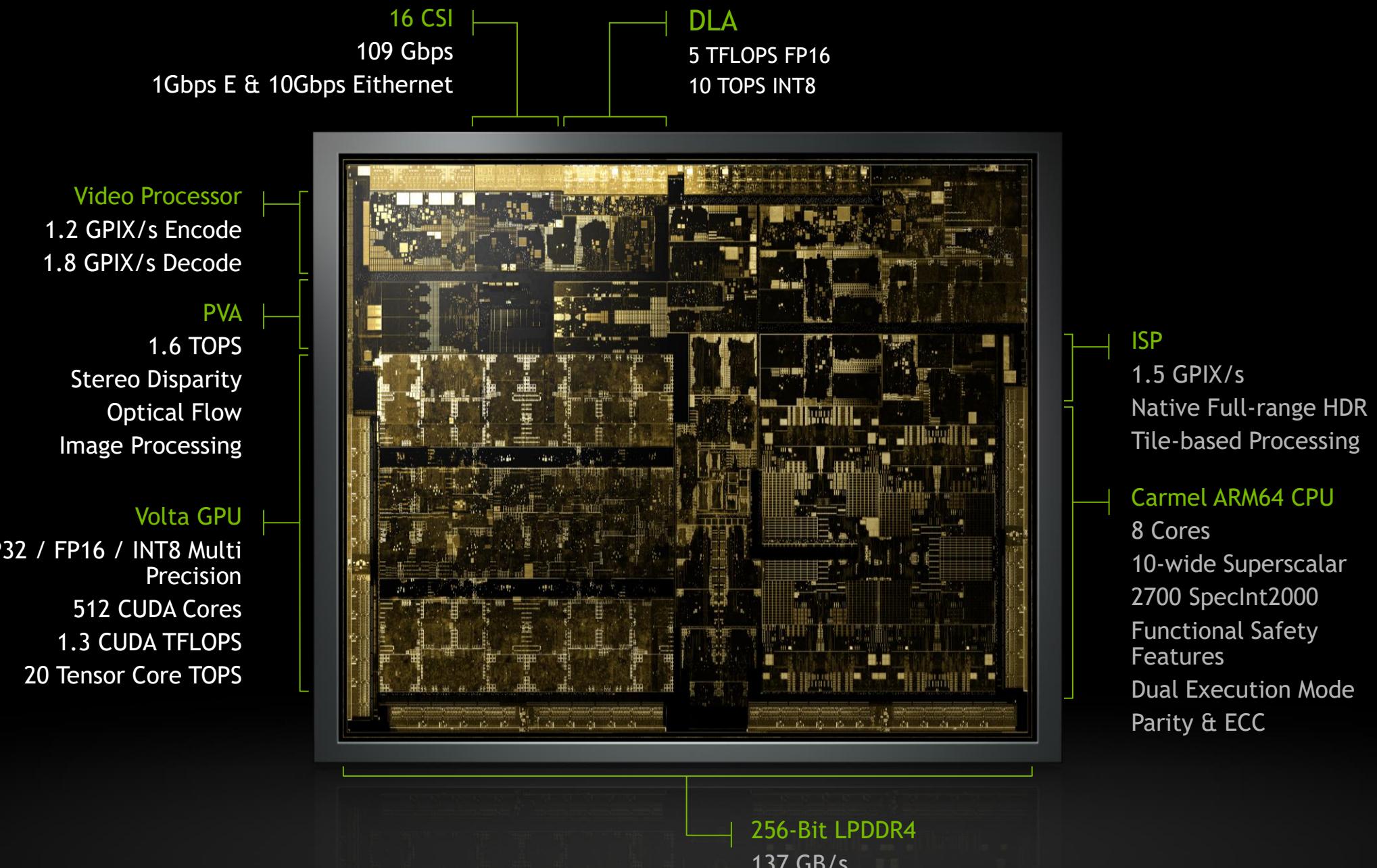
- Carmel ARM64 CPU has a dual execution mode (duplicate instruction streams)

ECC/Parity

- On chip SRAMs, caches, registers
- External DDR memories

Diagnosis, BIST

- SCE (Safety Cluster Engine) with Lock-step ARM Cortex R5 processor pair



DL: Deep Learning

CV: Computer Vision

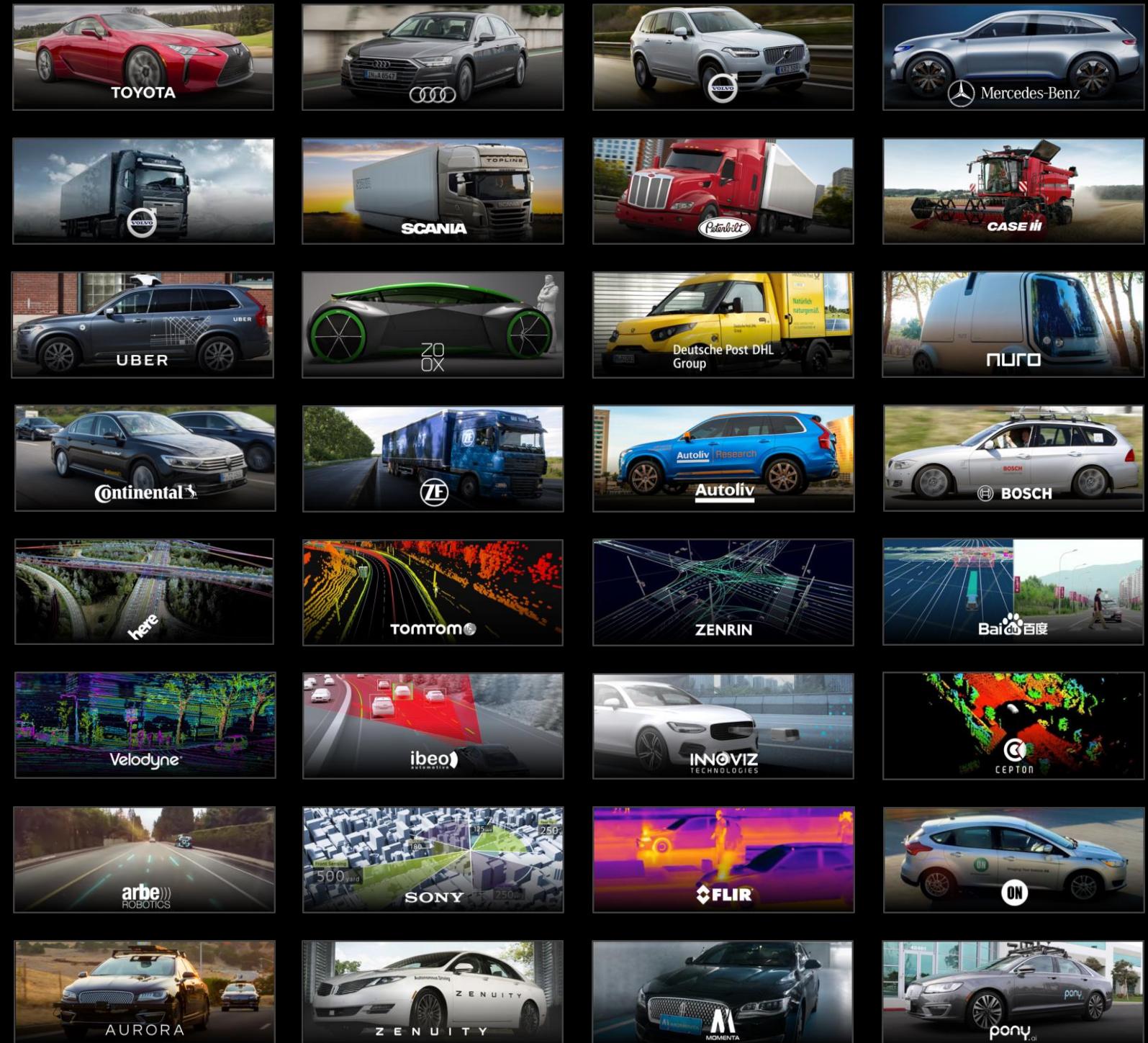
DLA: Deep Learning Accelerator

PVA: Programmable Vision Accelerator

ISP: Image Signal Processor

Most Complex SOC Ever Made | 9 Billion Transistors, 350mm², 12nFFN
TÜV SÜD's team determined Xavier's architecture meets the ISO 26262 requirements
to avoid unreasonable risk in situations that could result in serious injury.

370 のパートナーが NVIDIA DRIVE で開発中



自動車

トラック

モビリティ
サービス

サプライヤー

マッピング

LIDAR

カメラ /
レーダー

スタートアップ

ボッシュとダイムラー：自動運転の実証試験をカリフォルニアで行う計画 自動運転車両のコントロールユニットのカギとなるプラットフォームはNVIDIAから

- 2019年の下半期に特定のルートを走る自動運転シャトルのサービスを開始
- ごくわずかな時間でセンサーからの情報を分析・解釈し、走行に必要な情報として入力するには膨大な計算処理能力が求められる。つまり完全に自律でドライバーレスの自動運転車は、いわば移動するスーパーコンピュータのようなもの
- 実現には多用途かつ冗長なシステム構造と最大限の安全機能を兼ね備えなければならない。これを実現するには必要な計算処理を複数の回路で平行して行う構造
- ボッシュとダイムラーが車両制御用に開発したアルゴリズムを実行するためのプラットフォームはアメリカ半導体企業Nvidiaから供給されている



MotorFan TECH 8/3/2018

<https://motor-fan.jp/tech/10005085>

DRIVE PX XAVIER

開発時：DRIVE PX2
量産時：DRIVE PX XAVIER



DRIVE PX 2

2 PARKER SoC + 2 PASCAL GPU

| 20 TOPS DL | 120 SPECINT | 80W

DRIVE PX XAVIER

20 TOPS DL | 160 SPECINT | 20W

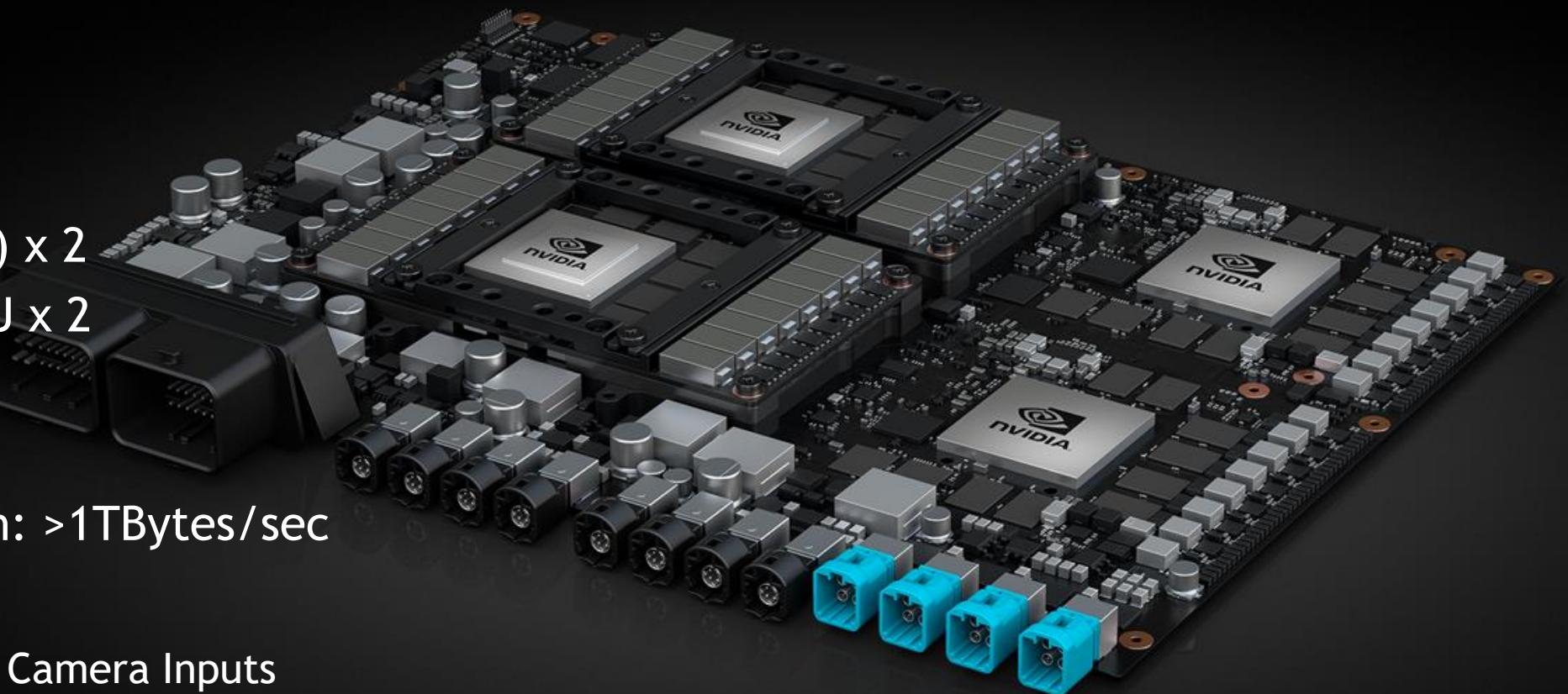
(20 TOPS DL/20Wは高電力効率使用の場合)

DRIVE AGX PEGASUS

ROBOTAXI DRIVE PLATFORM

レベル5 完全自動運転

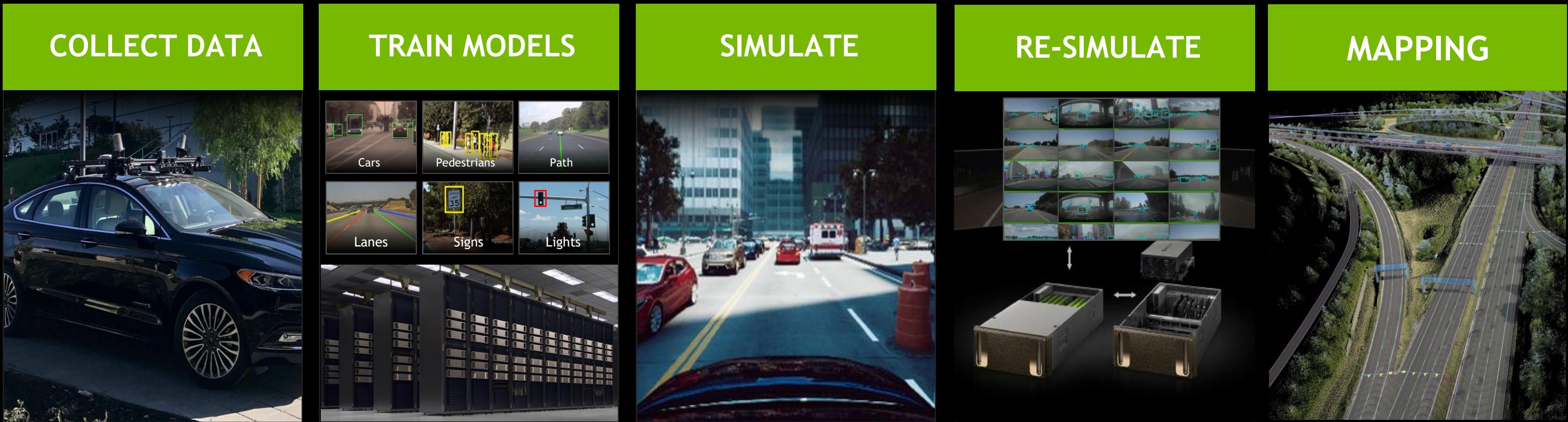
- Xavier (Volta GPU integrated) x 2
- Next generation discrete-GPU x 2
- 320 TOPS CUDA TensorCore
- ASIL D Certification
- Combined Memory Bandwidth: >1TBytes/sec
- Automotive I/Os
 - 16x GMSL High-speed Camera Inputs
 - Multiple 10Gbit Ethernet
 - CAN, Flexray
- 400W
- Late Q1 Early Access Partners
- Supercomputing Data Center in your Trunk



講演目次

- 1. FROM THE CARS: NVIDIA自動運転プラットフォームの最新状況**
 - 1) NVIDIA 高性能、高効率GPUの広範な応用
 - 2) ビッグGPUと同一のGPUアーキテクチャをもつ自動運転用SoC
 - 3) レベル3 - レベル5（完全自動運転）用SoC Xavier及びDRIVEプラットフォーム
- 2. TO THE CLOUD: 学習、シミュレーション、検証までのEnd to Endソリューション**
 - 1) 自動運転車開発に必須なEnd-to-Endソリューションとは
 - 2) 自動運転用AIモデルの学習
 - 3) 自動運転用シミュレーションとテスト
 - 4) 自動運転システム開発及び量産インフラストラクチャ

自動運転の全ての開発工程をサポート

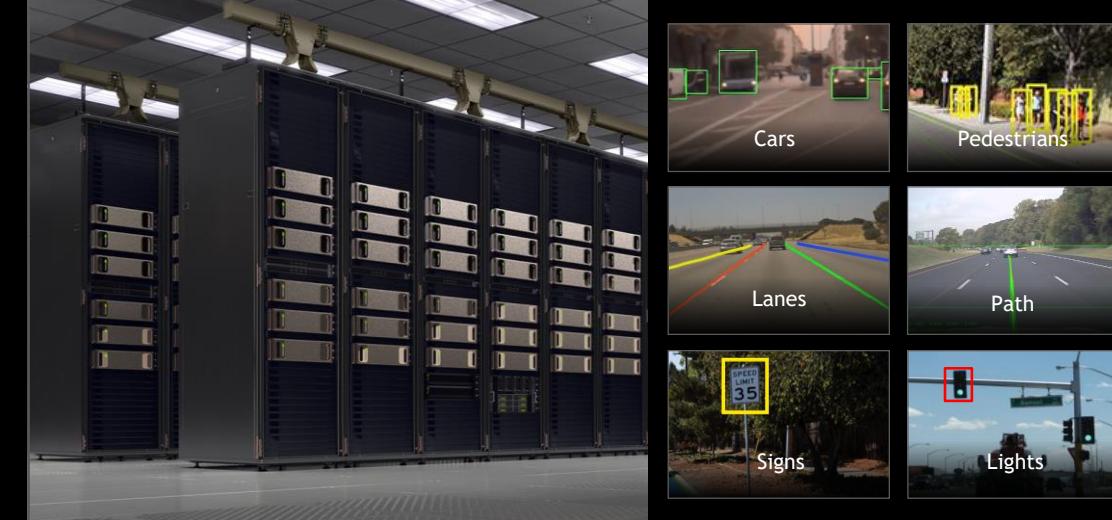


NVIDIA DRIVE END-TO-END プラットフォーム

COLLECT & PROCESS DATA



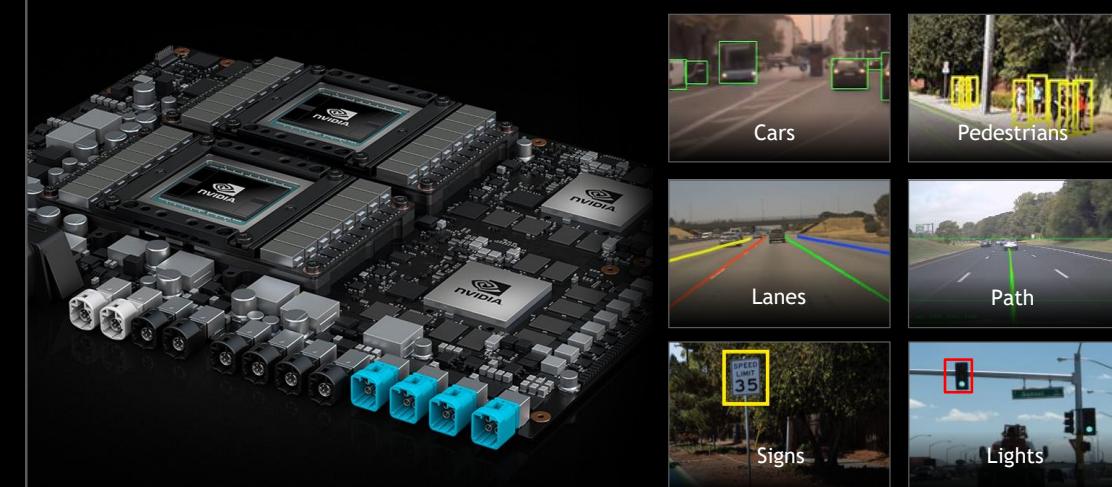
TRAIN MODELS



SIMULATE



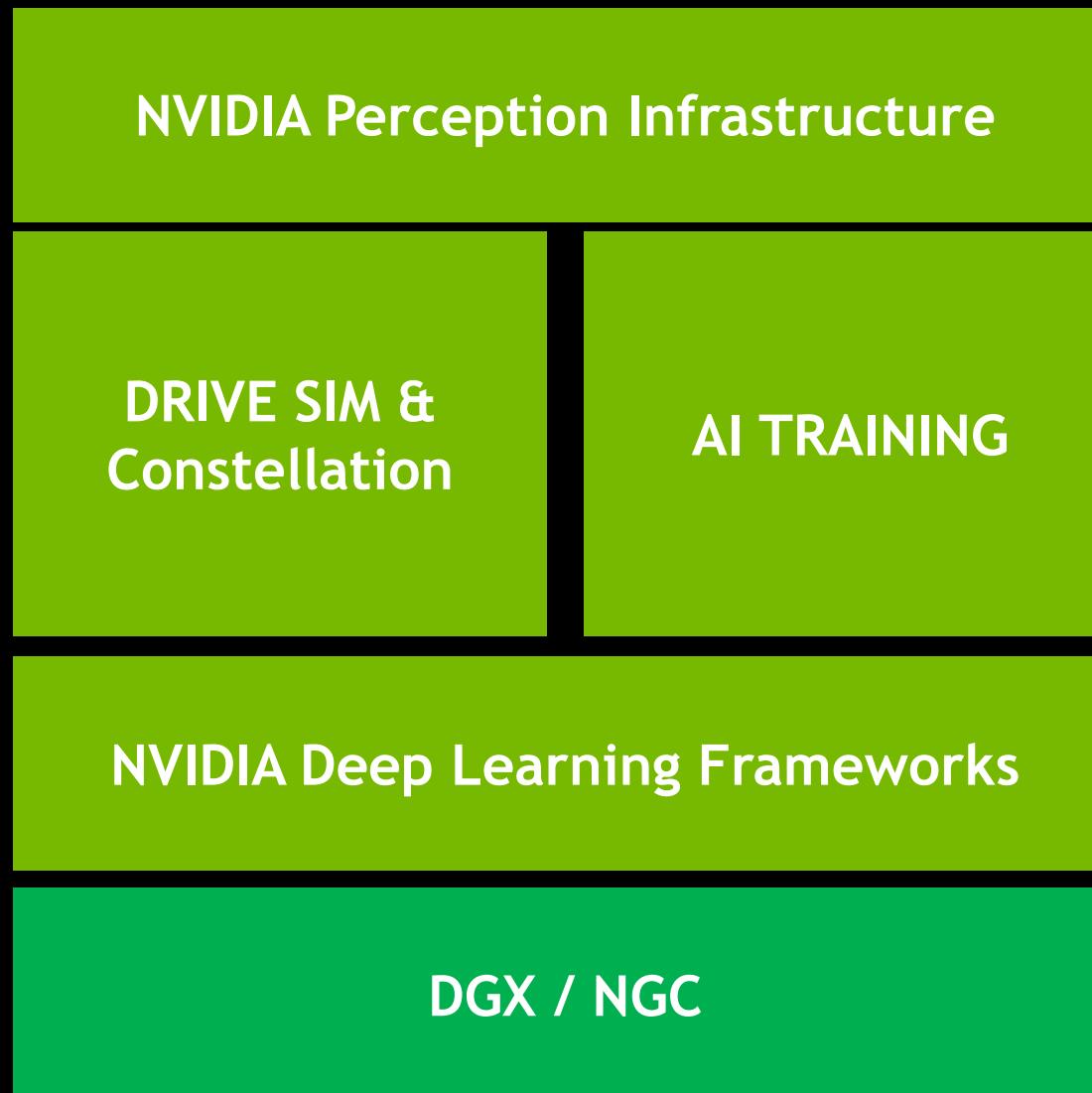
DRIVE



NVIDIA END TO END DRIVE プラットフォーム

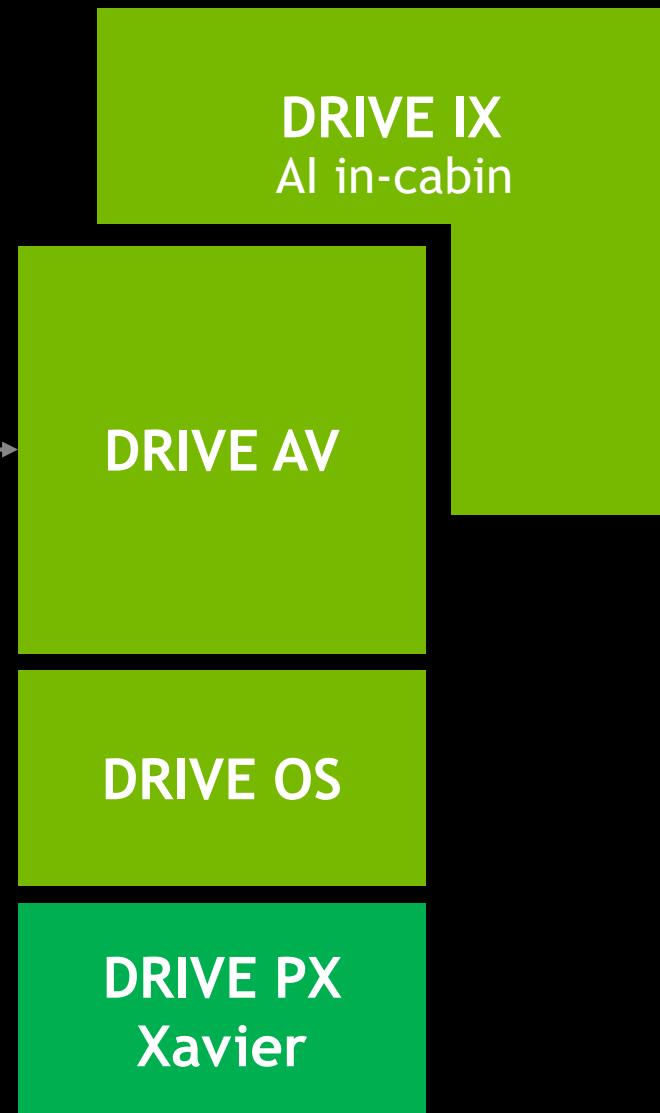
DRIVE TSTADI

(Training, Simulation, Testing for Autonomous Driving Infrastructure)



データセンタ

DRIVE AI

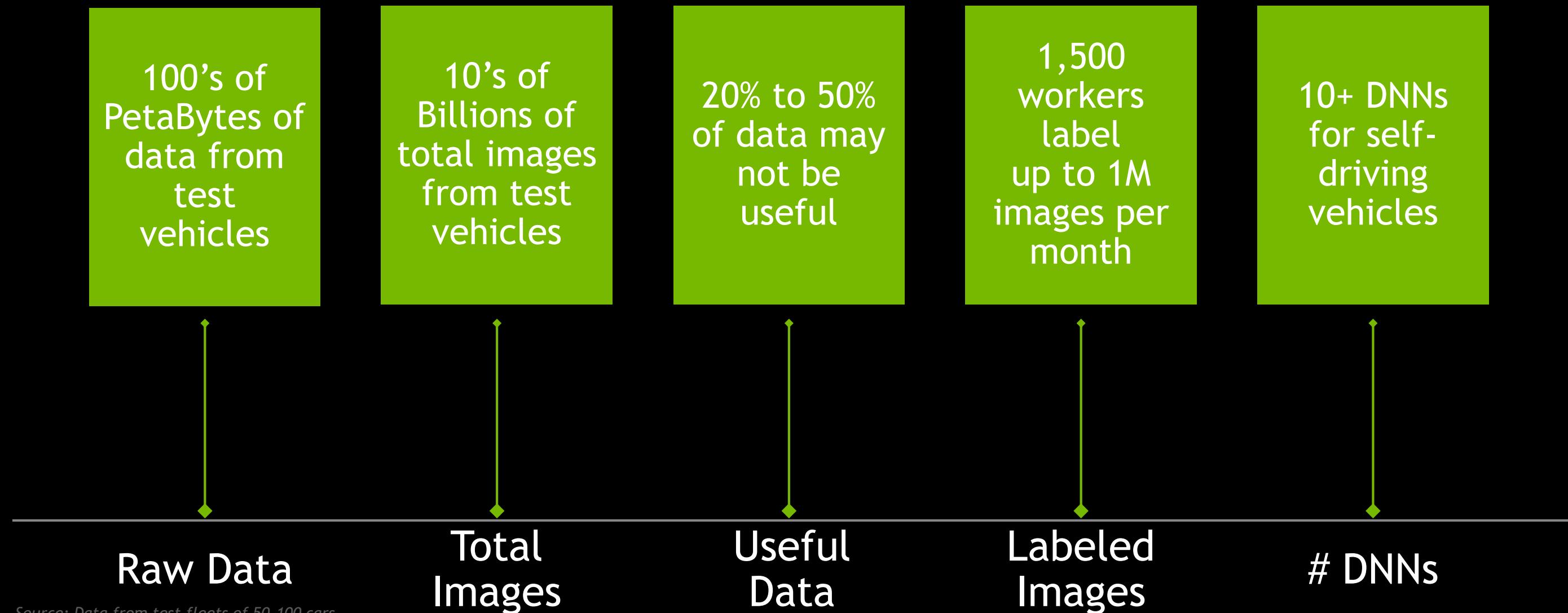


自動運転車

講演目次

- 1. FROM THE CARS: NVIDIA自動運転プラットフォームの最新状況**
 - 1) NVIDIA 高性能、高効率GPUの広範な応用
 - 2) ビッグGPUと同一のGPUアーキテクチャをもつ自動運転用SoC
 - 3) レベル3 - レベル5（完全自動運転）用SoC Xavier及びDRIVEプラットフォーム
- 2. TO THE CLOUD: 学習、シミュレーション、検証までのEnd to Endソリューション**
 - 1) 自動運転車開発に必須なEnd-to-Endソリューションとは
 - 2) 自動運転用AIモデルの学習
 - 3) 自動運転用シミュレーションとテスト
 - 4) 自動運転システム開発及び量産インフラストラクチャ

AI学習用データ収集及びラベル付けの規模感



センサ種別による学習データ規模感

8-Hrs/day
250-days/year

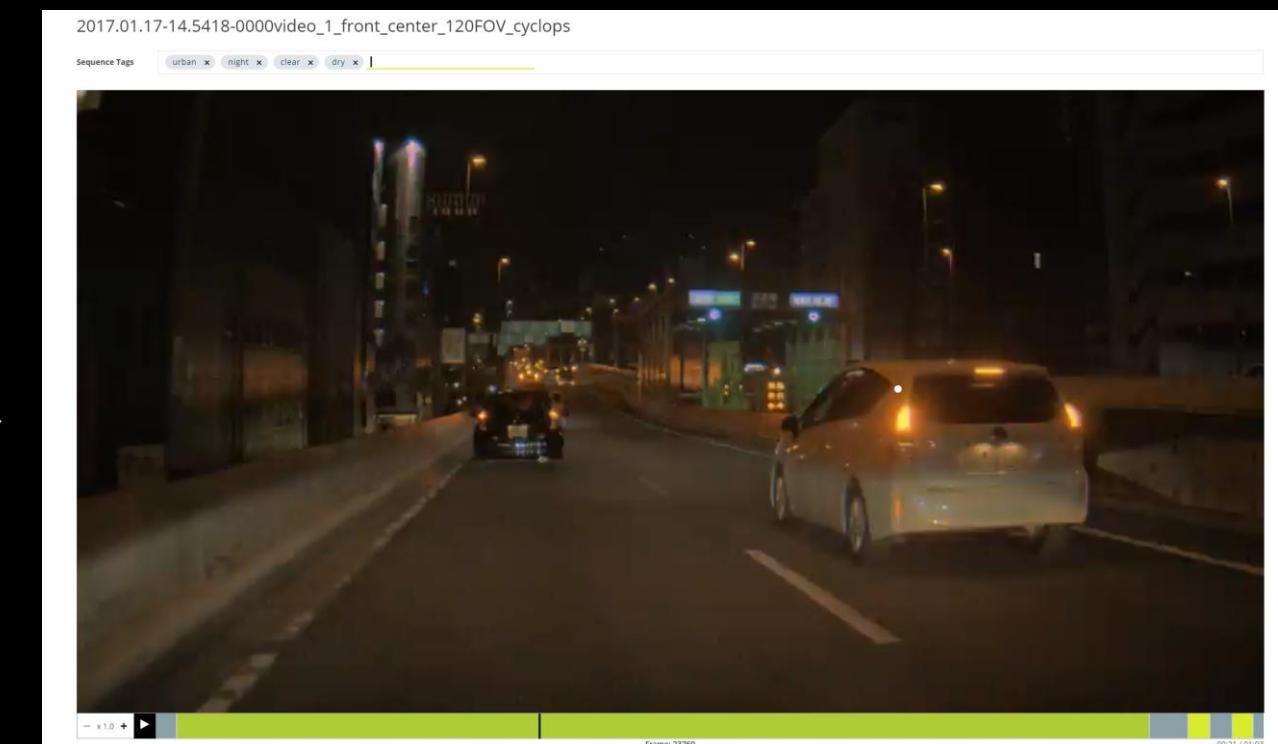
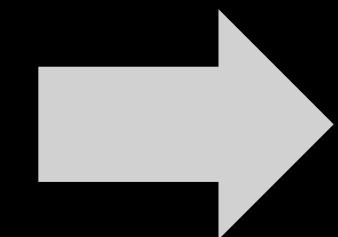
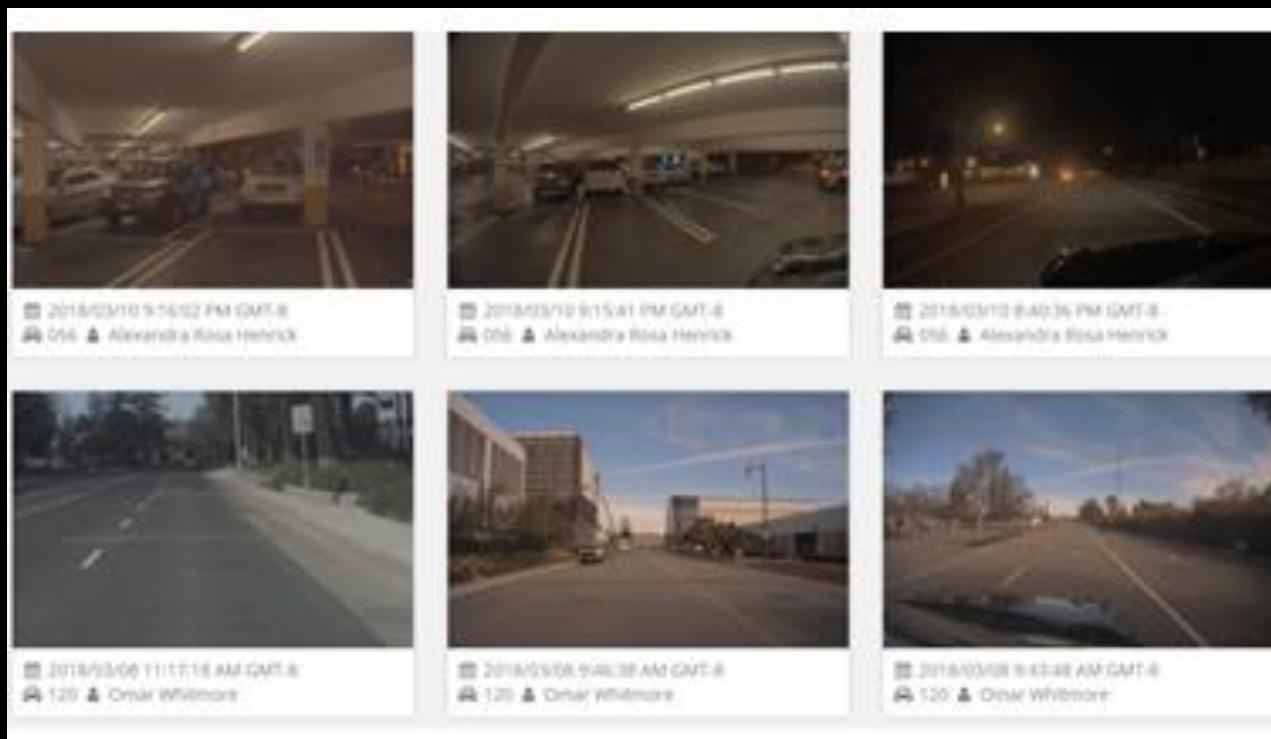
Sensor type	Resolution	Frames per second	Sensors per car	Images per year	Storage needed	Labeled images	DGX-1 systems
Camera	2 MPixel	30	5	1B	2 PB	1M	25
Lidar	0.13 MPixel	10	3	200M	400 TB	200K	5



Carefully Curated
Only 1/1000 to be labeled

DATA CURATION & INDEXING

SELECTING THE MOST INTERESTING DATA FOR LABELING

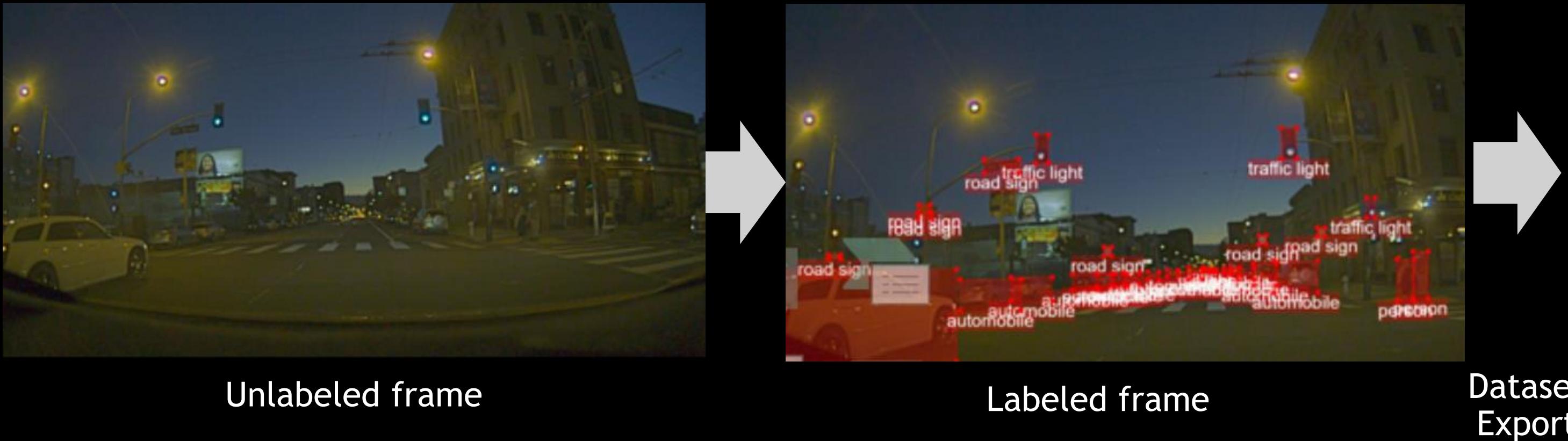


Search from recorded sessions

Frame selection

DATA LABELING & EXPORT

ENSURING THE BEST QUALITY LABELED IMAGES



Standard guidelines and processes are required to correctly annotate frames

Producing high quality labeled data exported for model training purpose

QA is important

自動運転車の外部環境認識及び内部環境認識AI群



Customer Application

外部環境認識、判断AI

DRIVE AV

Object, Freespace, Path/Lane,
Path Planning,
Wait, Map, Sign, Lights, Road Marking
Surround

内部環境認識AI

DRIVE IX

Gaze, Eye Openness, Head Pose,
Gestures, Emotions
Facial Recognition
Voice Recognition & Lip Reading



DRIVE OS

自動運転用の10種のAI例

外部環境認識、判断AI



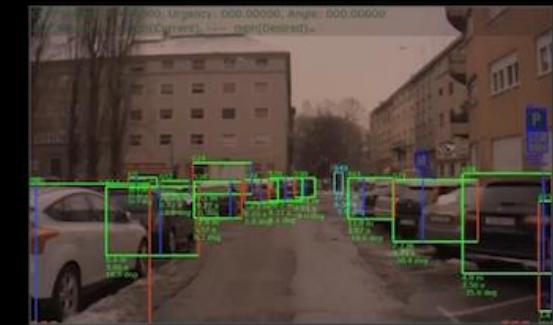
Perception



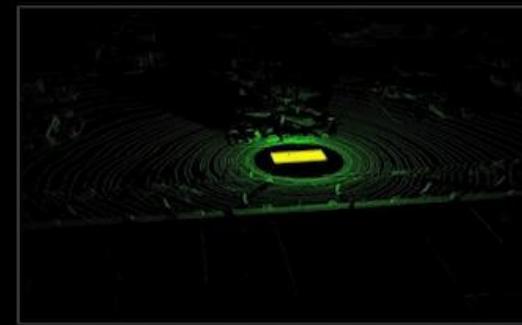
Free Space Perception



Distance Perception



Weather



LIDAR Perception



Camera-based Mapping



Camera Localization to HD Map



LIDAR Localization to HD Map



Path Perception

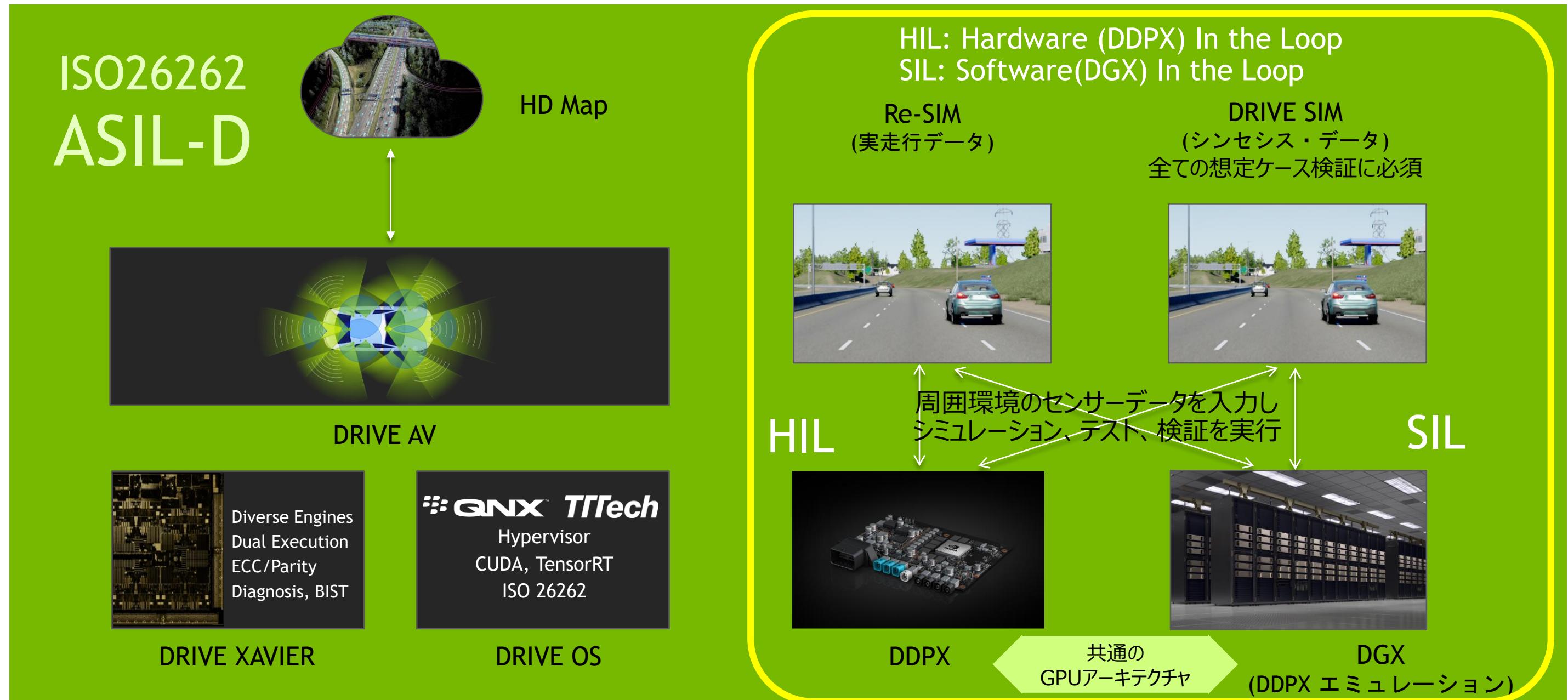


Scene Perception

講演目次

- 1. FROM THE CARS: NVIDIA自動運転プラットフォームの最新状況**
 - 1) NVIDIA 高性能、高効率GPUの広範な応用
 - 2) ビッグGPUと同一のGPUアーキテクチャをもつ自動運転用SoC
 - 3) レベル3 - レベル5（完全自動運転）用SoC Xavier及びDRIVEプラットフォーム
- 2. TO THE CLOUD: 学習、シミュレーション、検証までのEnd to Endソリューション**
 - 1) 自動運転車開発に必須なEnd-to-Endソリューションとは
 - 2) 自動運転用AIモデルの学習
 - 3) 自動運転用シミュレーションとテスト
 - 4) 自動運転システム開発及び量産インフラストラクチャ

NVIDIA DRIVE 機能安全アーキテクチャ



System Operates Safely Even when Faults Detected
Holistic System – Process & Methods, Processor Design, Software, Algorithms, System Design, Validation
ISO 26262 ASIL-D Safety Level | Partnership with BlackBerry QNX and TTTech | New AutoSIM Virtual Reality 3D Simulator

“ Autonomous vehicles need to be driven more than **11 billion miles** to be 20% better than humans. With a fleet of 100 vehicles, 24 hours a day, 365 days a year, at 25 miles per hour, this would take **518 years**.

Rand Corporation, Driving to Safety

シミュレーションの重要性

Waymo has racked up over **eight million self-driving miles since the former Google Car project kicked off almost a decade ago**. But that impressive realworld number is nothing compared to the mileage amassed in virtual reality.

"We do eight million miles every day in simulations," said James Stout, a senior staff software engineer at Waymo.

“A Waymo engineer told us why a virtual-world simulation is crucial to the future of self-driving cars”,
Business Insider Matthew DeBord Aug 16, 2018,

ウーバーの死亡事故について取り上げた[このコラムの過去の記事](#)で、「自動運転車の免許証」について触れた。これは公道を走らせる自動運転車に、**国家がシミュレーションによる自動運転プログラムの「バーチャルテスト」と、実験車両を使った「実技試験」を実施し、これに合格した車両だけを公道走行可能にする**という構想だ。今回の講演で河野氏が触れた「安全性評価」のコンセプトはこれに近い。

「ゴルフカートで自動運転？ 意外な可能性に驚く」 日経 xTECH (クロステック) 2018年8月

シミュレーション BILLIONSマイル走行検証の実現手段



NVIDIA DRIVE SIM AND CONSTELLATION AV VALIDATION SYSTEM

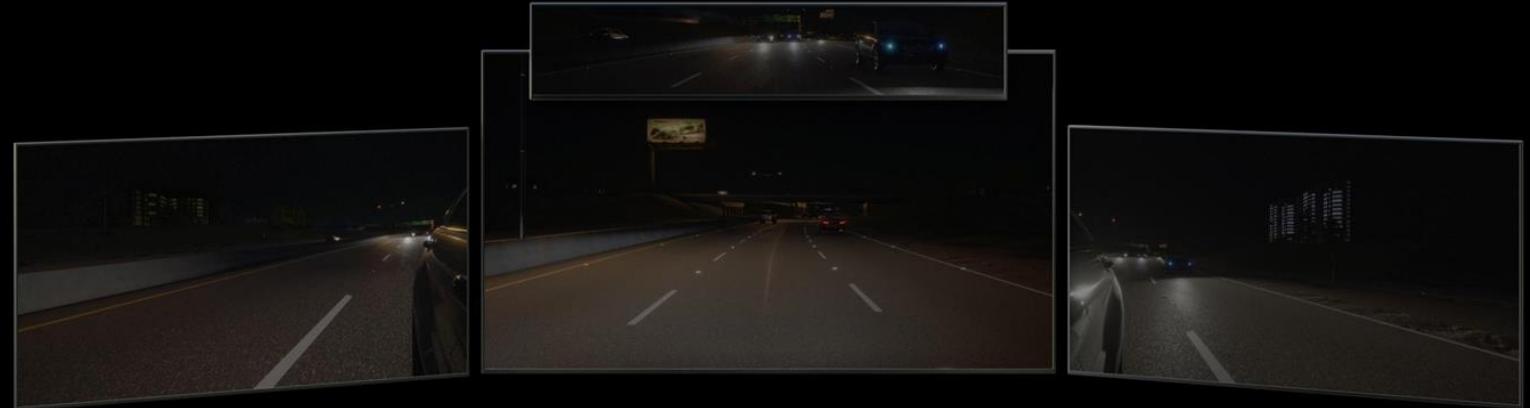
Virtual Reality AV Simulator

Same Architecture as DRIVE Computer

Simulate Rare and Difficult Conditions,

Recreate Scenarios, Run Regression Tests,
Drive Billions of Virtual Miles

10,000 Constellations Drive 3B Miles per Year



Runs NVIDIA DRIVE Sim SW to
simulate sensors on GPU

NVIDIA DRIVE SIM AND CONSTELLATION AV VALIDATION SYSTEM

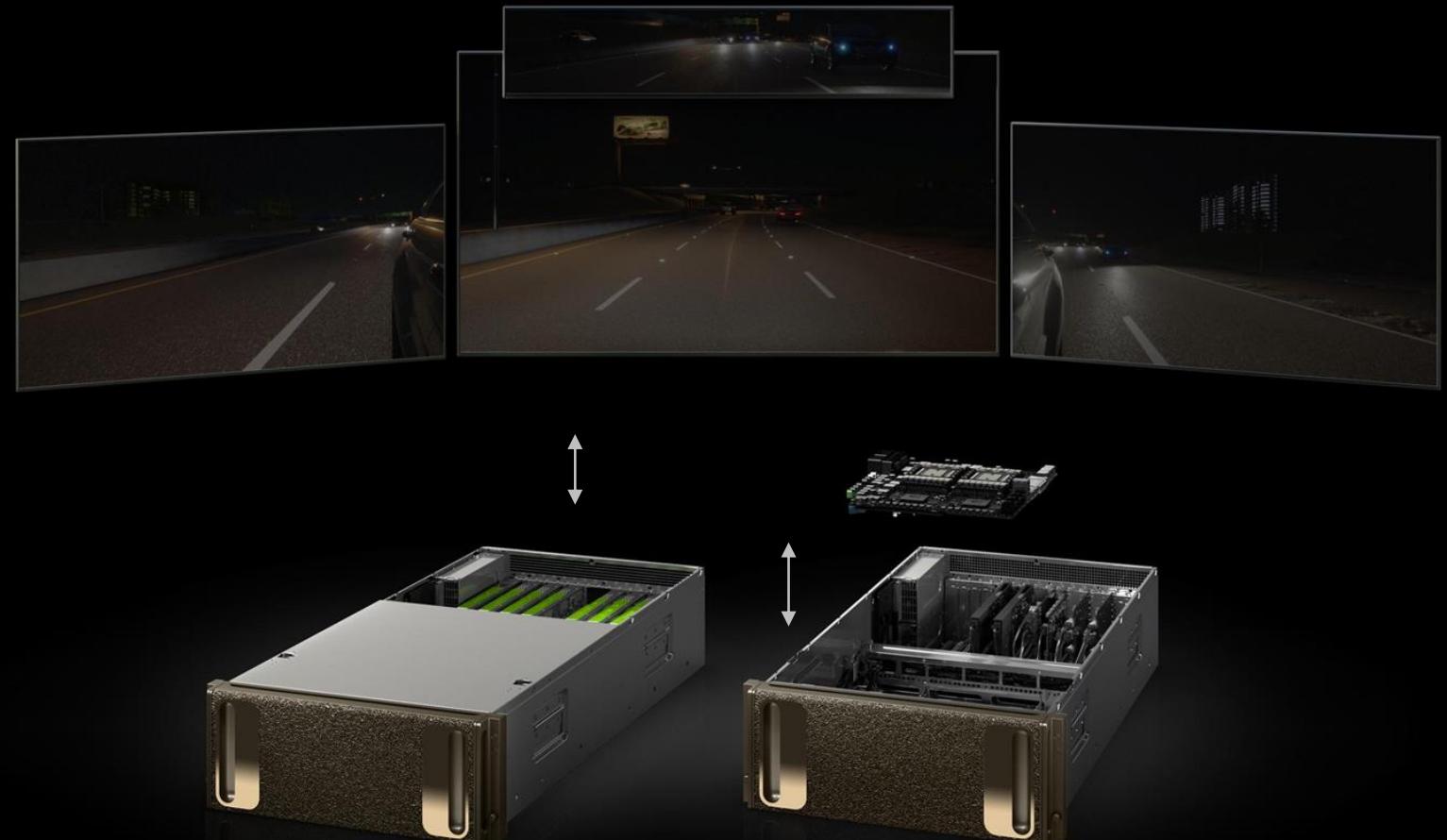
Virtual Reality AV Simulator

Same Architecture as DRIVE Computer

Simulate Rare and Difficult Conditions,

Recreate Scenarios, Run Regression Tests,
Drive Billions of Virtual Miles

10,000 Constellations Drive 3B Miles per Year



NVIDIA DRIVE SIM AND CONSTELLATION AV VALIDATION SYSTEM

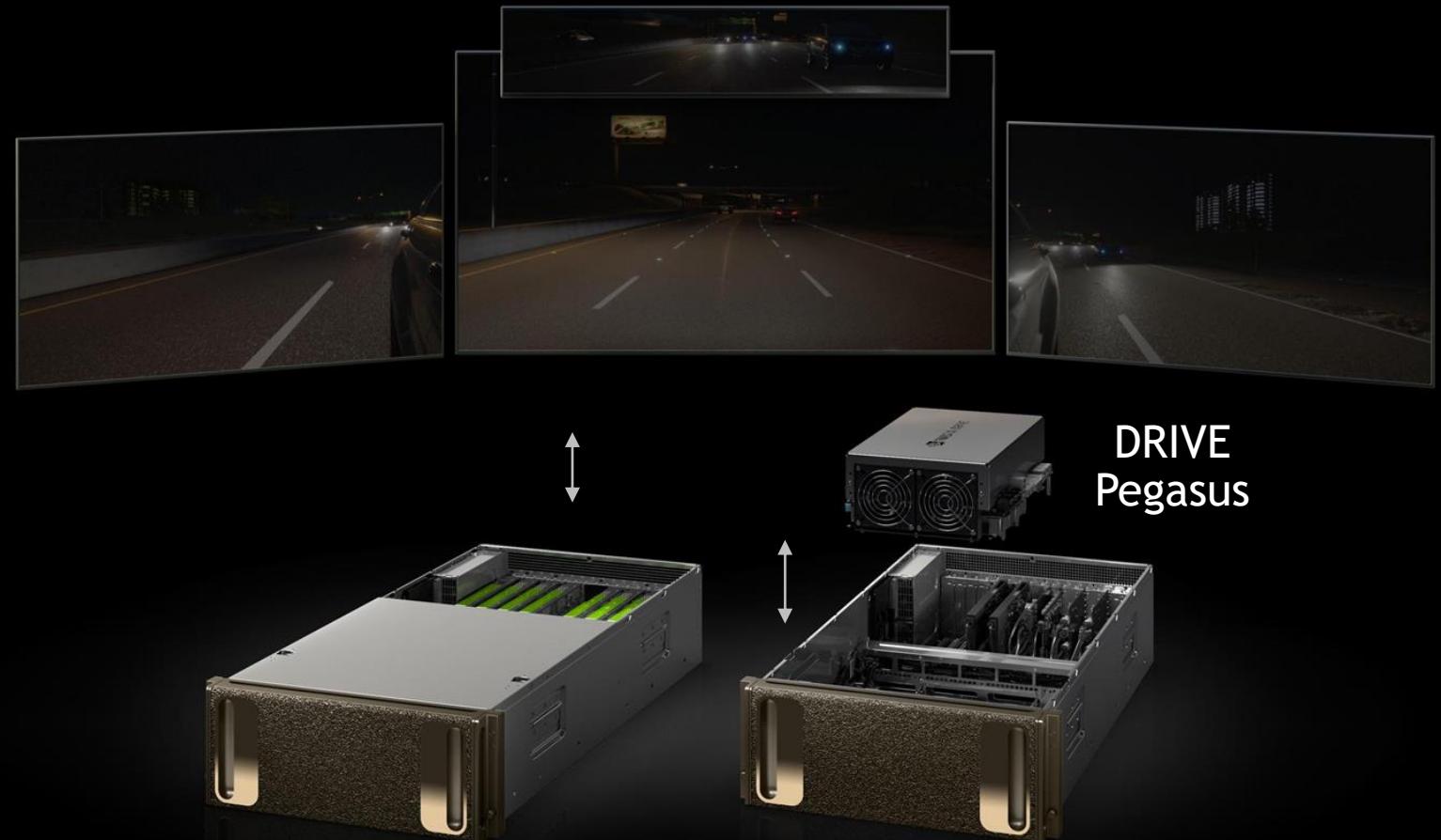
Virtual Reality AV Simulator

Same Architecture as DRIVE Computer

Simulate Rare and Difficult Conditions,

Recreate Scenarios, Run Regression Tests,
Drive Billions of Virtual Miles

10,000 Constellations Drive 3B Miles per Year



NVIDIA DRIVE SIM AND CONSTELLATION

AV VALIDATION SYSTEM

Virtual Reality AV Simulator

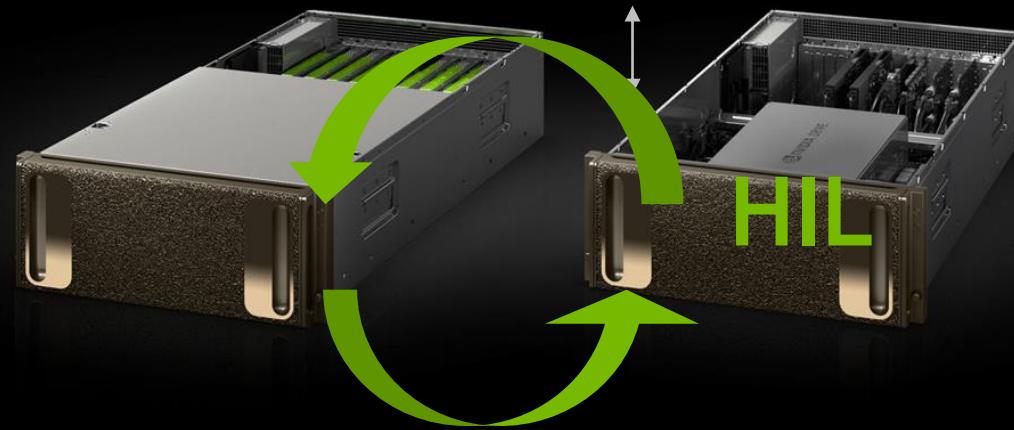
Same Architecture as DRIVE Computer

Simulate Rare and Difficult Conditions,
Recreate Scenarios, Run Regression Tests,
Drive Billions of Virtual Miles

10,000 Constellations Drive 3B Miles per Year



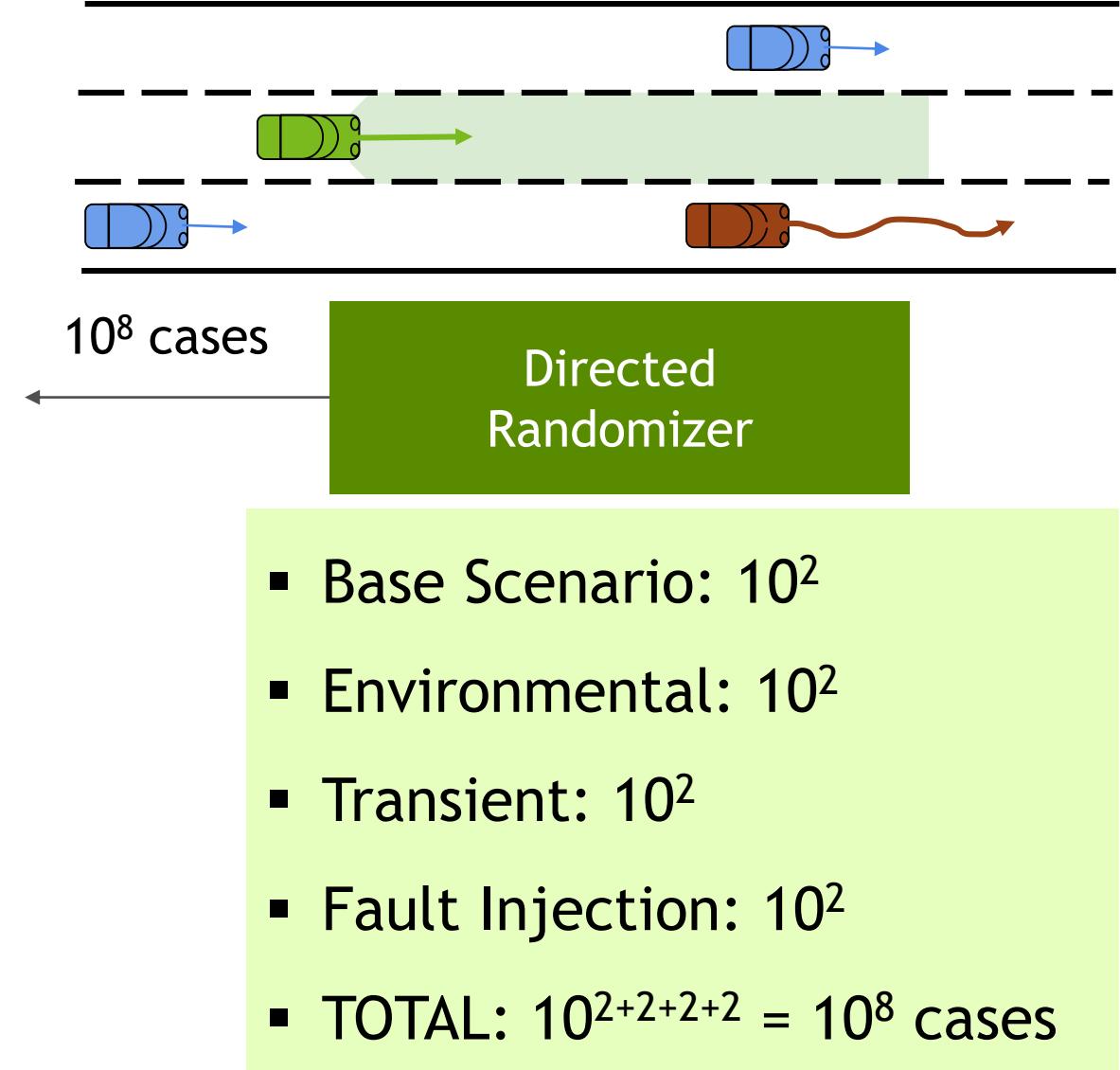
- Autonomous Vehicle Responses



- 8 Camera signals over GMSL2
- Radar and LIDAR signals over 1 Gbit Ether

シミュレーションのケース規模

- Simulation testing is scenario-based, and each scenario will have multiple trials to test with different directed random parameters.
- A trial is a single test of a scenario under a single set of randomized parameters.
- Figure shows the estimated base scenarios and trials for directed randomized testing



SIMULATION SIZING

1. 100 environmental conditions - Such as rain, snow, sunlight, clouds, etc.
2. 100 transients - Such as an adversarial car moving to close off space during a lane change
3. 100 fault injections - Introduced via test code paths to trigger error handling routines

This implies **every base scenario has one million trials** (100x100x100).

Base scenarios	Trials	Running time per trial (seconds) ¹	Total time of testing (hours) ¹	TAT (hours)	DRIVE Constellation Systems ^{1,2,3}
10	1M	30	100k	100	1,000

1. Rounded numbers

2. Average uptime of 20 hours per day

3. DRIVE Constellation systems running at 1x speed



Because HIL runs on real-time,
 $[100k \text{ Hr total test time}] / [100 \text{ Hr/Constellation}] = 1,000 \text{ Constellations}$

RESIMULATION/REGRESSION TESTING

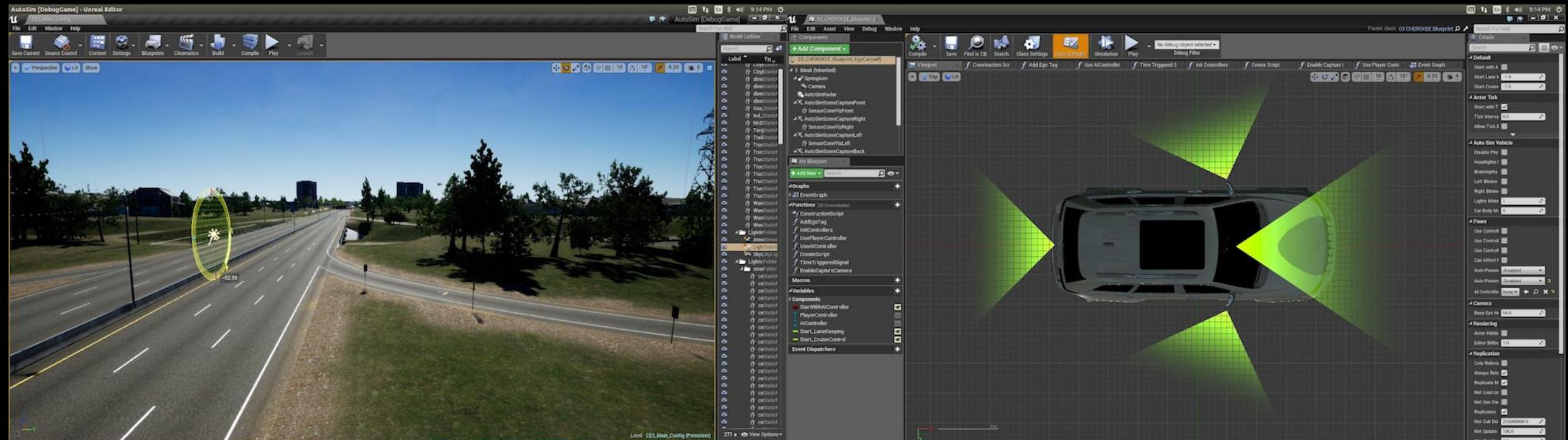
HIL Sizing

Regression testing plays back previously recorded sensor data to the driving stack.

Regression testing is used to test against labeled ground-truth, and to test against recorded and expected vehicle behavior (for example, we test AEB for zero false-positives using regression testing).

Raw Data (Hours)	TAT	# DRIVE Constellations
2000	100 hours	20
20000	200 hours	100

MULTI-SENSOR SIMULATION



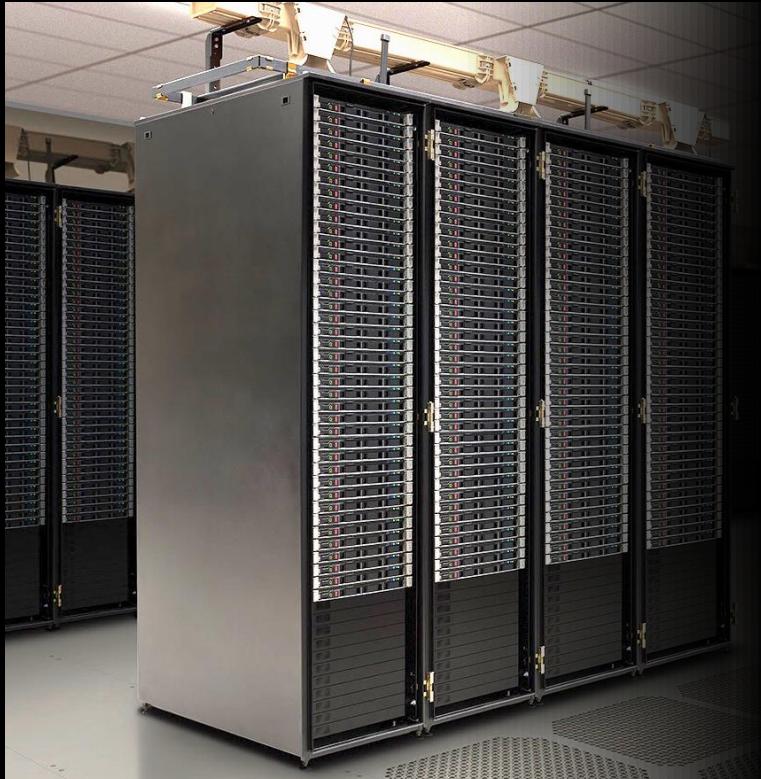
Setting Up The Environment

講演目次

- 1. FROM THE CARS: NVIDIA自動運転プラットフォームの最新状況**
 - 1) NVIDIA 高性能、高効率GPUの広範な応用
 - 2) ビッグGPUと同一のGPUアーキテクチャをもつ自動運転用SoC
 - 3) レベル3 - レベル5（完全自動運転）用SoC Xavier及びDRIVEプラットフォーム
- 2. TO THE CLOUD: 学習、シミュレーション、検証までのEnd to Endソリューション**
 - 1) 自動運転車開発に必須なEnd-to-Endソリューションとは
 - 2) 自動運転用AIモデルの学習
 - 3) 自動運転用シミュレーションとテスト
 - 4) 自動運転システム開発及び量産インフラストラクチャ

NVIDIA GPU PLATFORM SAVES MONEY

Enable Deep Learning Training at Scale



\$1.25M
in
servers

140 Skylake Gold CPU



One DGX-1

10x
TCO

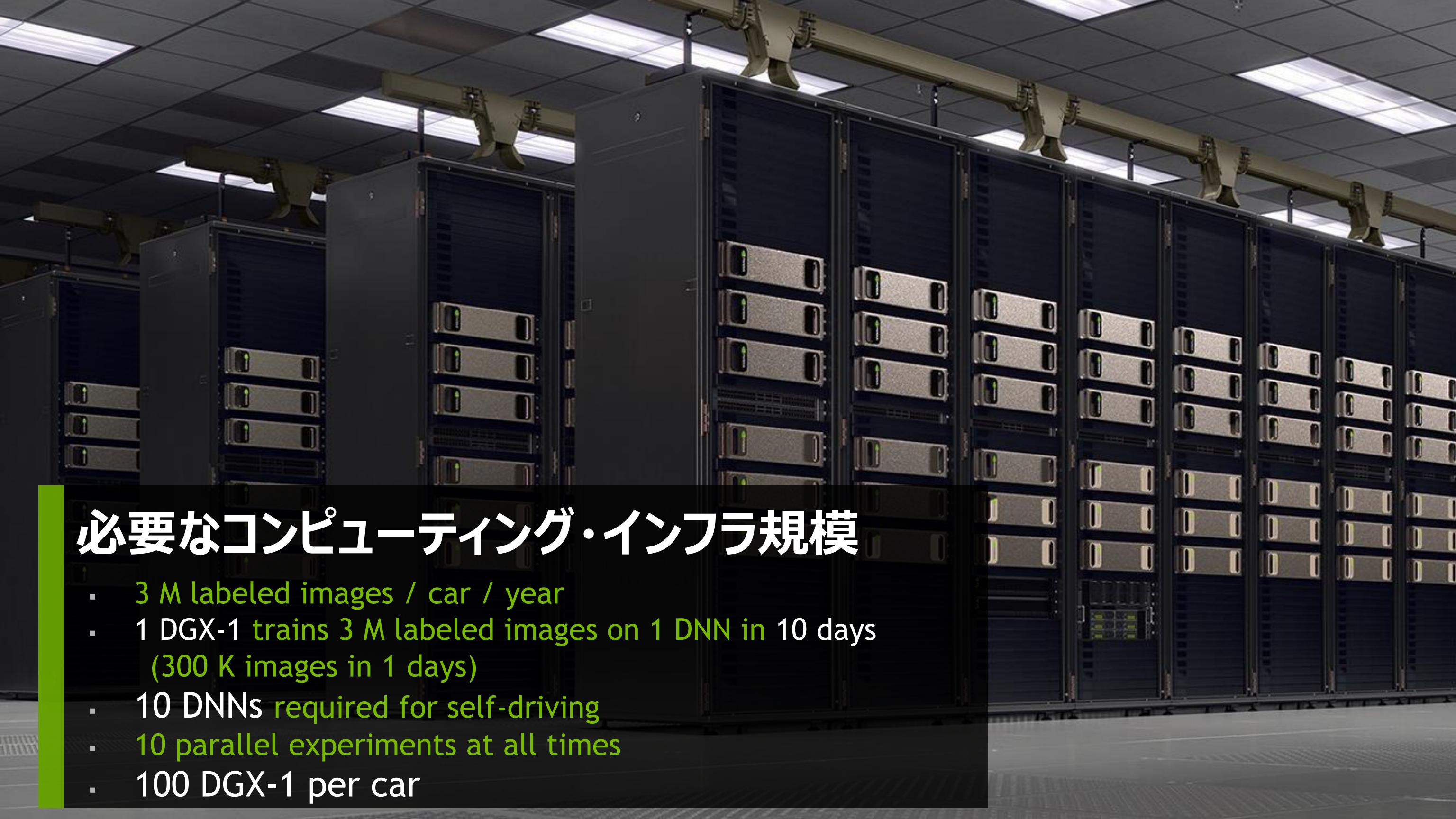
1/8
THE COST

10X
LESS SPACE

20X
LESS POWER

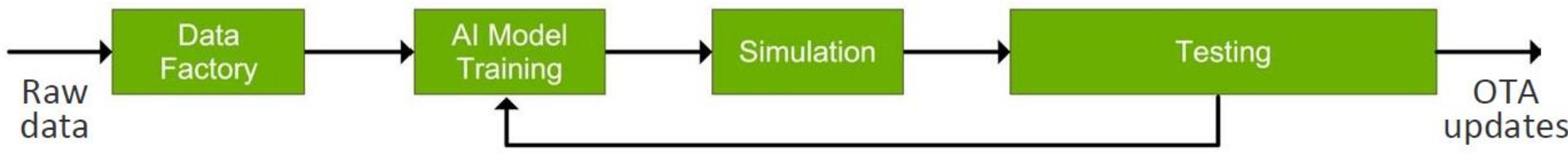
必要なコンピューティング・インフラ規模

- 3 M labeled images / car / year
- 1 DGX-1 trains 3 M labeled images on 1 DNN in 10 days
(300 K images in 1 days)
- 10 DNNs required for self-driving
- 10 parallel experiments at all times
- 100 DGX-1 per car



TSTADI WORKFLOW

(Training, Simulation, Testing for Autonomous Driving Infrastructure)



DGX POD
(9 DGX-1 Servers)



DGX POD



DRIVE
Constellation
POD



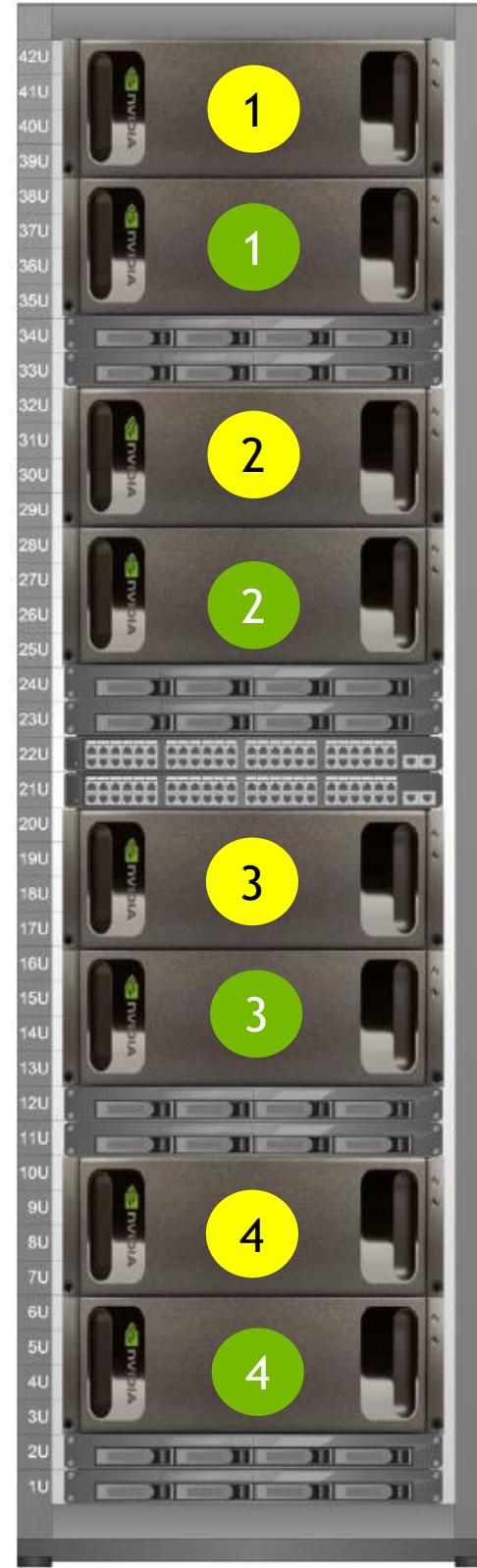
DRIVE
Constellation
POD



DGX POD

POD (Point of Delivery)
A module of computer,
network, storage and
apps that deliver
network service.

DRIVE CONSTELLATION POD



Four sets of DRIVE Constellation systems
 $(4 \times 8 = 32 \text{ RU})$

Eight storage servers
 $(8 \times 1 \text{ RU} = 8 \text{ RU})$

In-band switch (data)
(1 RU)

Out-of-band switch (management)
(1 RU)

Yellow circle: n runs NVIDIA DRIVE Sim to simulate sensors on GPU

Green circle: n contains NVIDIA DRIVE Pegasus AI Car Computer

NVIDIA DRIVE END TO END



THANK YOU