G2M RESEARCH

G2M Research Multi-Vendor Webinar: AI, Self-Driving Cars, and Advanced Storage

Tuesday July 21, 2020 (v2.0)

Webinar Agenda

9:00-9:05	Ground Rules and Webinar Topic Introduction (G2M Research)
9:06-9:35	Sponsoring Vendor presentations on topic (10 minute each)
9:36-9:46	Key Question 1 (2-minute question; 3 minutes response per vendor)
9:47-9:49	Audience Survey 1 (1 minutes)
9:49-9:59	Key Question 2 (2-minute question; 3 minutes response per vendor)
10:00-10:01	Audience Survey 2 (2 minutes)
10:02-10:12	Key Question 3 (2-minute question; 3 minutes response per vendor)
10:13-10:23	Audience Q&A (11 minutes)
10:24-10:25	Wrap-Up

G2M Research Introduction and Ground Rules

Mike Heumann

Managing Partner, G2M Research

Panelists

Rob Davis
VP, Storage Technology
www.nvidia.com

Alexander Noack
Head of Automotive Electronics
www.b-plus.com

Shailesh Manjrekar
Head of AI, Strategic Alliances
www.weka.io

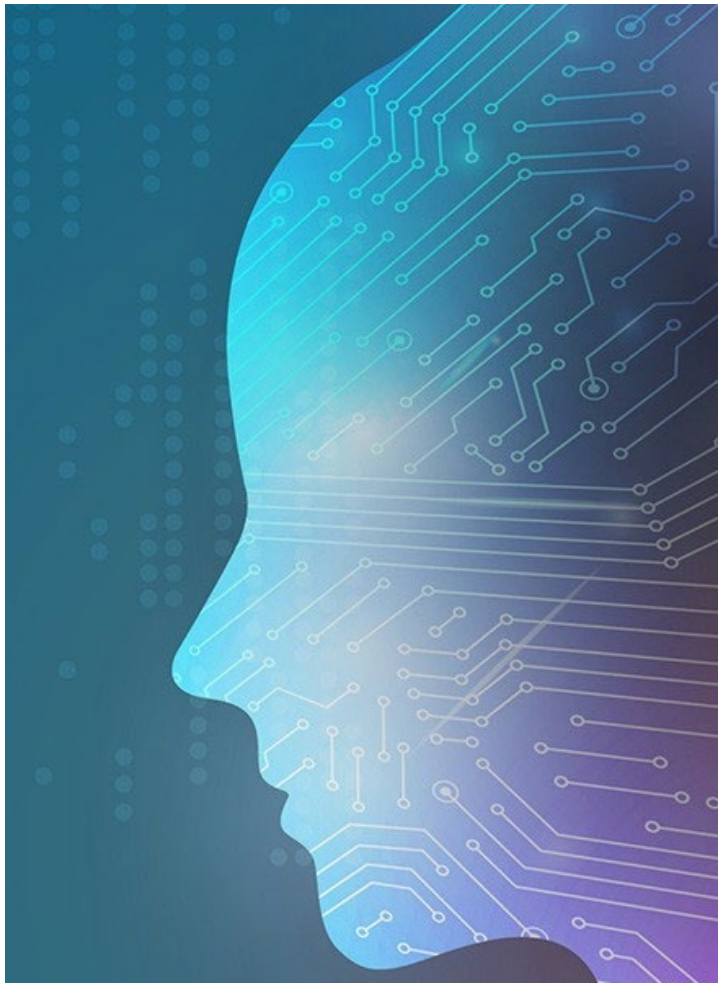
Mike Heumann
Managing Partner
www.g2minc.com

Autonomous Vehicles Success Based on AI/ML

- Safety is Critical for Self-Driving Cars and Autonomous Vehicles
  - Tracking objects, pedestrians, and other vehicles is job 1 for self-driving cars
  - Methods such as high-definition maps, path planning, and SLAM drive these capabilities
  - Computer vision, RADAR/LIDAR, and similar technologies support these methods
- Artificial Intelligence (AI) and Machine Learning (ML) turn this data into vehicle commands
  - Huge amounts of data is required for both training and vehicle operation

How Much Data Do Self-Driving Cars Need?

- Training Demands the Most Data
  - One instrumented car can over 30TB of data per day
  - A fleet of 100 vehicles could generate 780PB of raw data
  - Even with extensive sampling and pre-processing, the data volume will be 100TB-1PB, which would be spread across hundreds of compute nodes during training
- During driving, storage is used for both operation and for data collection
  - Data collection likely continues during driving, but it would have to be highly selective (30TB/day is a LOT)
  - Running out of storage has been an issue for Tesla because of this specific issue



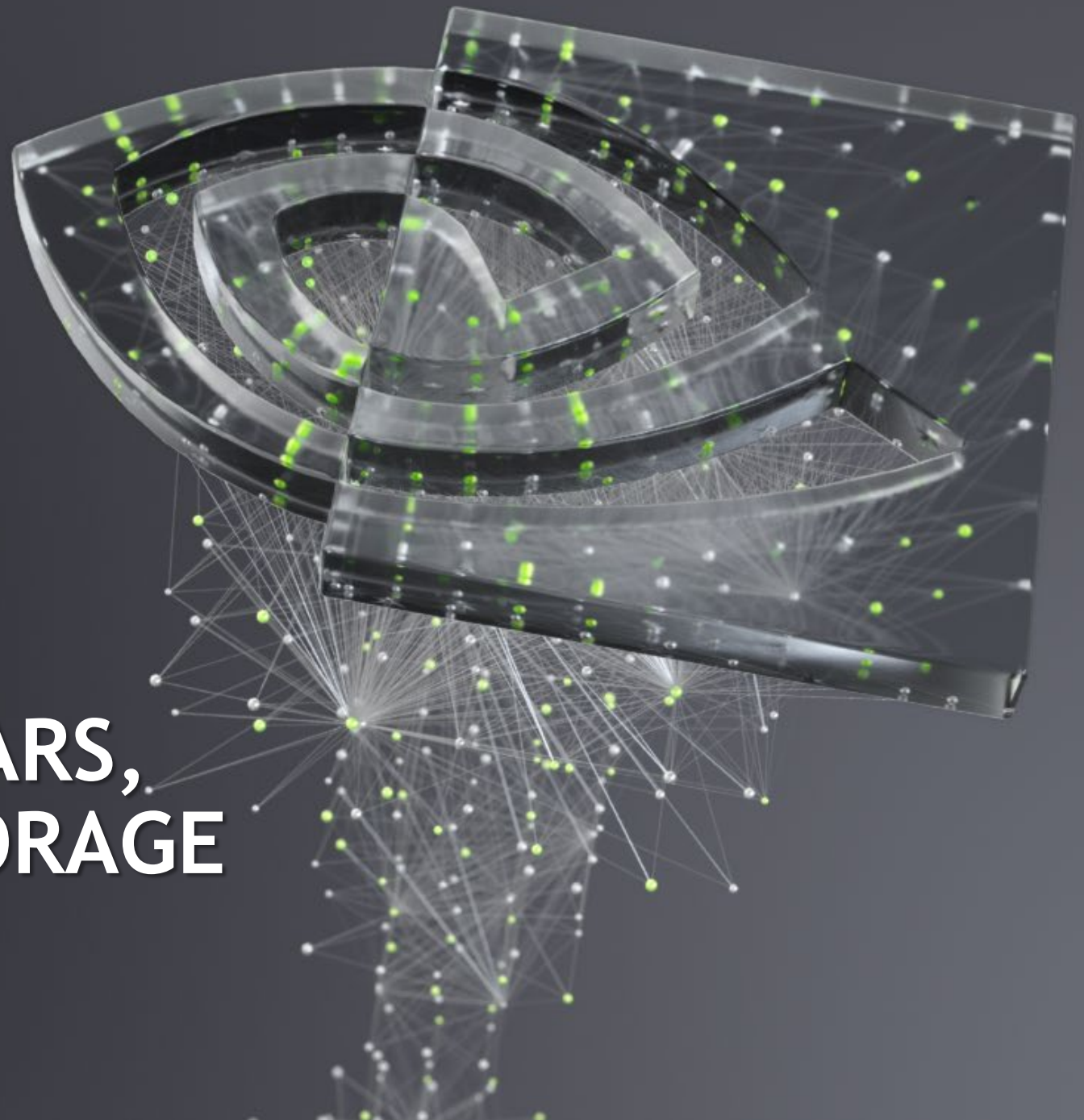
G2M RESEARCH

NVIDIA

- Rob Davis

Vice President Storage Technology, Nvidia Worldwide Networking Business Unit

www.nvidia.com



NVIDIA

AI, SELF-DRIVING CARS, AND ADVANCED STORAGE

Rob Davis, July 2020

NVIDIA AND SELF-DRIVING CARS

NVIDIA

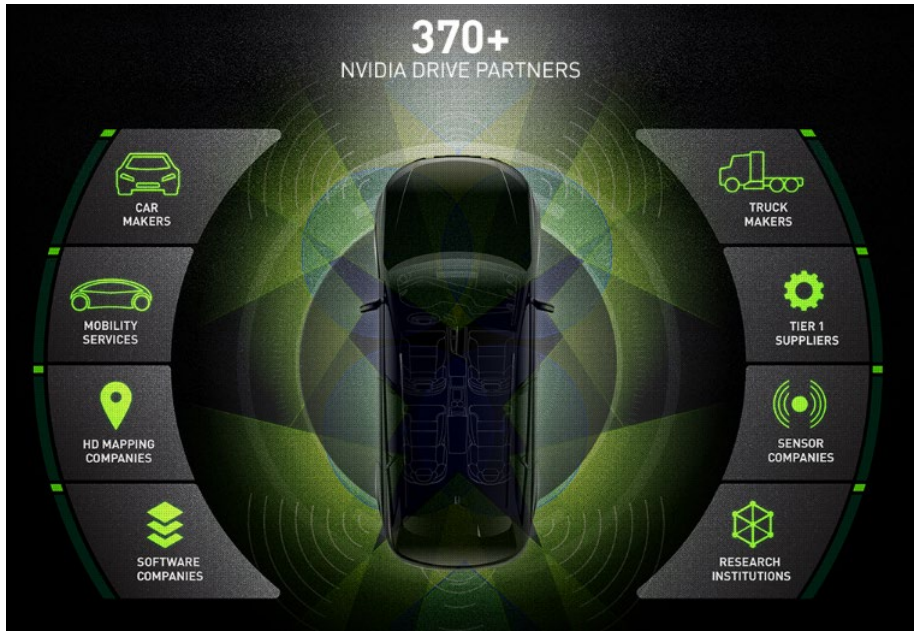
SELF-DRIVING CARS PRODUCTS SOLUTIONS DRIVE LABS PARTNERS FOR DEVELOPERS

NVIDIA DRIVE Infrastructure DNN TRAINING SIMULATION/REPLAY CONTACT US

NVIDIA DRIVE INFRASTRUCTURE

End-to-End Solutions for Training, Development, and Validation of Autonomous Vehicles.

https://www.nvidia.com/en-gb/self-driving-cars/infrastructure/

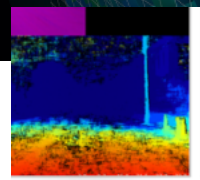


https://www.nvidia.com/en-us/self-driving-cars/partners/

NVIDIA DRIVE SOFTWARE

DRIVE OS

The foundation of the DRIVE Software stack, DRIVE OS is the first safe operating system for accelerated computing. It includes NvMedia for sensor input processing, NVIDIA CUDA® libraries for efficient parallel computing implementations, NVIDIA TensorRT™ for real-time AI inference, and other developer tools and modules to access hardware engines.



DriveWorks

The NVIDIA DriveWorks® SDK provides middleware functions on top of DRIVE OS that are fundamental to autonomous vehicle development. These consist of the sensor abstraction layer (SAL) and sensor plugins, data recorder, vehicle I/O support, and a deep neural network (DNN) framework. It's modular, open, and designed to be compliant with automotive industry software standards.

DRIVE AV

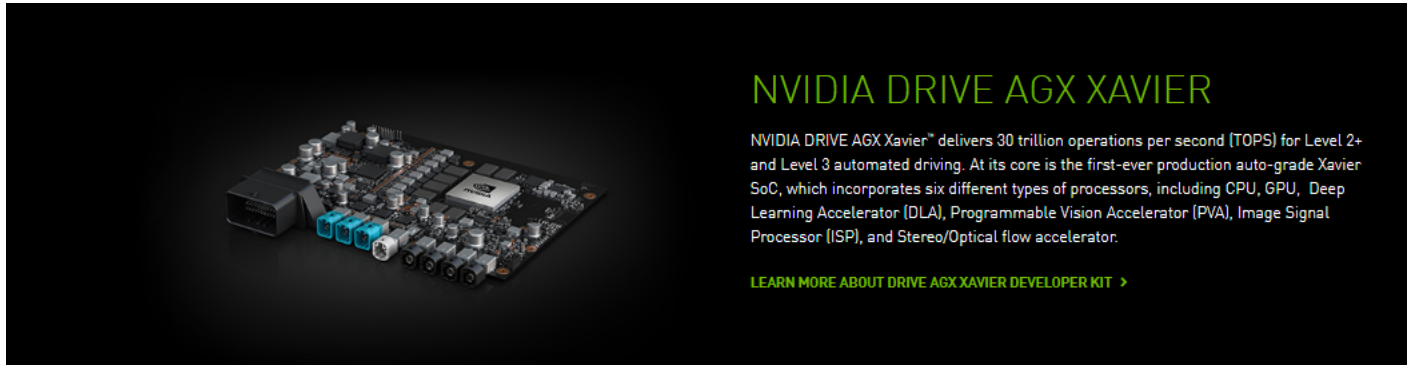
The DRIVE AV software stack contains the perception, mapping, and planning layers, as well as diverse DNNs trained on high-quality real-world driving data. These rich perception outputs can be used for both autonomous driving and mapping. In the planning and control layer, the NVIDIA Safety Force Field™ computational module checks the actions of the primary planning system to shield the vehicle from collisions.

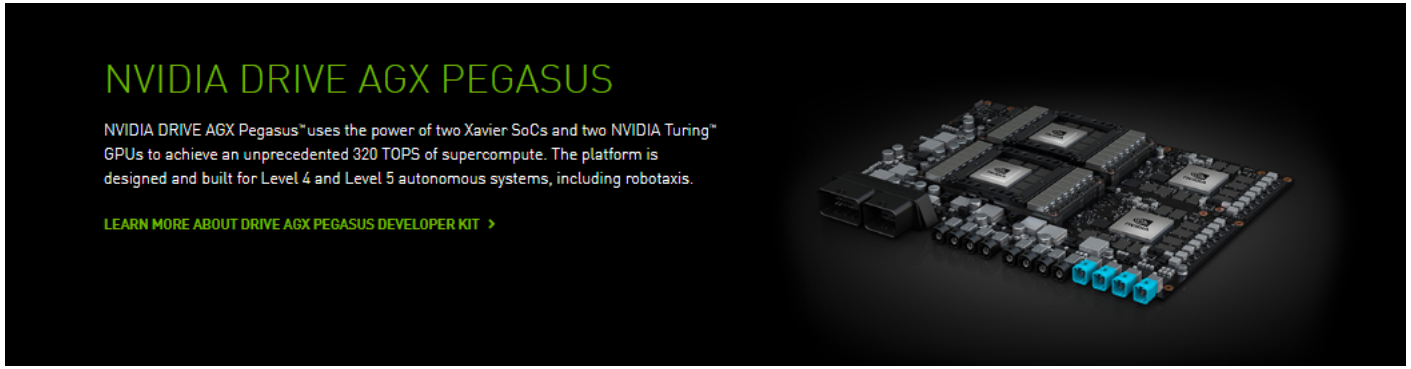
DRIVE IX

DRIVE IX is an open software platform that delivers interior sensing for innovative AI cockpit solutions. It provides perception applications to access features and DNNs for advanced driver and occupant monitoring, AR/VR visualization, and natural language interactions between the vehicle and passengers. DRIVE IX also enables fleet monitoring, event-based in-cabin recording, and remote operation capabilities in trucks.

https://www.nvidia.com/en-us/self-driving-cars/drive-platform/software/

NVIDIA DRIVE HARDWARE – AGX





https://www.nvidia.com/en-us/self-driving-cars/drive-platform/hardware/

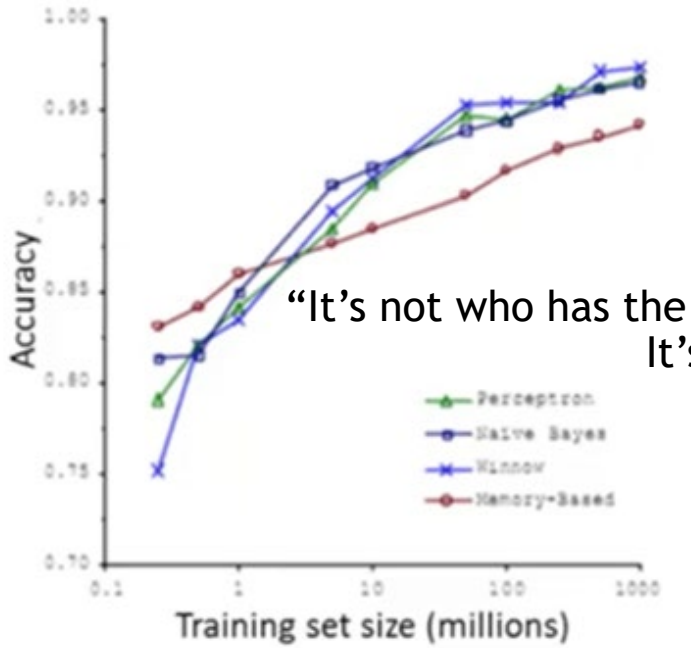
Forbes Billionaires Innovation Leadership Money Business Small Business Lifestyle

Don't Forget About Storage When Planning For AI And ML

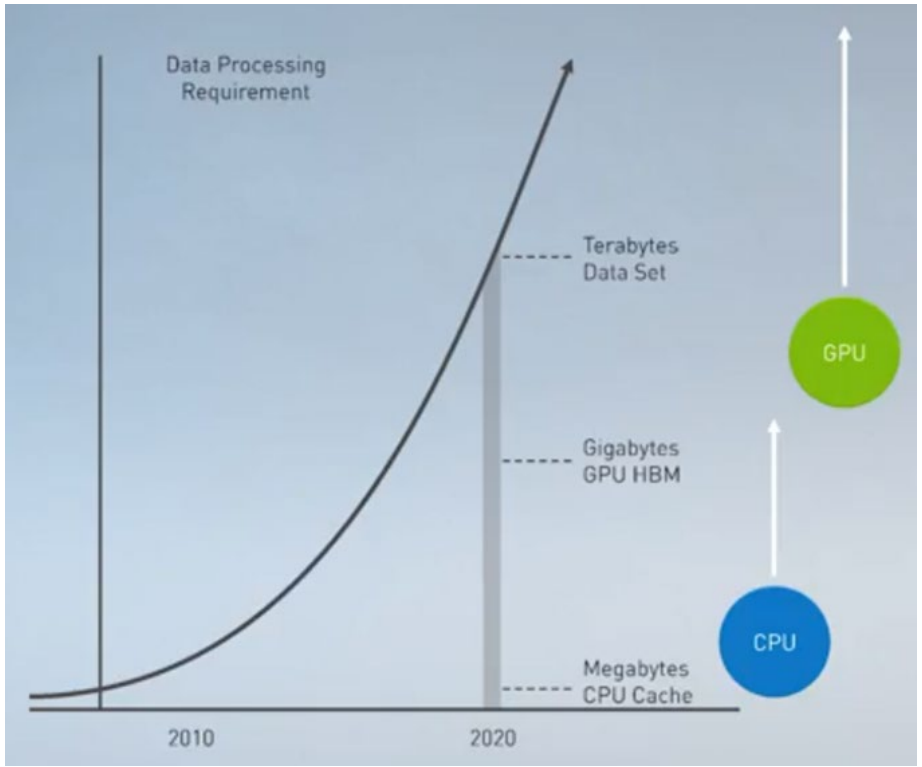
Patrick Moorhead Contributor ⓘ

Enterprise & Cloud

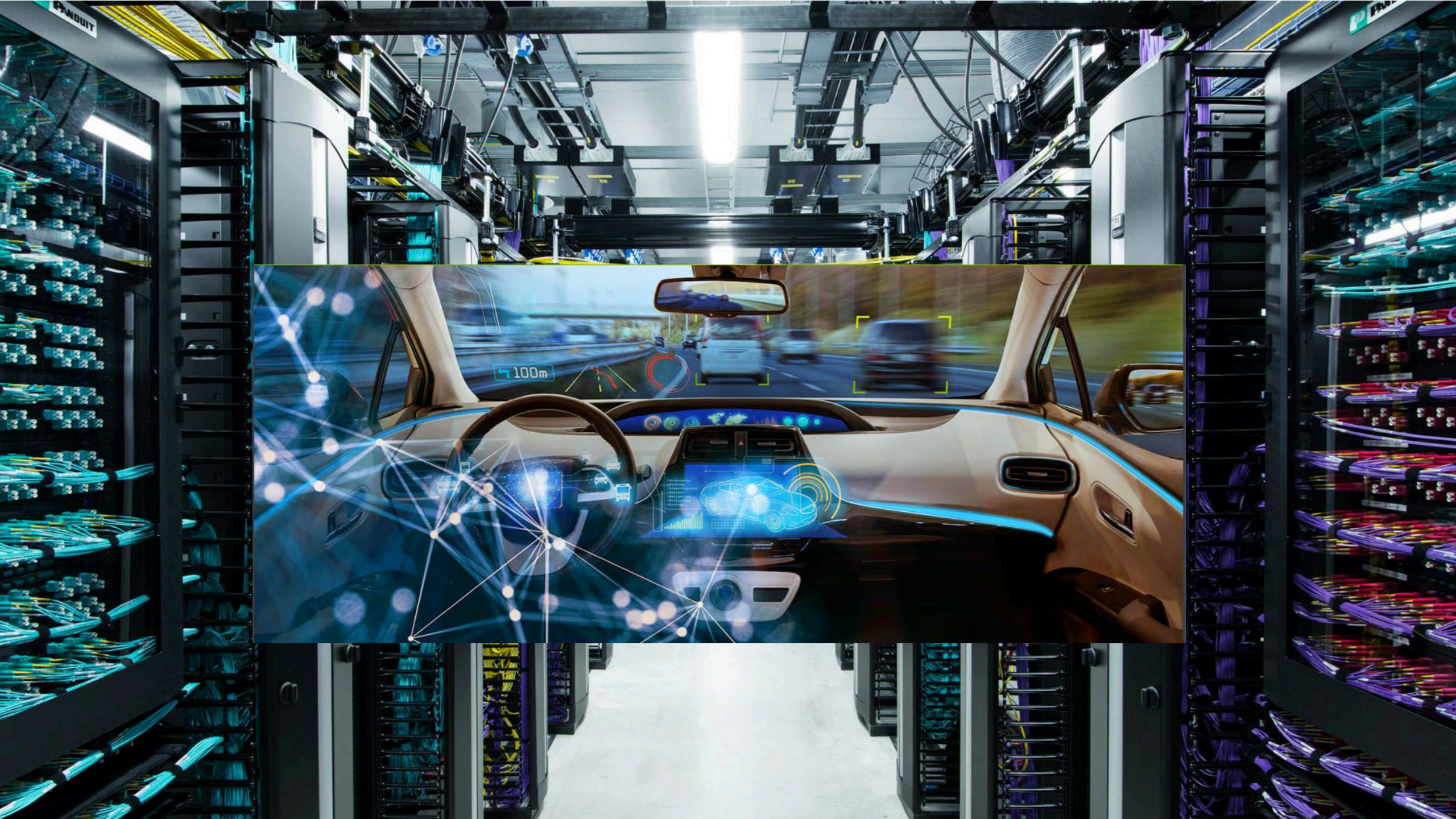
I write about disruptive companies, technologies and usage models.



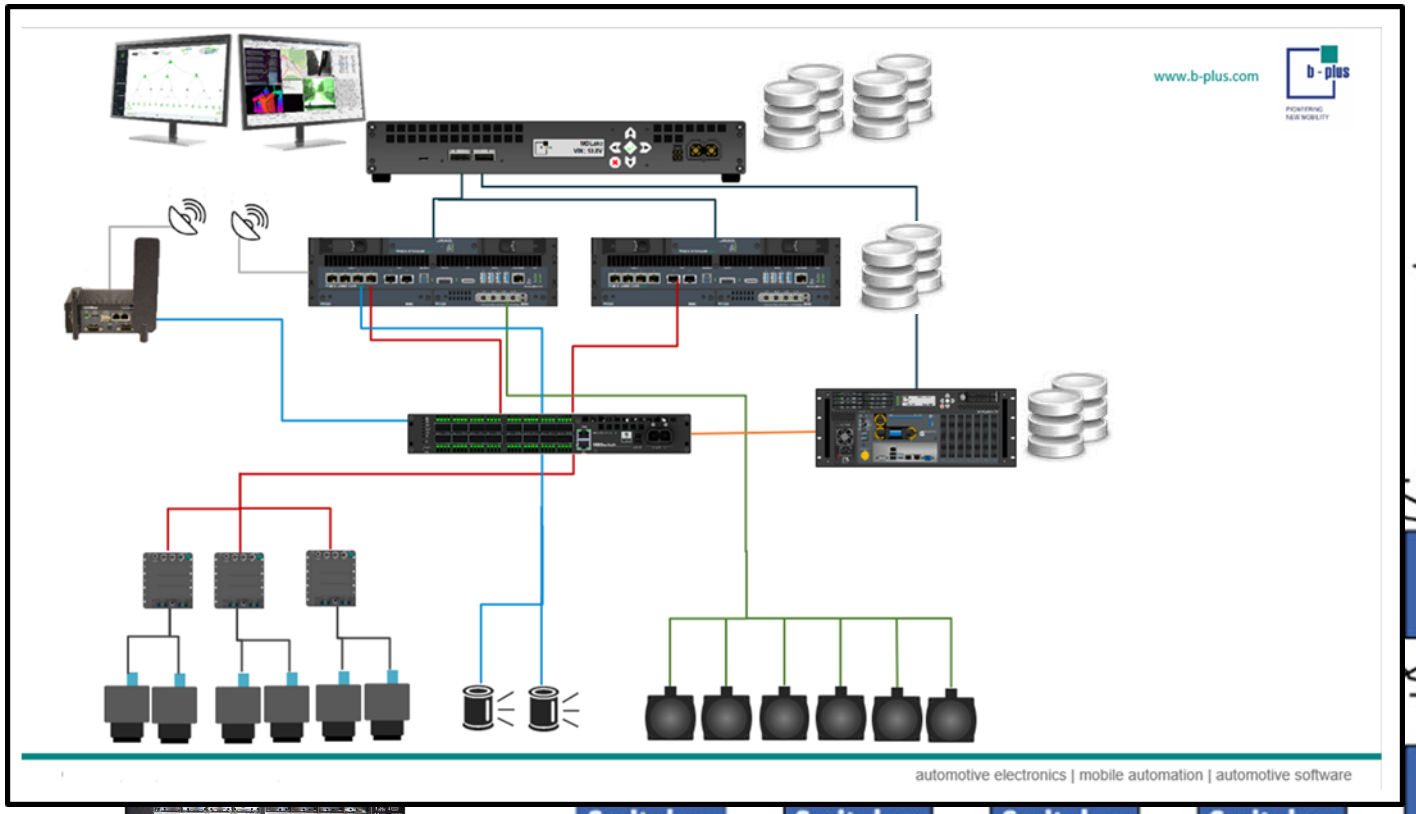
"It's not who has the best algorithm that wins.
It's who has the most data."
Andrew Ng



NVIDIA.


PANDUIT
100m

STORAGE NETWORK TOPOLOGY



Key
CR=Core Router
AS=Aggregation Switch
ToR=Top-of-Rack Switch

AS

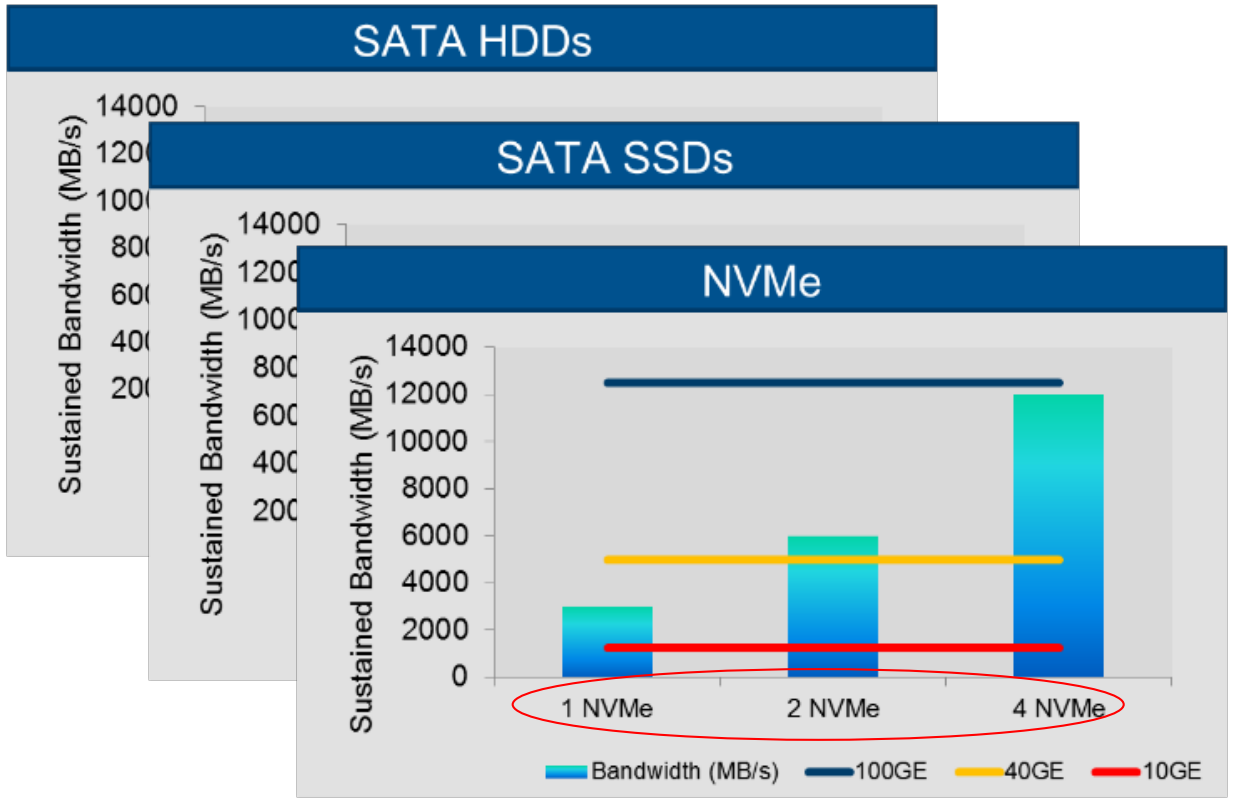
Switches
Switches
Switches
Switches

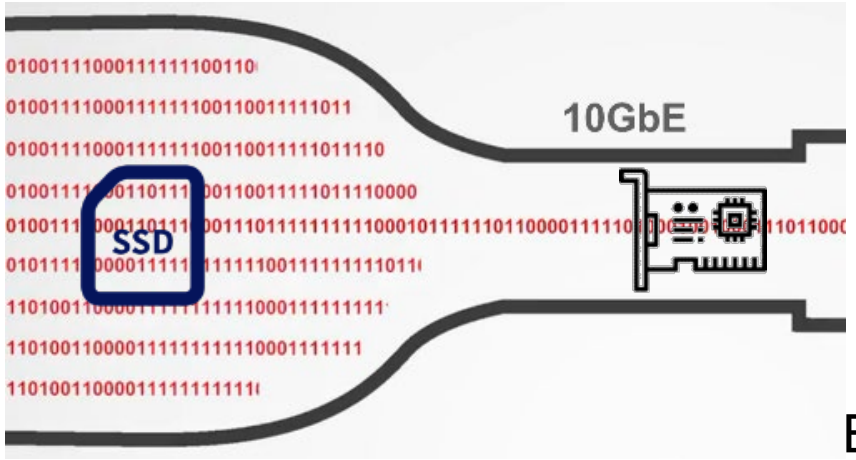
ToR Switches
ToR Switches

Servers
Servers
Servers
Servers

GPU
GPU

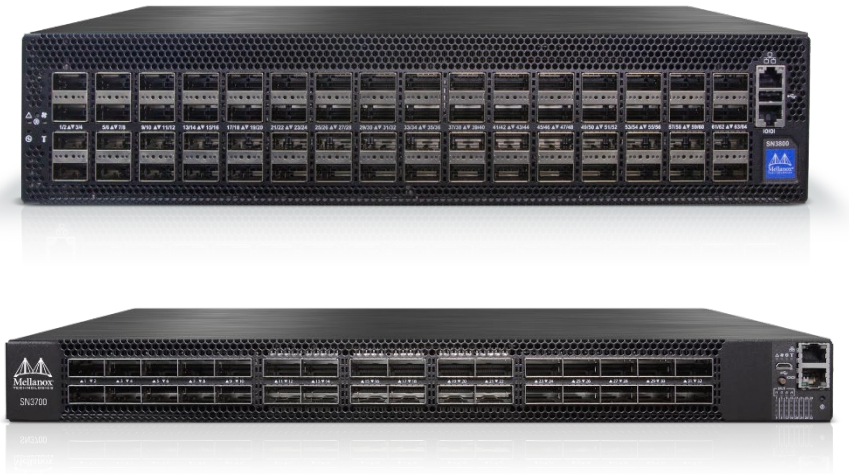
SSDs Have Changed Networked Storage





SSDs move the Bottleneck from the Disk to the Network

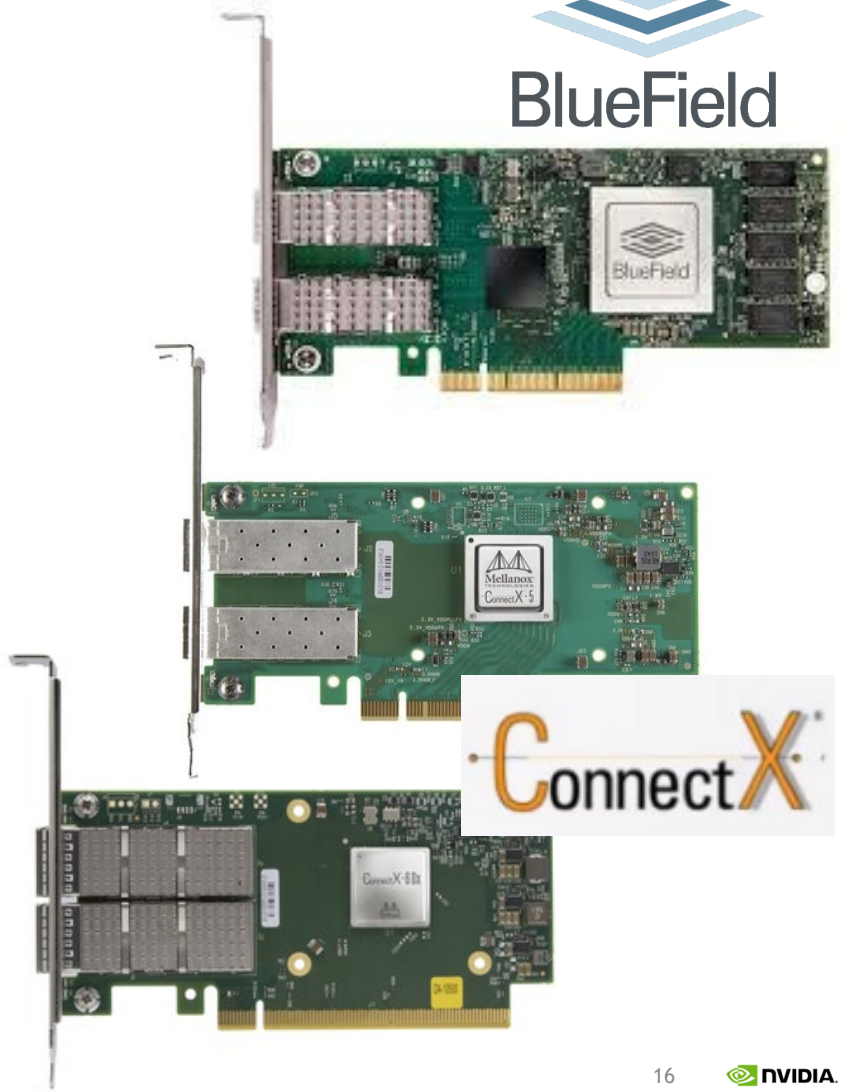
Faster Network Products from Nvidia Solves Half the Network Bottle Neck Problem…



Spectrum™

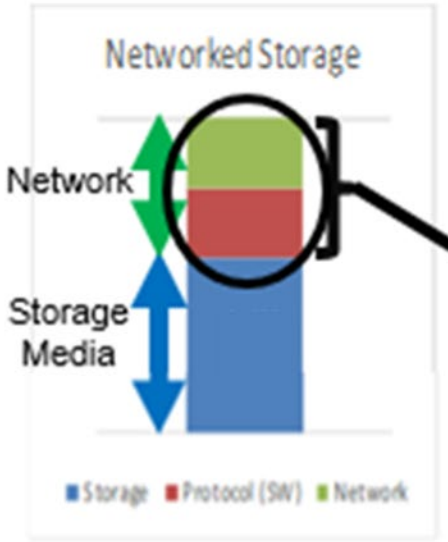
End to End 200Gb/s

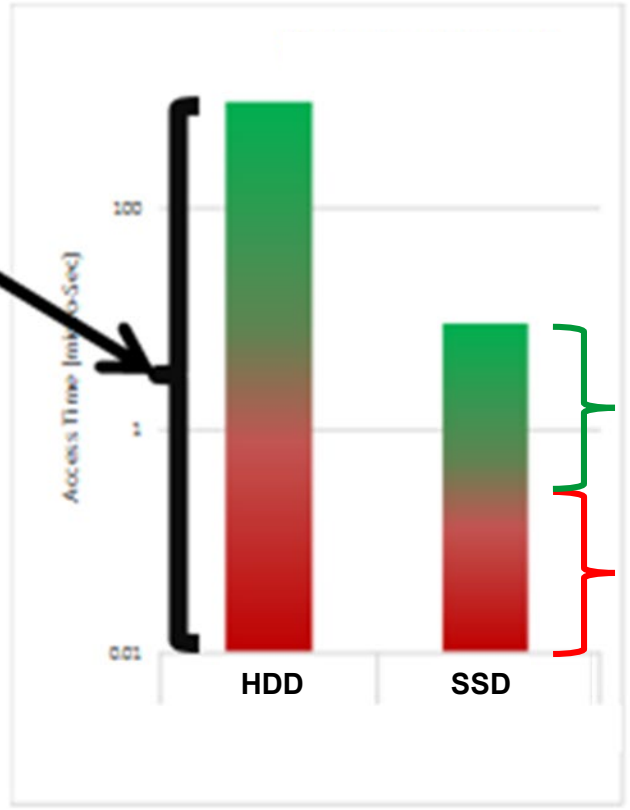
BlueField



NVIDIA

Faster Protocols from Nvidia Solves the Other Half

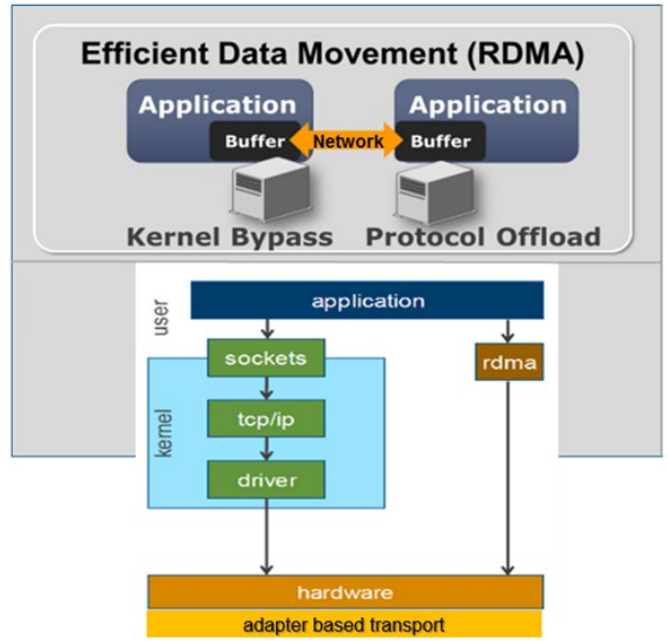




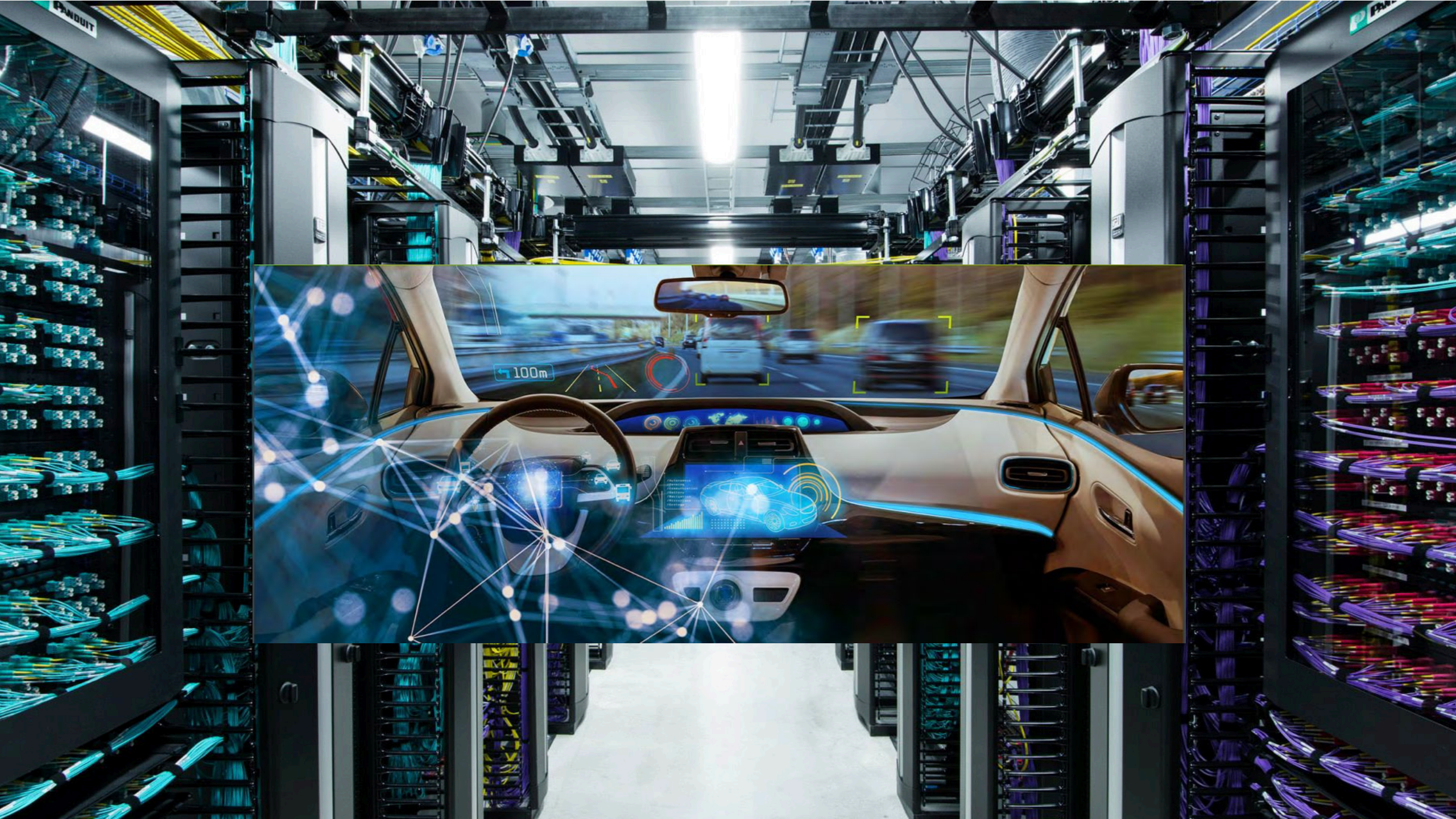
Faster Network

Faster Protocol

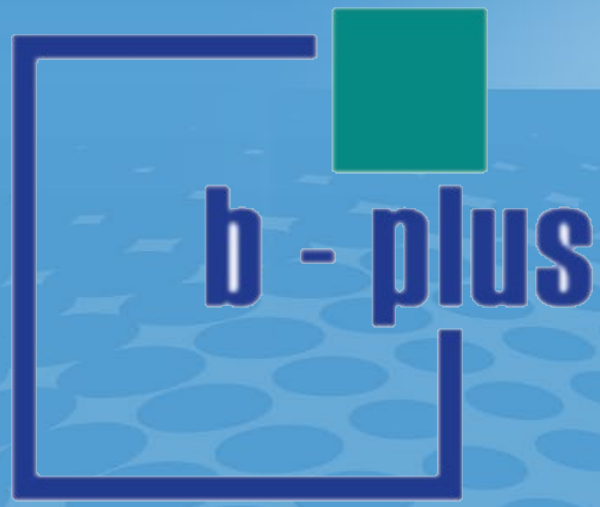
NVMe-oF



NVMe OVER FABRICS
POWERED BY RoCE


PANDUIT
100m

G2M RESEARCH



b-plus

Alexander Noack

Head of Automotive Electronics

www.b-plus.com

AI, Self-Driving Cars, and Advanced Storage

PIONEERING NEW MOBILTY.

Who we are

b-plus is partner for the automotive and automation industry.

We develop software and hardware solutions for applications such as driver assistance, autonomous driving, connected car and control of mobile machines.

Tools for autonomous driving

ADAS/AD Multi-Sensor Recording, Visualization and Recording framework, Hardware in the Loop raw data simulation

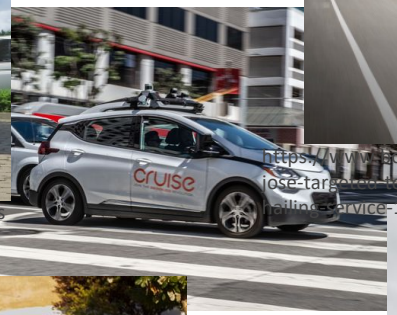
b-plus in numbers

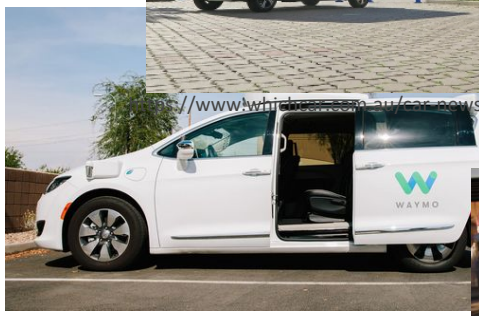
Founded: 1996
Employees: 200+

Locations

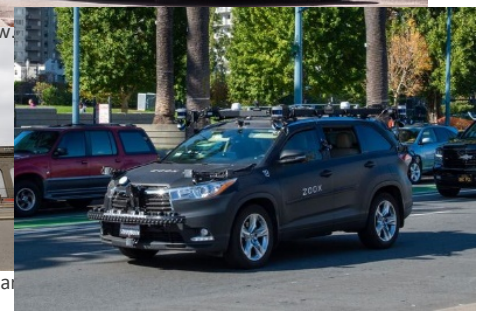
Deggendorf (Headquarters)
Regensburg
Lindau

AUTONOMOUS CARS

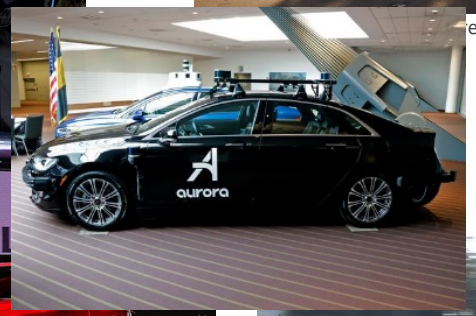


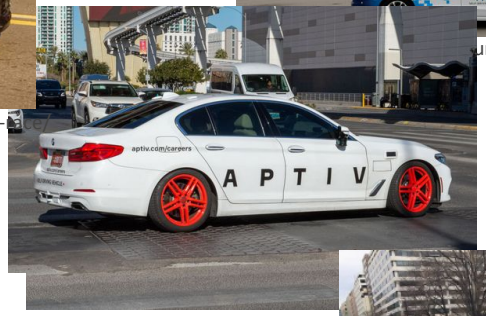


https://www.bloomberg.com/features/2020-self-driving



https://www.bloomberg.com/features/2020-self-driving-car-race/





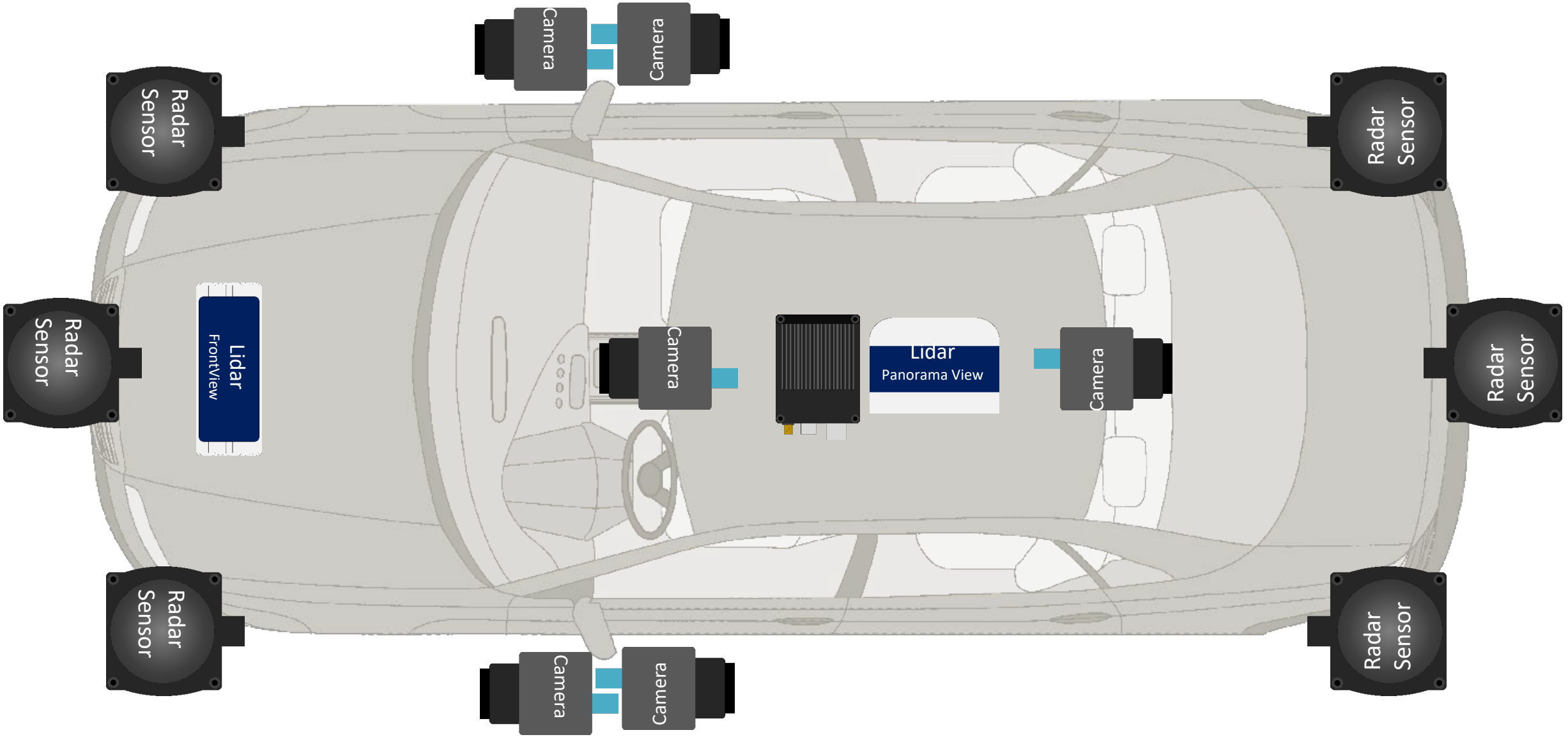
https://www.bloomberg.com/features/2020-self-driving-car-race/

https://www.bloomberg.com/features/2020-self-driving-car-race/

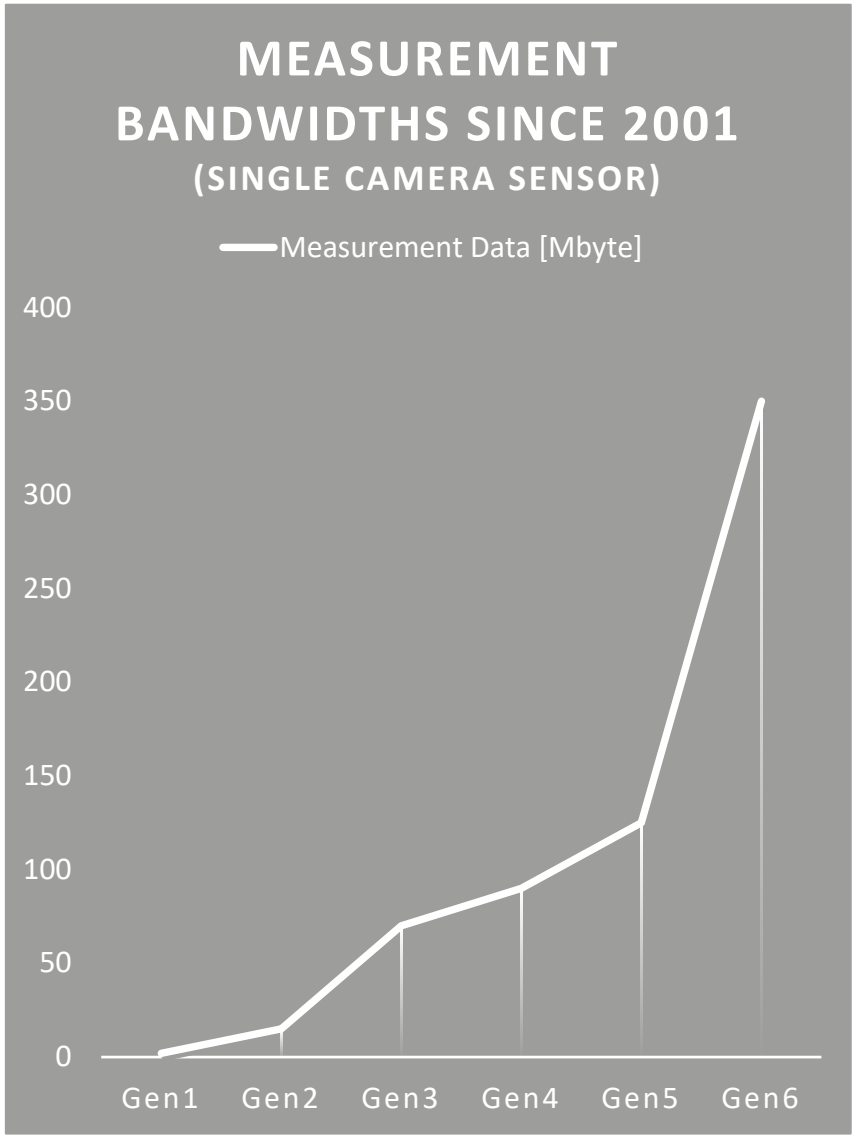
https://www.zf.com/site/magazine/en/articles_16141.html

https://www.bloomberg.com/features/2020-self-driving-car-race/

SAMPLE CONFIGURATION



BANDWIDTHS SEEN FROM THE PAST TO NOW



Terabytes generated per vehicle

Data rate examples:

1x Camera (Full HD/RAW12/40fps) ≈ 120 MB/s

1x Radar ≈ 220 MB/s

Example setup:

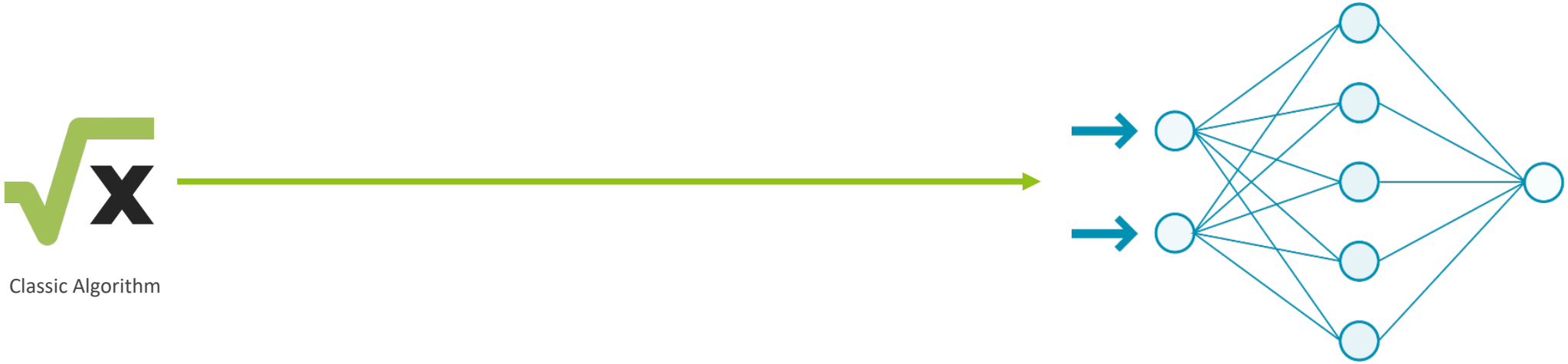
6 Cameras and 6 Radars

Complete vehicle RAW data ≈ 2.040 MB/s

≈ 58 TB in a 8h test drive shift

Data driven development (Neural networks) needs accurate data

- Therefore accurate data is the key to a successful function
- How accurate is my training data?
- Do I have frame drops?
- Do I have enough training data?



Challenges when dealing with multi source data

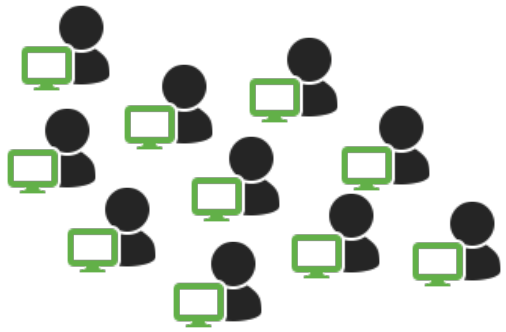
Various sources need to be synchronized and sorted accurate in time

Data must not be lost.

Corrupt data have to be marked for analysis and data replay/simulation.

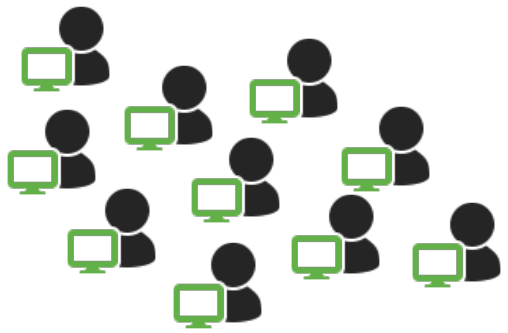

In-vehicle recording
Analysis of the sensor data
Replay of sensor data

Software / Algo Development



ECU Software

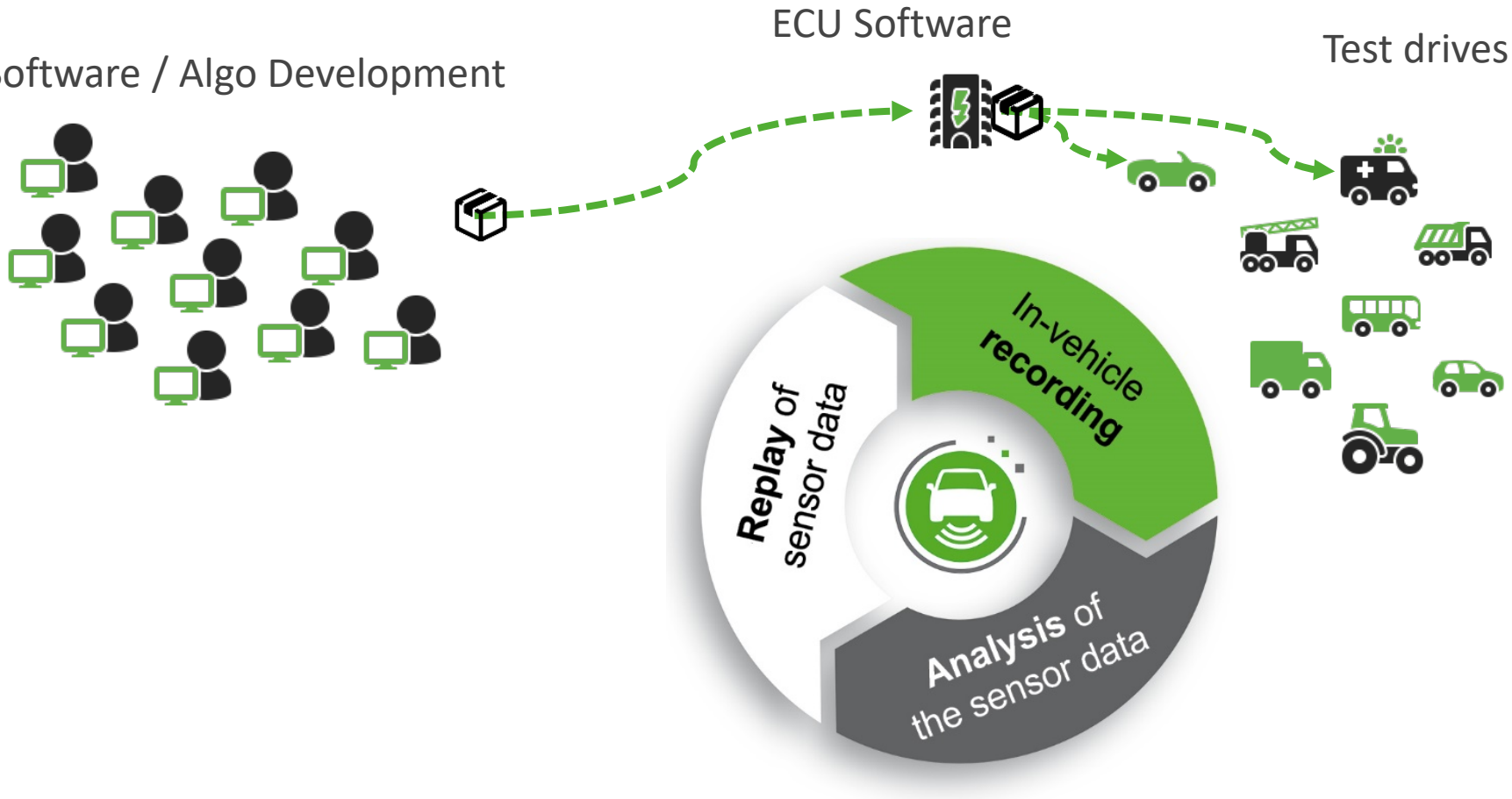
Software / Algo Development



Software / Algo Development

ECU Software

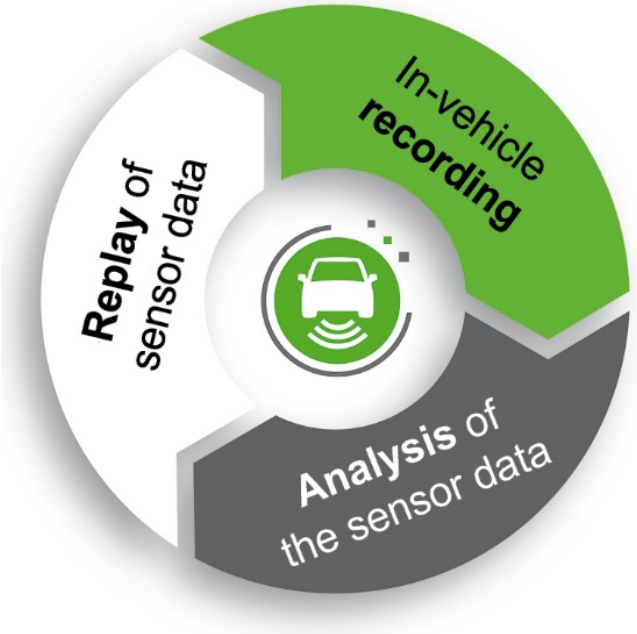
Test drives

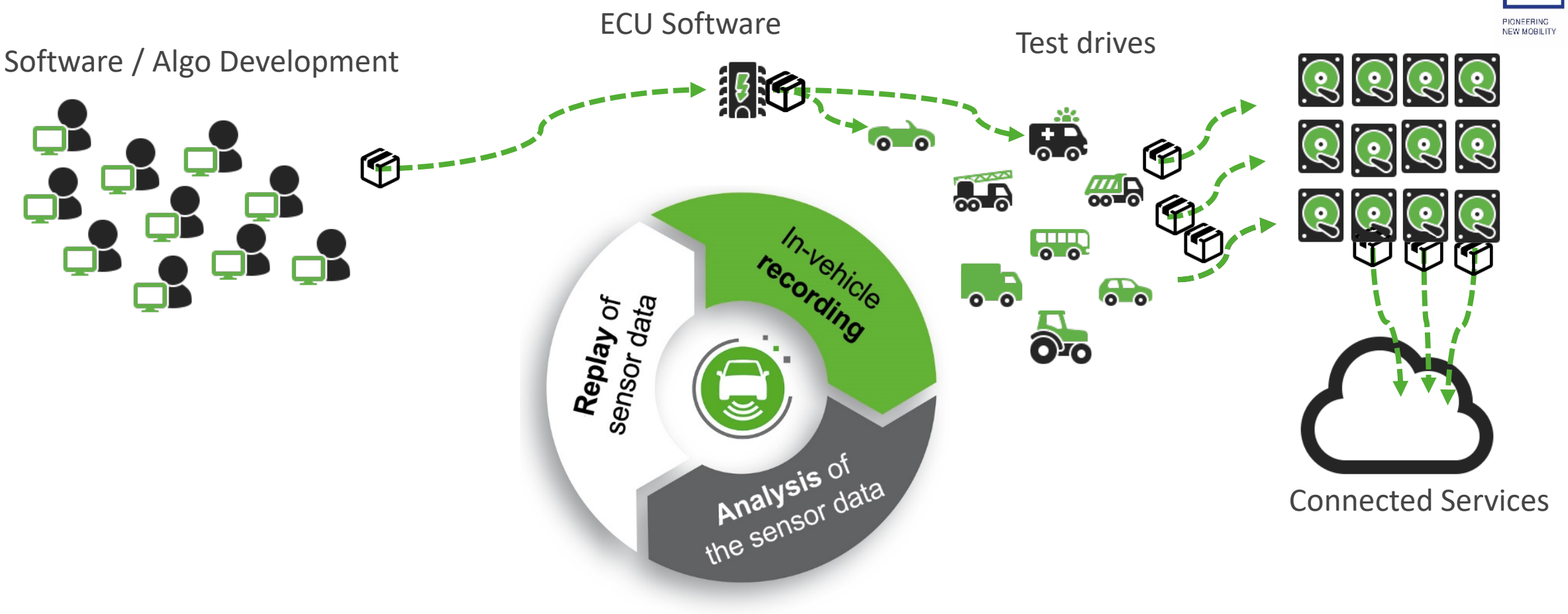


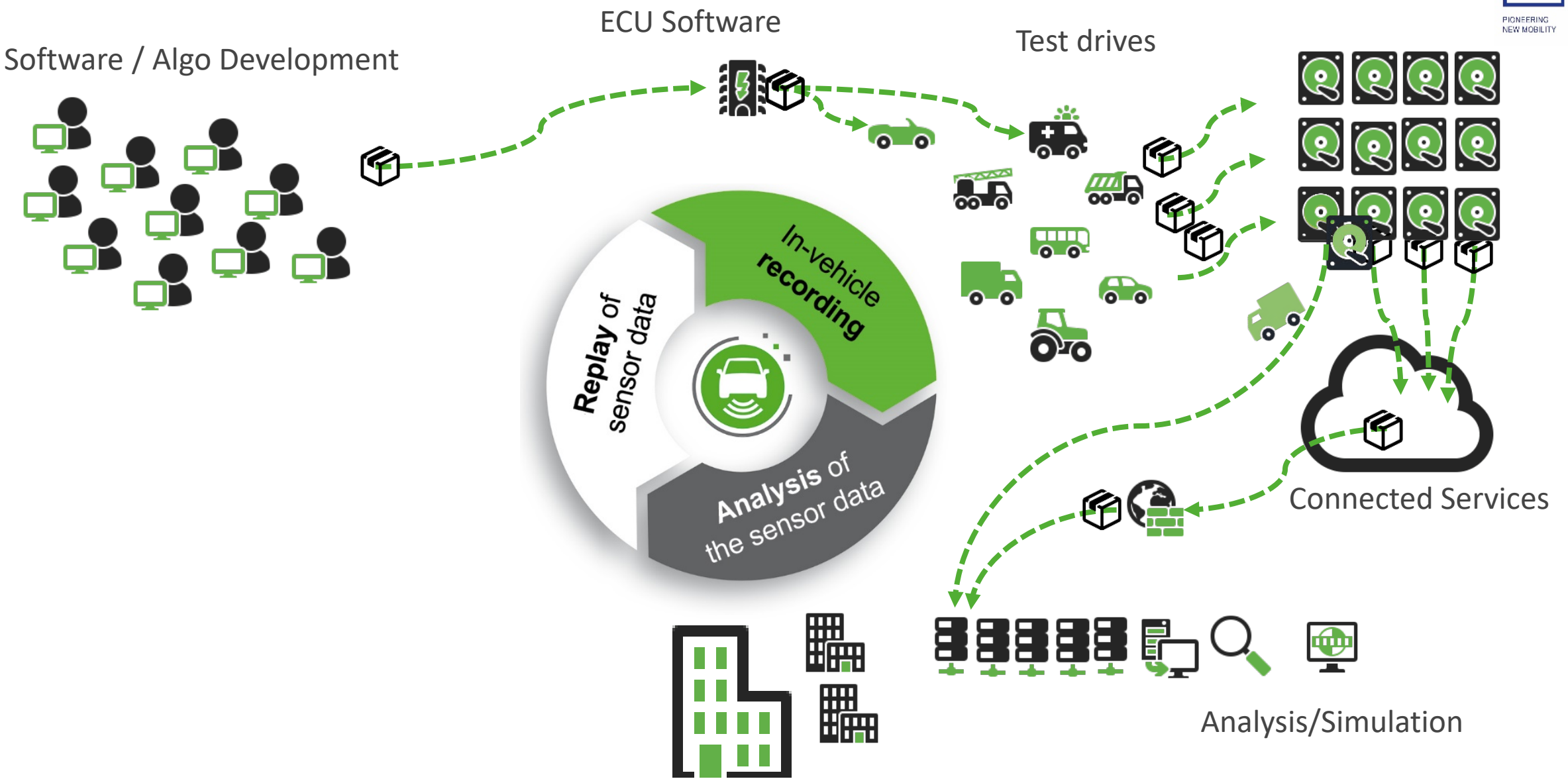
Software / Algo Development

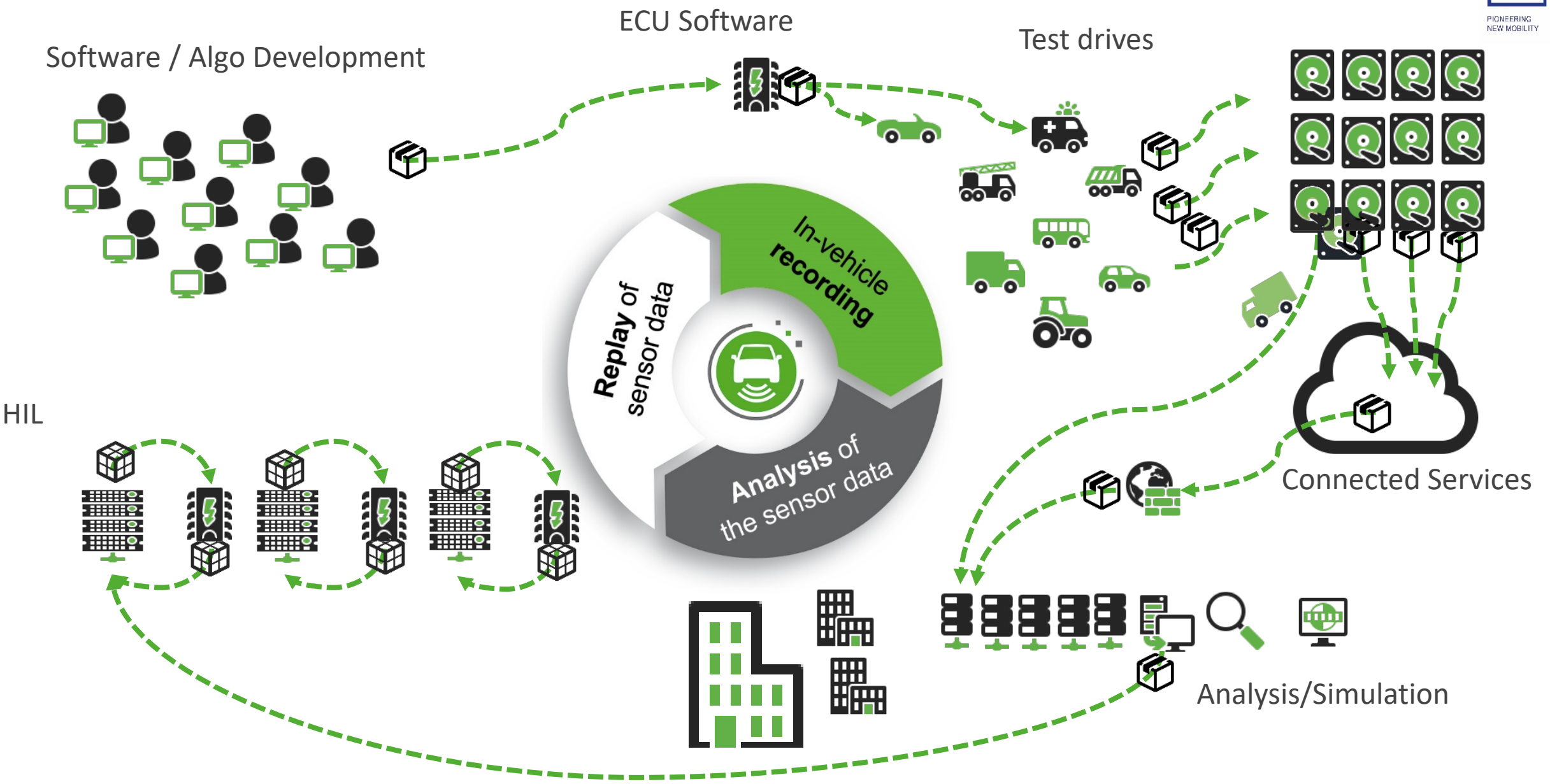
ECU Software

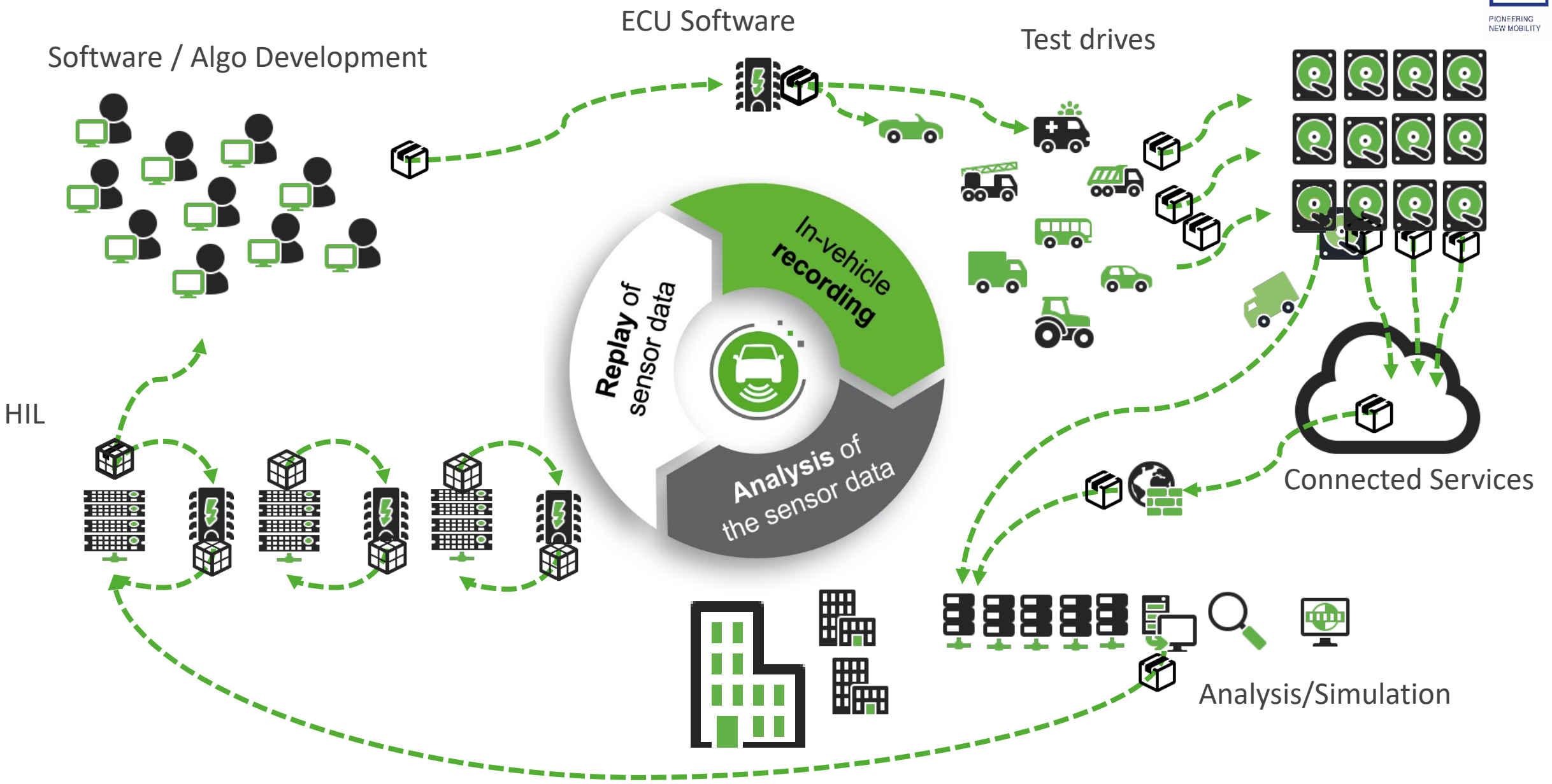
Test drives




Software / Algo Development
ECU Software
Test drives
In-vehicle recording
Analysis of the sensor data
Replay of sensor data
Connected Services


Software / Algo Development
ECU Software
Test drives
In-vehicle recording
Replay of sensor data
Analysis of the sensor data
Connected Services
Analysis/Simulation


Software / Algo Development
ECU Software
Test drives
HIL
In-vehicle recording
Replay of sensor data
Analysis of the sensor data
Connected Services
Analysis/Simulation


Software / Algo Development
ECU Software
Test drives
HIL
In-vehicle recording
Replay of sensor data
Analysis of the sensor data
Connected Services
Analysis/Simulation

OVERVIEW

Data recording

Test fleets collecting real environment test data

- Data decoupling
- Chronological recording of all data sources
- Clock synchronization and time stamping
- Monitoring of data integrity

Data processing

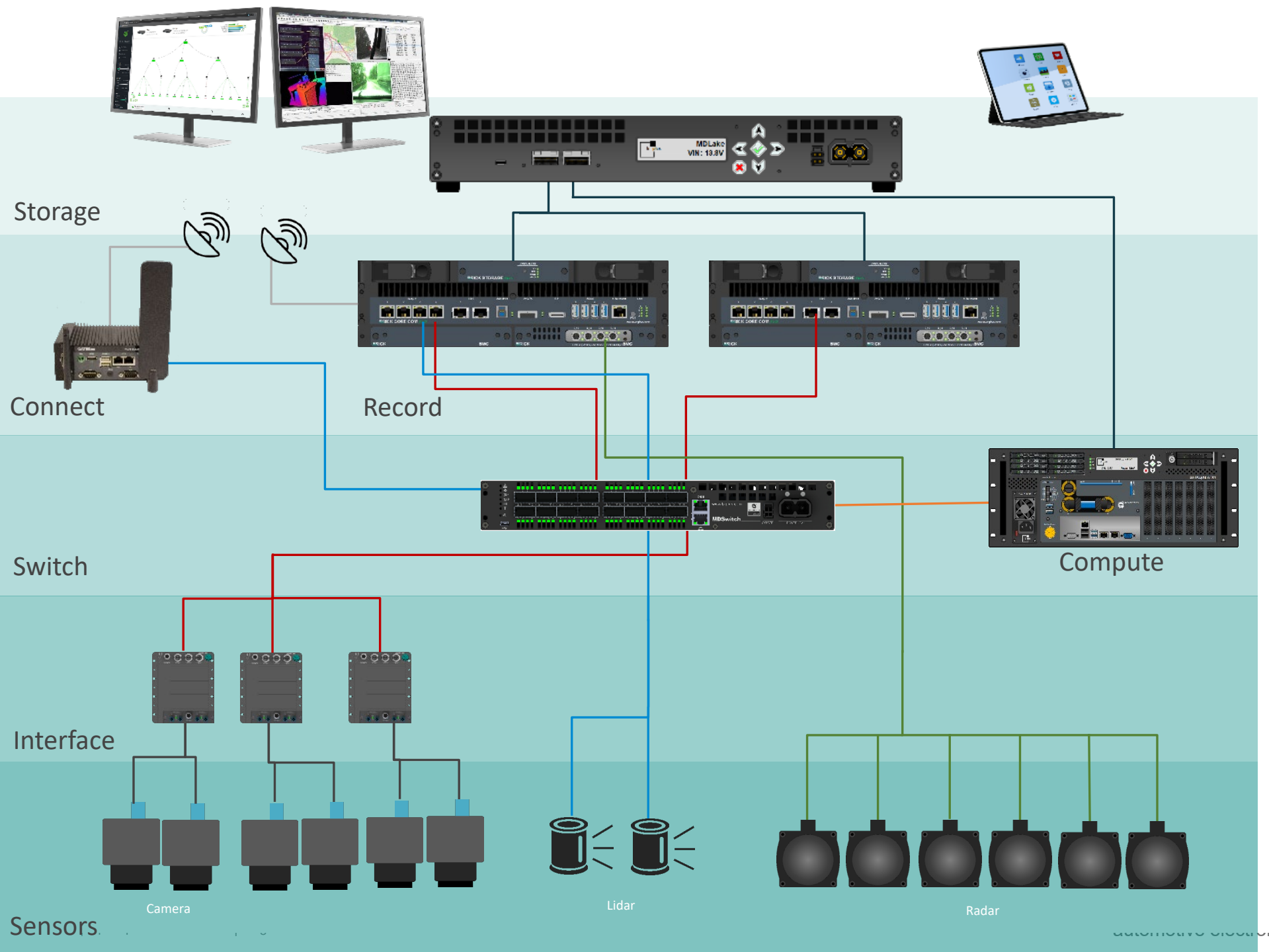
Data quality assurance & enrichment

- Analytics / Scene detection
- Data enrichment
- Automated Labeling
- Live Processing

Data ingestion

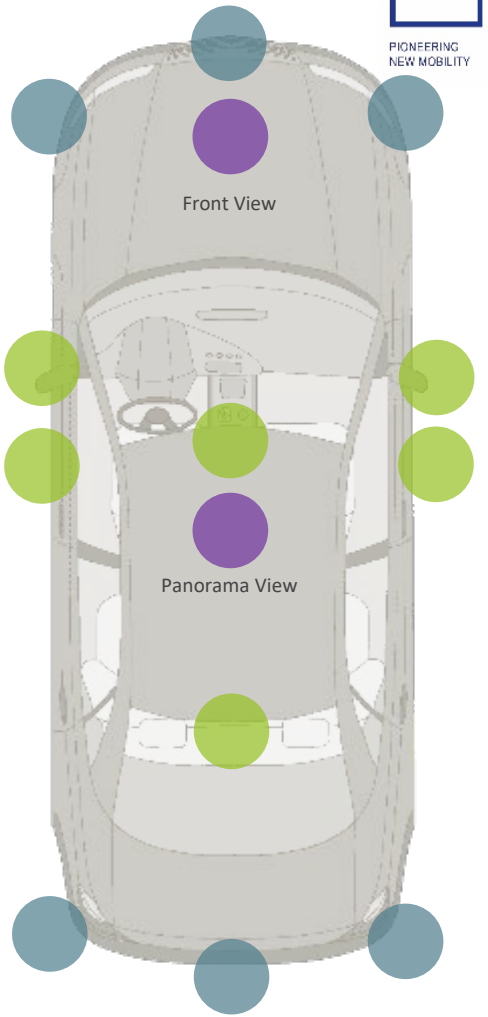
Upload into data center

- Data transfer from test cars to data center
- Copies from various sources via copy station or storage unit
- Secure hash support



www.b-plus.com

b - plus
PIONEERING NEW MOBILITY



100 GBE
40 GBE
10 GBE
CAN
1 GBE
Camera
Lidar
Radar

DATALynx ATX4 Car Server

- Latest Server grade CPUs (AMD EPYC™ Processors or Intel® Xeon® Scalable Processors) (up to 2 CPUs)
- Liquid cooled processors and power supply (external cooling option)
- 6 – 32V DC automotive PSU with Ignition
- Up to 768GB DDR4-2933 ECC RDIMM
- Up to 80TB SSD space
- 7 PCIe Slots
  - Tailor made configurations
  - up to 2 NVIDIA Tesla cards
- Support for Multi-CPU/GPU configurations
- Add-ons for up to 4 GPUs
- -10°C to +60°C full operating

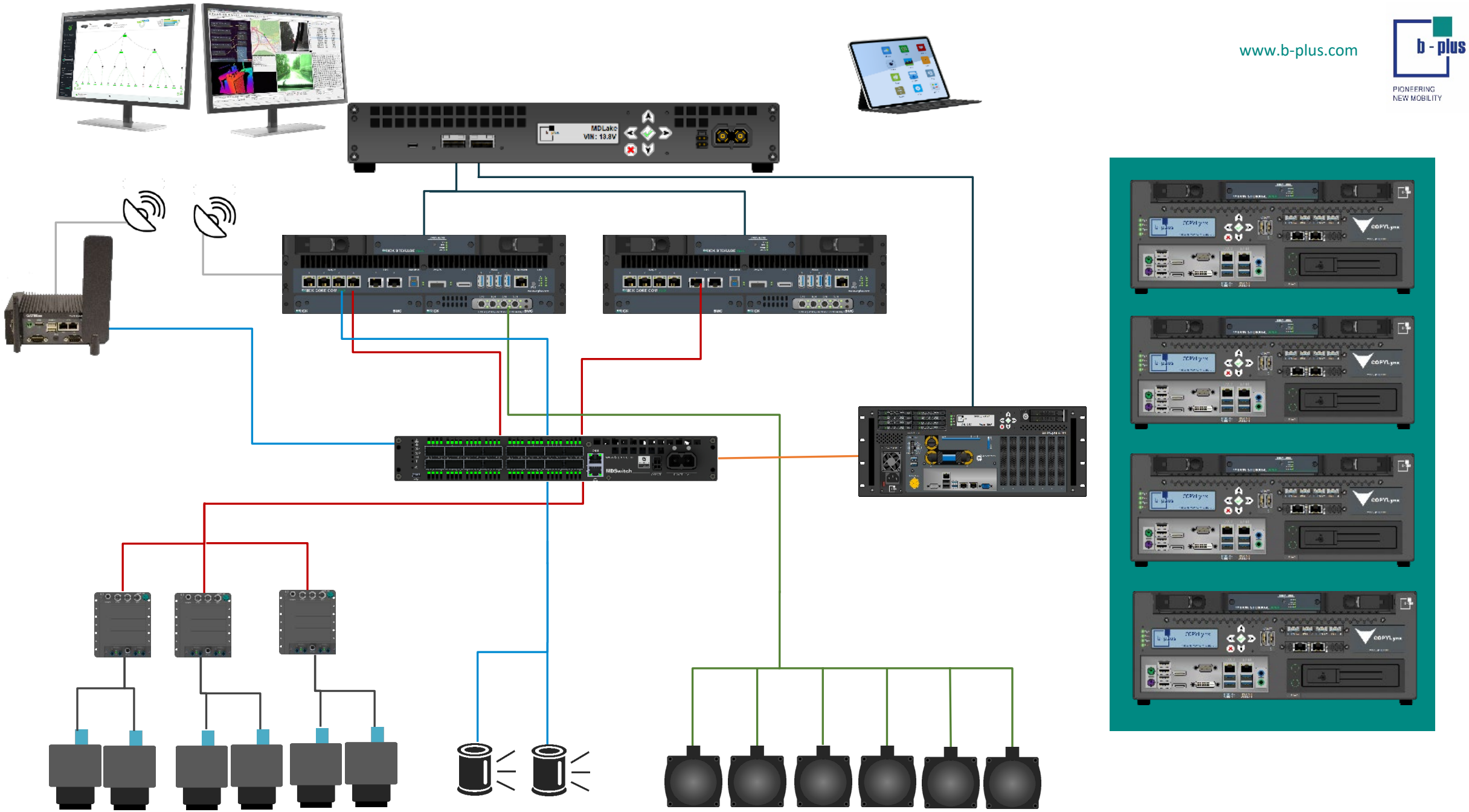
Mobile Datacenter Switch / MDSwitch

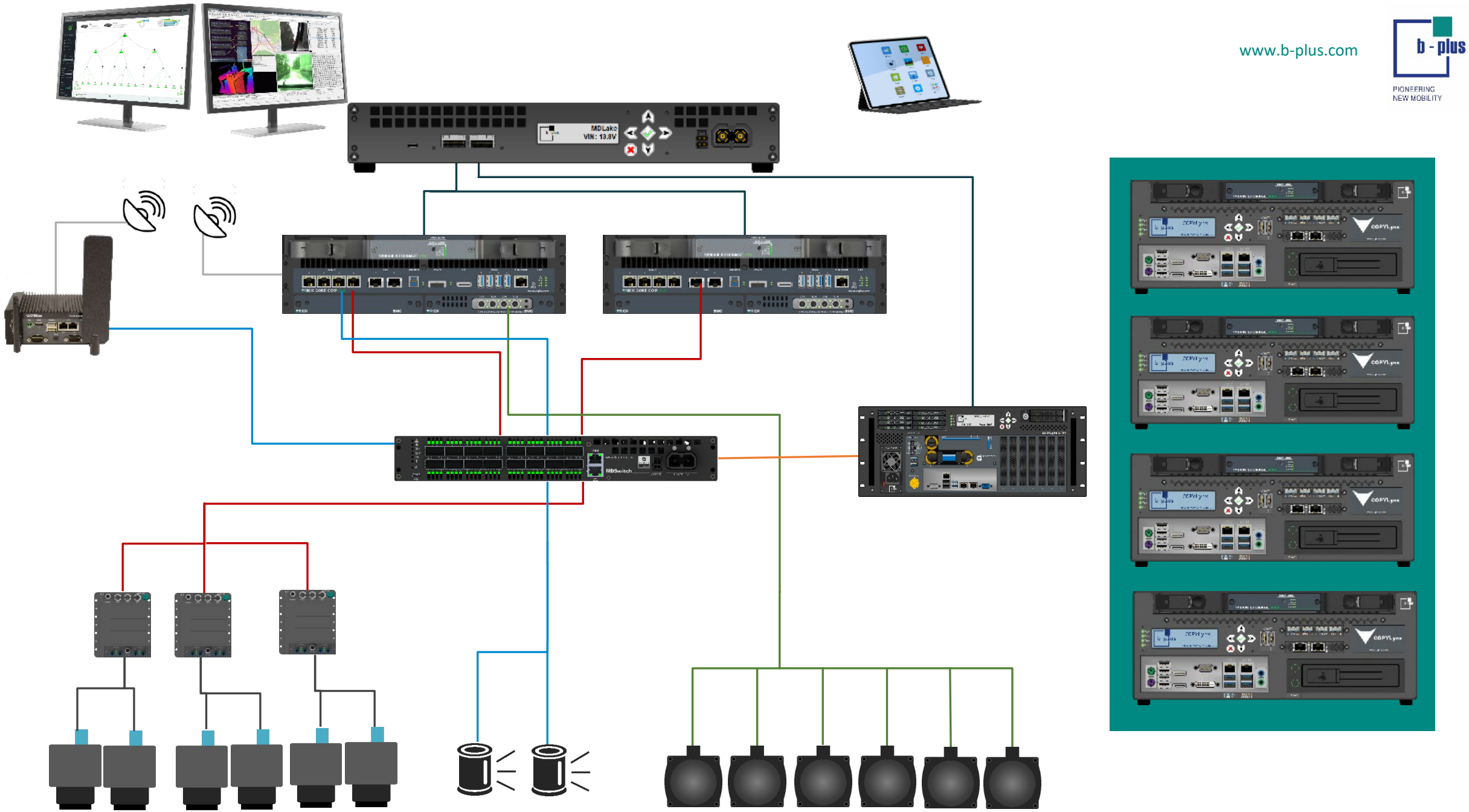
- Automotive Timesync Switch with IEEE 1588 802.1AS
- 16x 100GbE e.g. for upstream to Mobile Data Lake
  - Or 32x 50GbE
  - Or 64x 25GbE/10GbE
- Configuration via Website and Remote Update (Firmware and Config)
- Automotive power supply
- 16x QSFP28 interfaces for Fibre or Twinax cables

Mobile Data Lake / MDLake

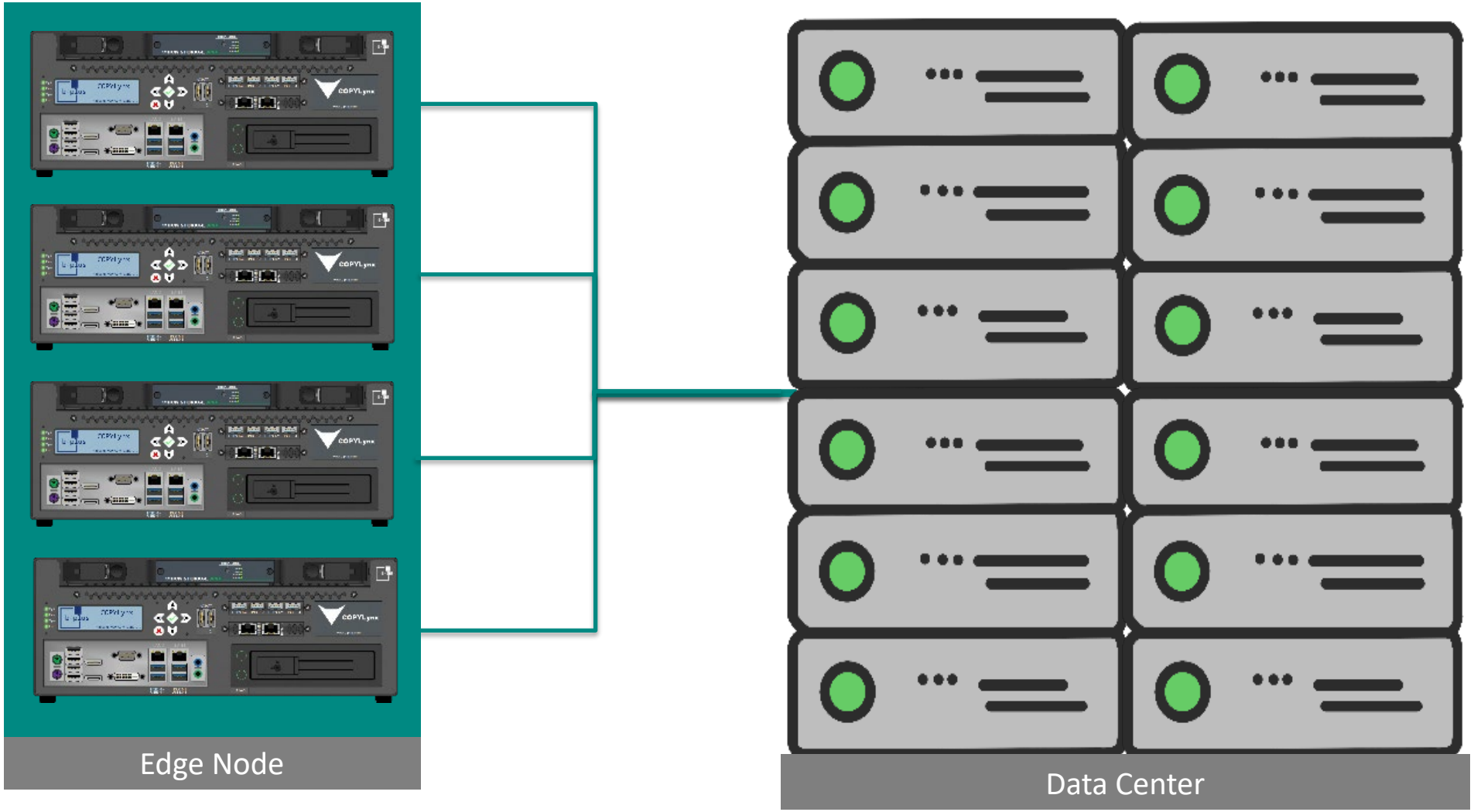
- Automotive DataLake System
- 16-Core ARM SoC with 16GB DDR4 and PCIe Gen4 Switch
- 16 NVMe SSDs connected via PCIe 3.0/4.0 x2

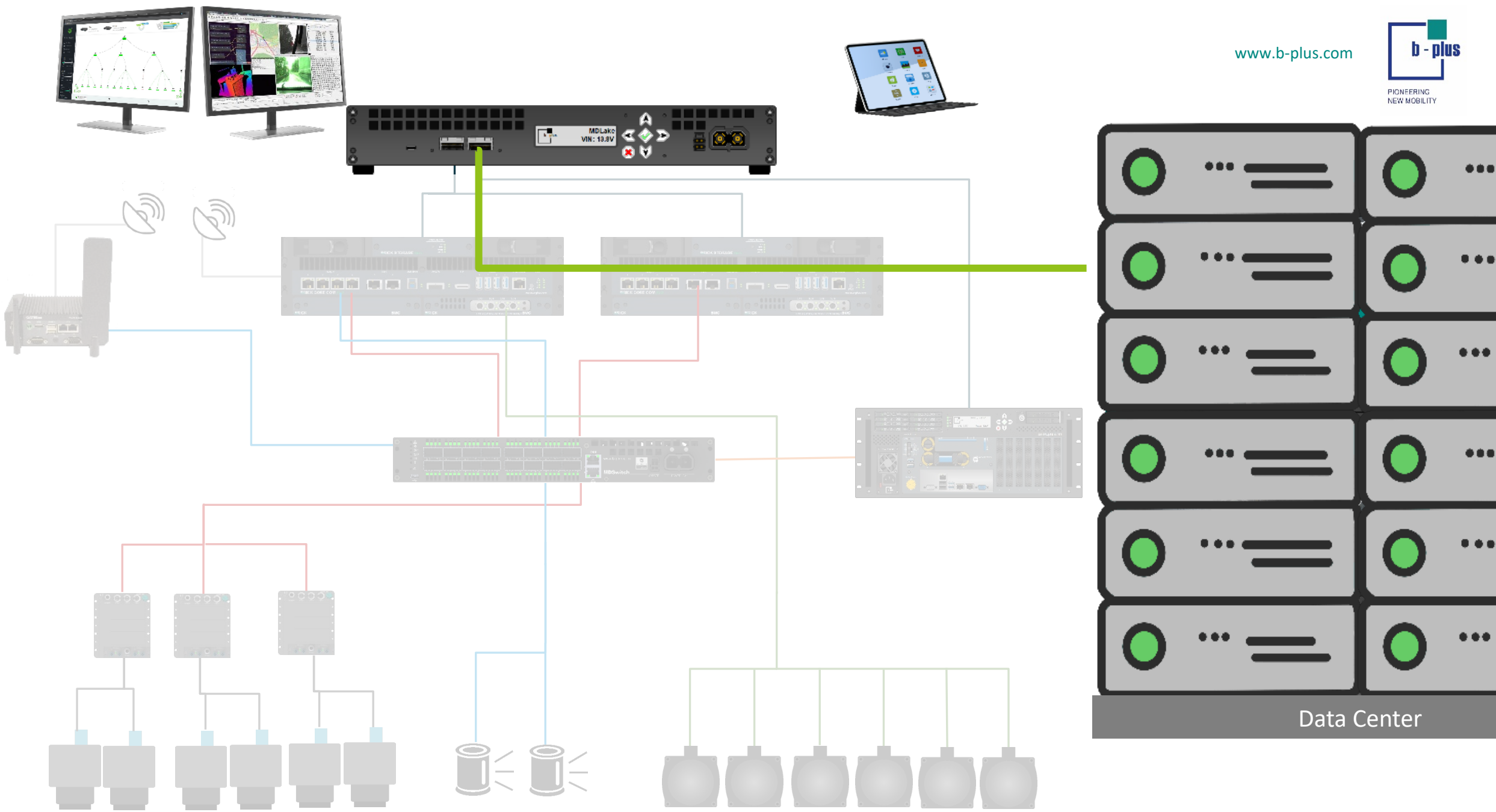
- Enables a 30TB to 180TB Storage JBOF
- 2x 100GbE QSFP28 Ethernet interfaces
- Low latency and highest throughput with RDMA technology
- NVMe-oF PCIe interface for Linux and Win for any Recorder
- 350W TDP, ~210W typical power consumption
- HxWxD: 44x320x380mm

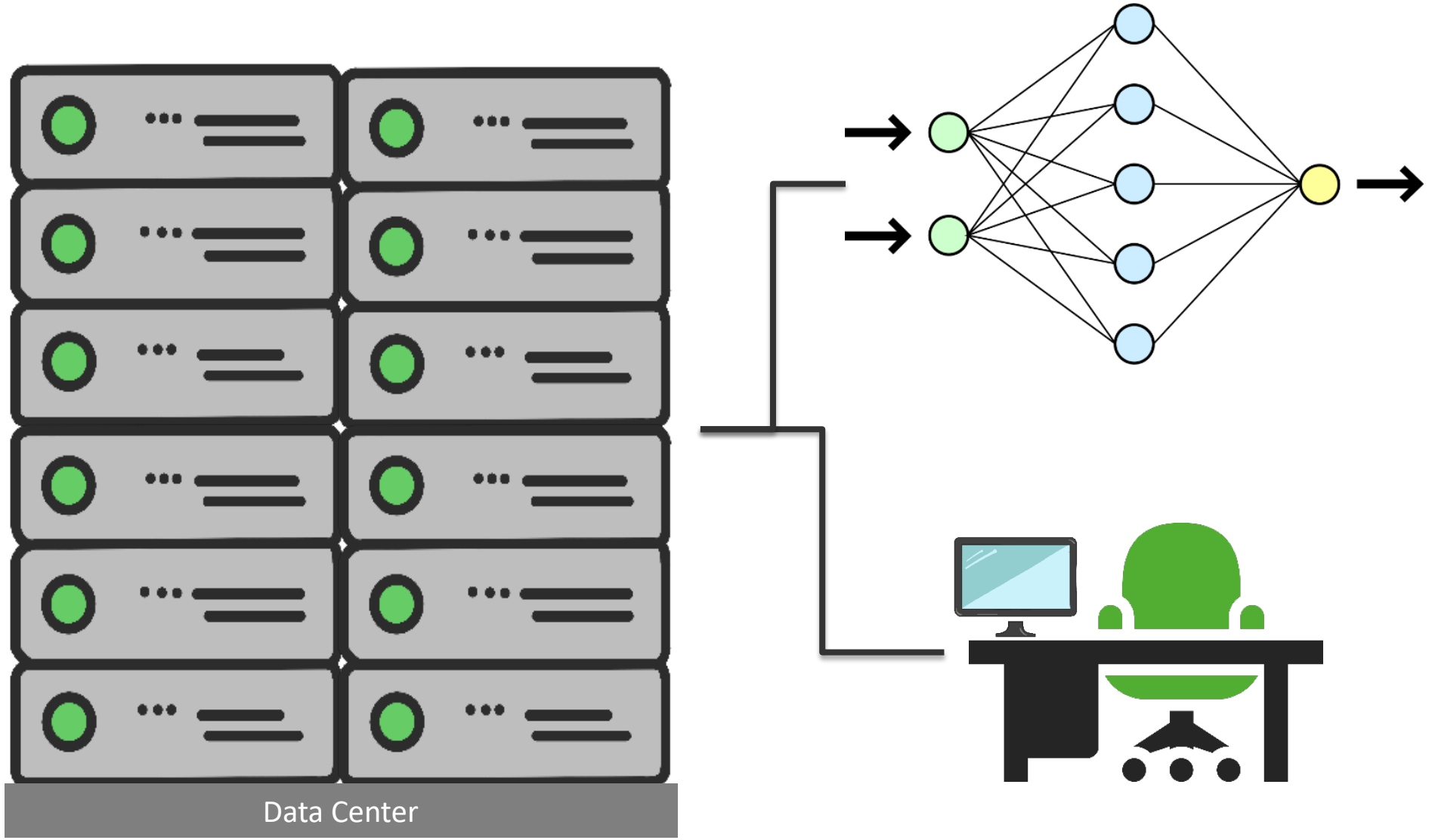

MDLake
VIN: 13.8V
MDSwitch


MDLake
VIN: 13.8V
MDSwitch

DATA INGEST




MDLake
VIN: 13.8V
Data Center


Data Center

b - plus

PIONEERING
NEW MOBILITY

Alexander Noack
Head of Automotive Electronics bei b-plus GmbH

G2M RESEARCH

WEKA

Weka

- Shailesh Manjrekar
- Head of AI and Strategic Alliances
- www.weka.io

WEKA

NVIDIA PREFERRED SOLUTION ADVISOR

Mobility-as-a-service with WEKA ai

Shailesh Manjrekar,
Head of AI and Strategic Alliances

Agenda

- Mobility-as-a-service – MaaS, paradigm change!
- New Architecture – Edge to Core to Cloud
- New Approach – Fuel your Digital Transformation with Accelerated DataOps

“Data is the new source code”

01

Mobility-as-a-service

$54B in 2019 to $556B market by 2026 at 39% CAGR – Globe News wire June 2019

THE FUTURE OF TRANSPORTATION STACK

COMET LABS

SERVICES

ROUTE PLANNING
SPATIAL, CITIViZ, waze, CommuteMaster, X, moovit, Citymapper

PARKING
Waypals, Cloud Parc, Aware, VoicePark, SpotHero, PEARL, seeveder, Parko, FenSens, LUXE, ginxsys

CAR HAILING + POOLING
TESLOOP, lyft, RideCell, 滴滴

OTHER: AFTERMARKET, REPAIR, RENTAL
CARVUE, upshift, WRENCH PATROL, zipcar, Getaround, URGENTLY, Tantalum, bitkar

SAFETY & SECURITY

PHYSICAL CAR & DRIVER SAFETY +ACCIDENT DETECTION
drivebetter, splitsecond, nexar, GUARDIAN Optical Technologies, phrame, LyfeLens, SMART AUTOLABS, tourmalinelabs, iproc, eLSys, FORESIGHT

EMOTION, FATIGUE & ALCOHOL DETECTION + DISTRACTION AVOIDANCE
CORTEXLABS, eyeris, ALMA care, CARMA, saverone, SMART wheel, Radiomize, EASYC, carvi

CYBERSECURITY
CYMOTIVE, cognomotiv, VisualThreat, ARGUS, Olympus Sky, Karamba Security, TriAlium, Arilou

INTRUSION, TRACKING & RECOVERY
trackstop, nonda, gmt Connect, ROADEYE, caruma

IN-CAR INTELLIGENCE + ASSISTANCE

VEHICLE DIAGNOSTICS & PREDICTIVE MAINTENANCE + SENSOR-BASED VEHICLE SAFETY
TrueMotion, BRAIQ, IQ, metromile, LIGHT METRICS

PASSENGER-FOCUSED SENSORS (INCLUDING USAGE-BASED INSURANCE)
Cartics, Zubie, nauto, SONICLUE, PREDICTSYSTEMS, NEBULA, engie

INFOTAINMENT + DISPLAY
link motion, HUDIFY, CARROBOT, PonnCar, swedspot, OverDrive, ROAV, FICONIC SOLUTIONS, RIGHTWARE, PELAGICORE, CLOUDCAR

PERSONAL / VOICE ASSISTANCE
navideck, AUTOMATIC, dash, GRIN, cariq

NAVIGATION ASSISTANCE + PEDESTRIAN ANALYSIS & COMMUNICATIONS
WAYRAY, PERCEPTIVE AUTOMATA, navdy, hudly

AUTONOMY

AUTOMATION SYSTEM
AutoX, drive.ai, POLYSYNC, OXBOTICA, oryx, ARGO AI, AIMOTIVE, PILOT, cruise, nuTonomy, Vector.ai, BWV, MOMENTA, cognivue, AIDRIVING

MAPPING, SIMULATION, & IMAGE RECOGNITION / ANNOTATION
ivi, minds.ai, NOMOKO, Mapper.ai, CARMERA, TOMTOM, IMPROBABLE, UDACITY, DEEPMAP, understand.ai, Civil Maps

AUTONOMOUS VEHICLE MAKER + TOOLS
nuro, WAYMO, ZOOX, FARADAY FUTURE

INFRASTRUCTURE + CONNECTED CAR

SENSOR NETWORKING INFRASTRUCTURE (V2V, V2X) - LPWA, CELLULAR, WIFI
DIGIMONDO, SAVARI, ZiFiSense, EVVOS, Actility

CONNECTED CAR - DATA, PLATFORM, SOFTWARE
ACTIVESCALER, INRIX, SPARC, mobiliuz, carvoyant, vinli, otonomo, bright box, VEHICLE DATA SCIENCE, mojio

FLEET + TRAFFIC MANAGEMENT
BESTMILE, Cloud Your Car, synovia, Fleetmatics, urban engines, Motionlogic, Travelai, Automile, Zendrive, GLYDEL, locoNav, Immense Simulations, MOTORWAY BUDDY, mentis, FR, Brisk Synergies

OTA CAR SOFTWARE UPDATION + SMART PHONE ENABLEDTELEMATICS
FIXD, DRIVEBOT, VOYOMOTIVE, ZENE, aryuna

INTELLIGENT MANUFACTURING

NEW/ADVANCED MATERIALS
AQUARIUS, NAWA TECHNOLOGIES, SIRRUS, NANOSTEEL

RAPID PROTOTYPING - 3D PRINTING, MODULARIZATION, OPEN SOURCE
Carbon, nanocore, DIVERGENT, OSVehicle, stratasys, VADER

ADVANCED / AUTOMATED ASSEMBLY LINE
Grabit, SYMBIO, CLEARPATH, ROBOMOTIVE, rethink robotics.

MATERIAL CHARACTERIZATION & TESTING
SynTouch, SASTRA, VIDI, OPTOFIDELITY

ONBOARD SENSORS

LOCATION - GIS, PRECISION POSITIONING, PATH PLANNING
swift NAVIGATION, radiosense, EXO TECHNOLOGIES

VISION / CAMERA
DEEP VISION, Carnegie Robotics, machines with vision, ROADSENSE, DEEPSCALE, Chronocam, VAYAVISION

LIDAR
faceI, INNOVIZ, PHANTOM INTELLIGENCE, VOYAGE, QUANERGY, AEYE, TRILUMINA, LeddarTech, Velodyne LiDAR, STROBE

RADAR
OCULii, TOPOSENS, RFISEE, 6TH SENSE, SDS

SPECIALTY VEHICLES

2-WHEELERS
MOOVIKE, smarthalo, superpedestrian, Zagster

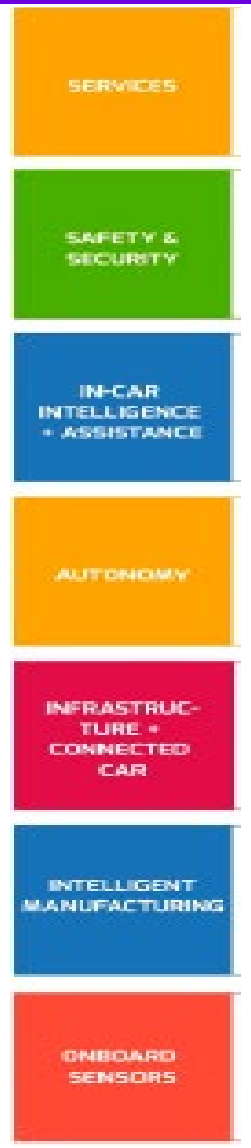
PUBLIC TRANSPORT
navya, Varden Labs, AURO, EASY MILE

TRUCKS / FREIGHT
JAYBRIDGE ROBOTICS, OTTO MOTORS, PELOTON, TESLA, FLEXPORT, TRUELITE TRACE, STARSKY ROBOTICS

FLIGHT
WRIGHT ELECTRIC, ZUNUM Aero, ZEE, JOBY, Kitty Hawk, LILIUM AVIATION, AEROMOBIL

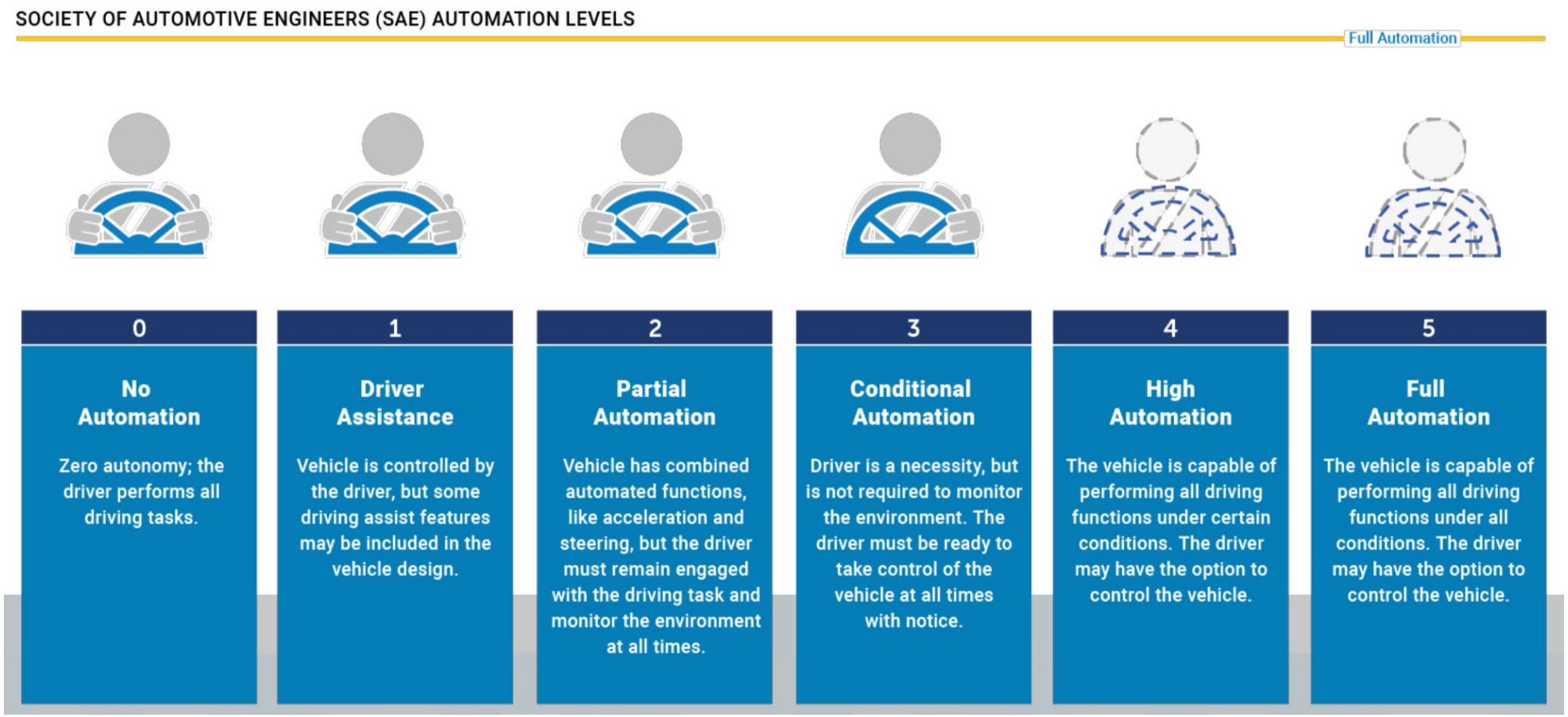
OTHER: HYPERLOOP, PERSONAL MOBILITY
hyperloop one, ninebot, RYNO, WHILL

WeksFS powers Mobility-as-a-service stack



- WE-ride
- (Telematics)
- tuSimple, largest electric car company, autonomous trucks in EMEA
- Cerence – AI, for World in Motion
- Innovize

SAE (Society of automotive engineers) – Levels of automation



Building AI for Autonomous vehicles is hard

Every neural net needs to handle thousands of conditions and geo-locations

- Safety is a non compromisable primary objective
  - Models need to be tested on huge data-sets to be confident; needs performance for faster iterations
- Needs tons of data
  - Innumerable scenarios is the key to building good AV models
- Inference on the edge
  - Edge has limited hardware capabilities, hence ability to chose right model, with multiple, faster iterations and access to tons of data is key
- Reproducibility to understand why a model behaved certain way
  - Proper version control of datasets, models and experimentation

Data has become the new source code and timely access to tons of data is cornerstone for success of autonomous vehicles

Production grade AV platform functionality

Scalable AI training

PB scale AI testing

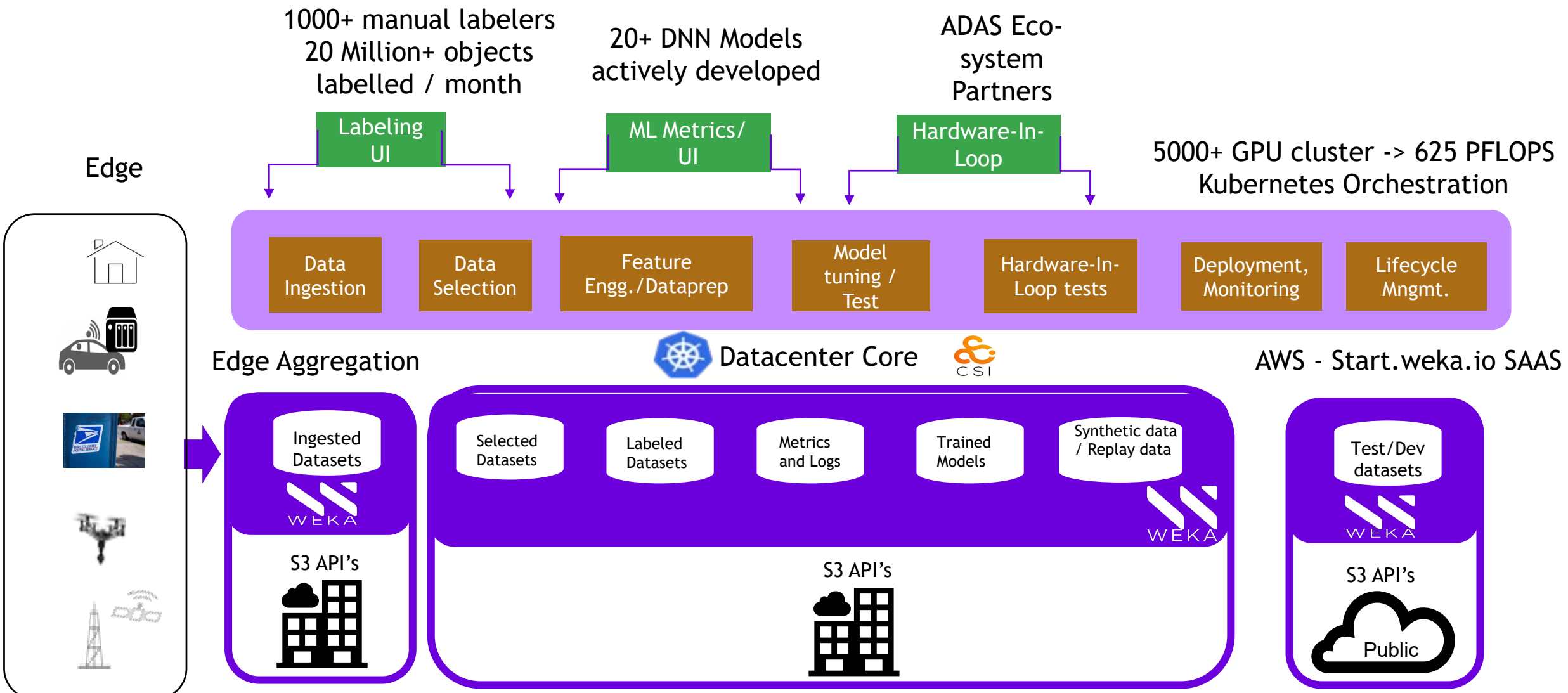
AI based data selection and mining

Traceability: model == code + data

Seamless PB scale data access

Workflow automation

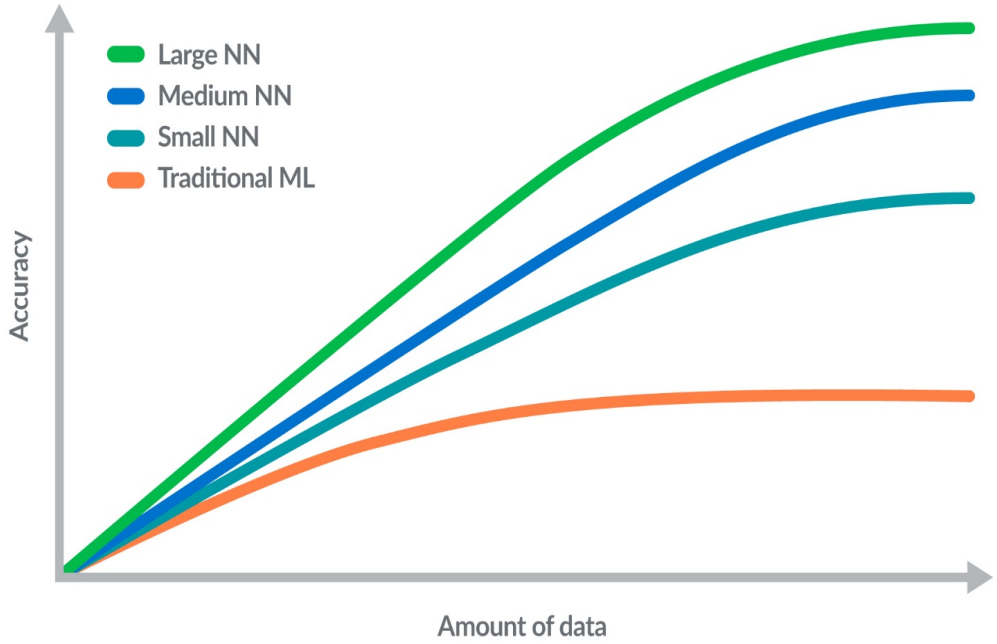
Software Defined Car – Advanced Driver Assistance System (ADAS) data pipeline



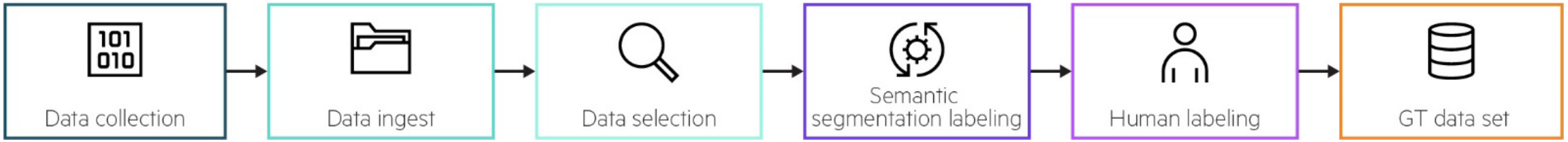
Deep Learning - Data Centric Use Cases

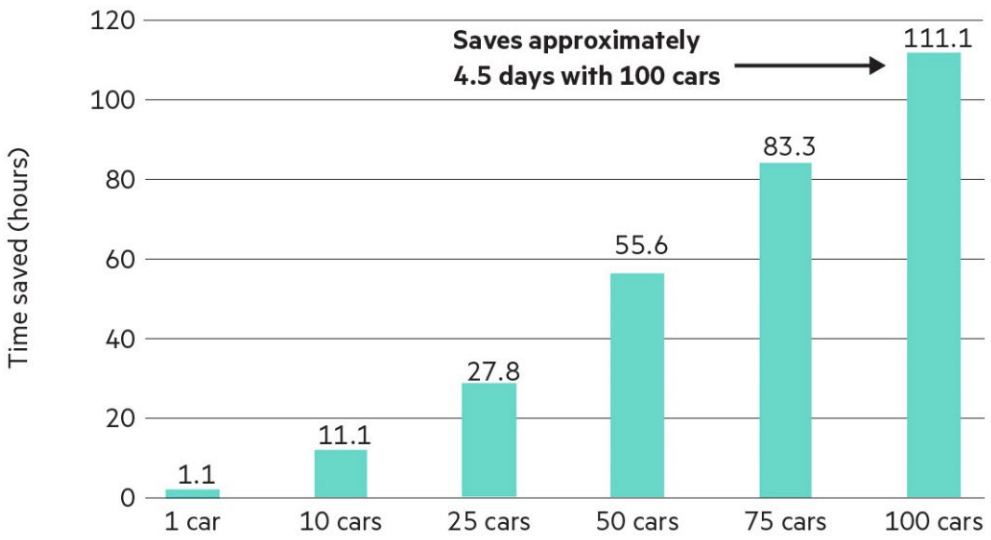
Advanced Driver Assist System (ADAS)

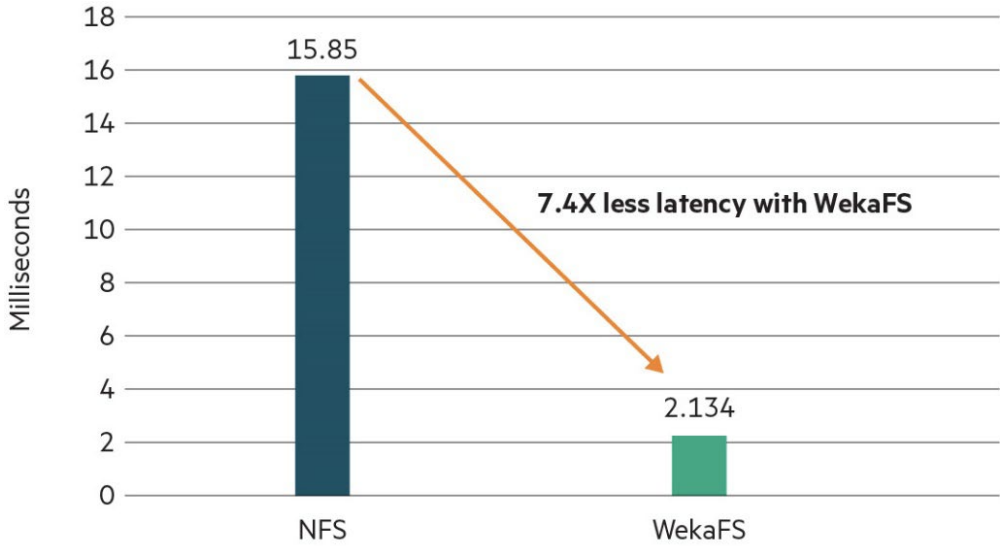
Level of Autonomy	Deep Neural Networks Needed	Survey Car	Data Storage	Data Processing
Level 2+	10	Equipped with 2MP Cameras	2 PB per Year per Car	100 GPU Servers per Car



Semantic segmentation – image classification at a pixel level – Ground Truth (Labeled Datasets)







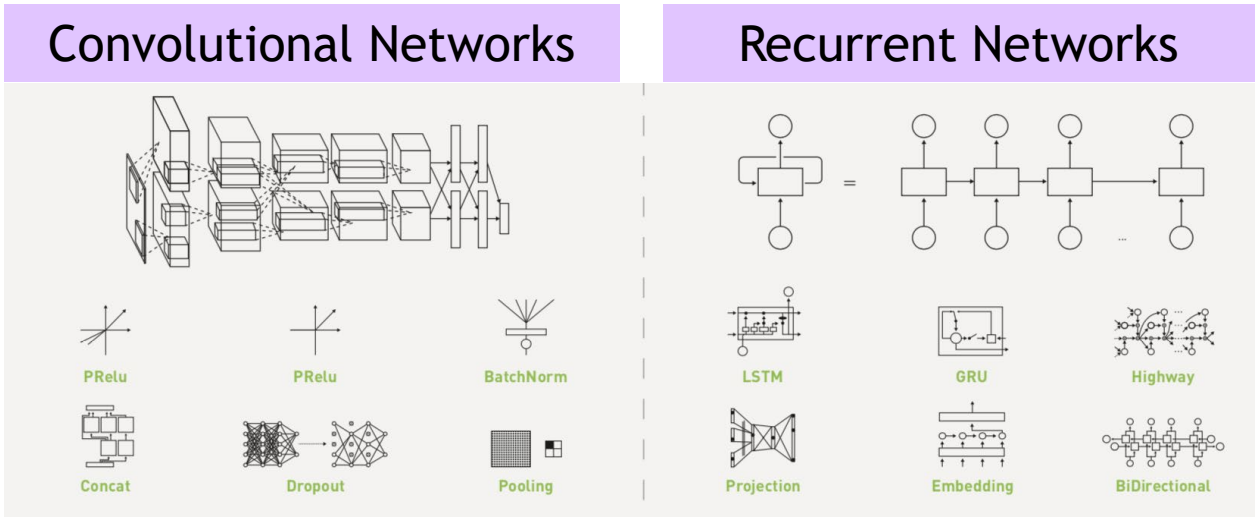
228 GB of City scape dataset used on WekaFS, Apollo 6500 running ENET CNN for semantic segmentation
https://www.weka.io/promo/2019-11-wp-hpe-storage-on-semantic-segmentation-2/

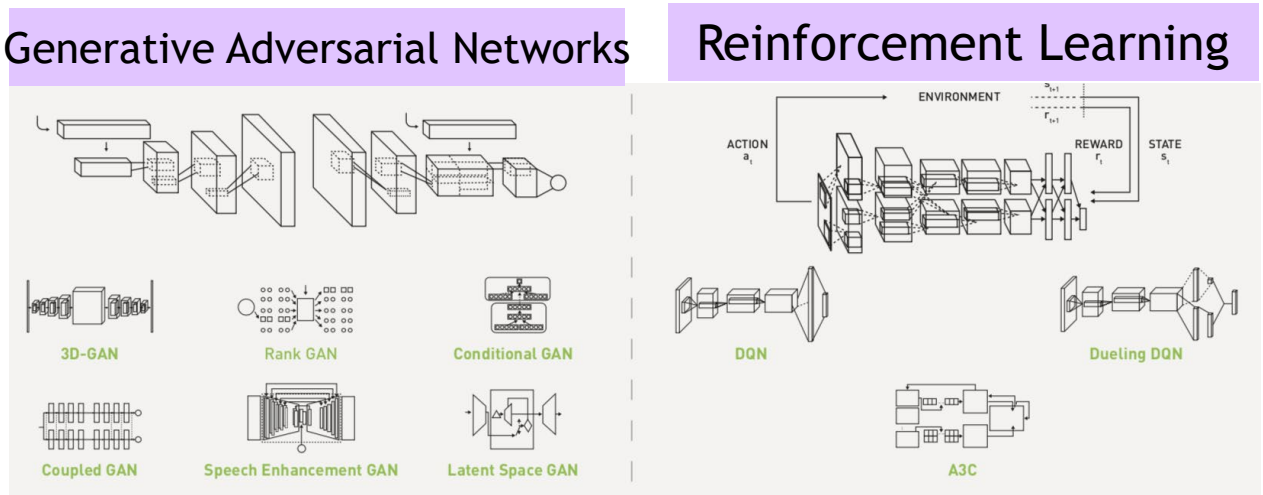
DNN's Becoming More Complex With Several Billion Hyperparameters

Use cases moving from Computer Vision to NLP/NLU and multi-modal

Advances in Deep Learning Methodologies:

- Deep Learning
- Transfer Learning
- Federated Learning
- Active Learning





02

New Architectures

GPUs Have "Densified" Compute into a Single Server Creating a Huge Data Bottleneck

100x More Compute
40x more network

GPU Accelerated Server

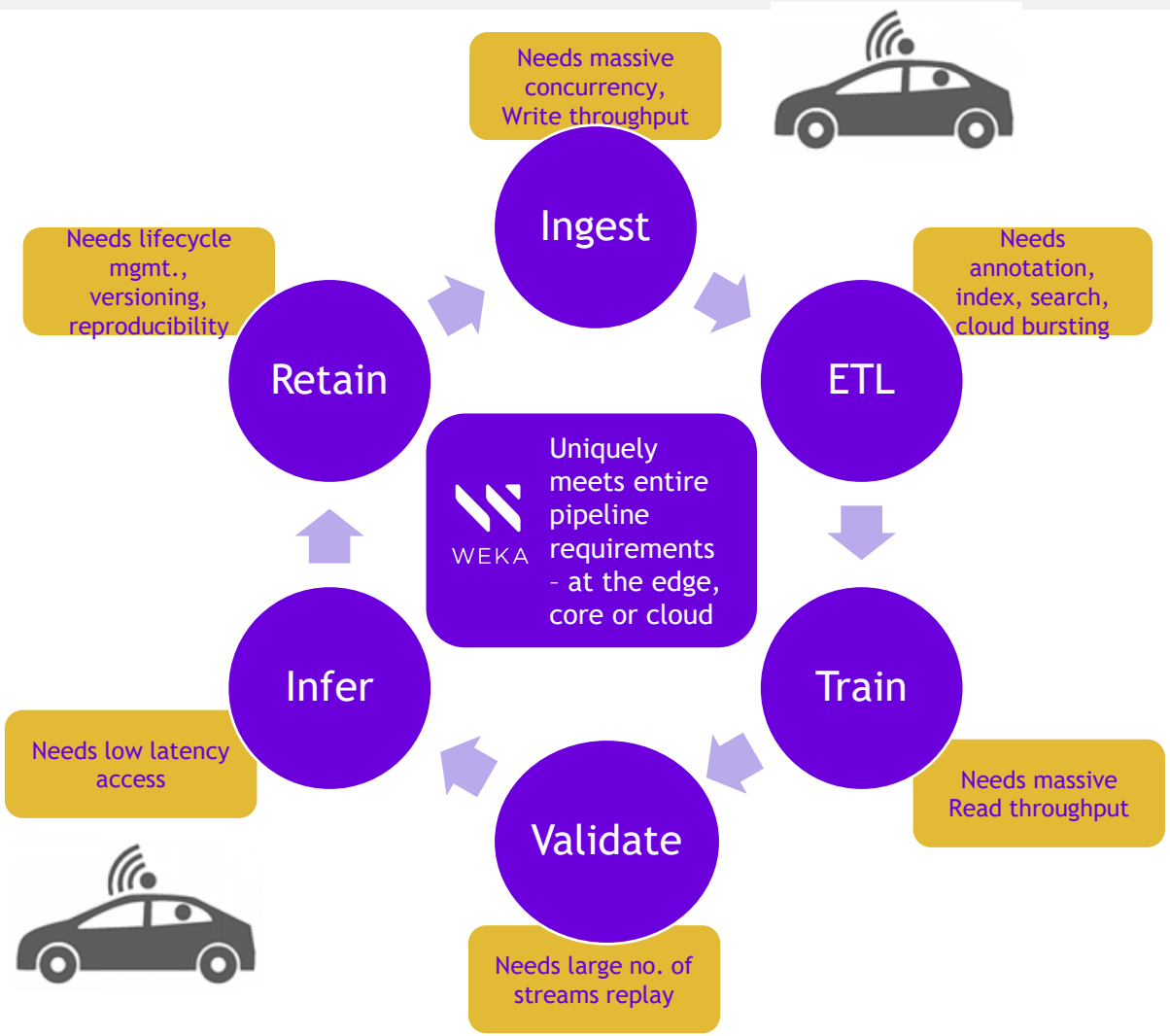
Current NAS solutions cannot feed these machines with enough data

CPU-Only Servers

- 100's of servers with CPUs
- 100's of low bandwidth network connections
- No one server was particularly demanding on storage

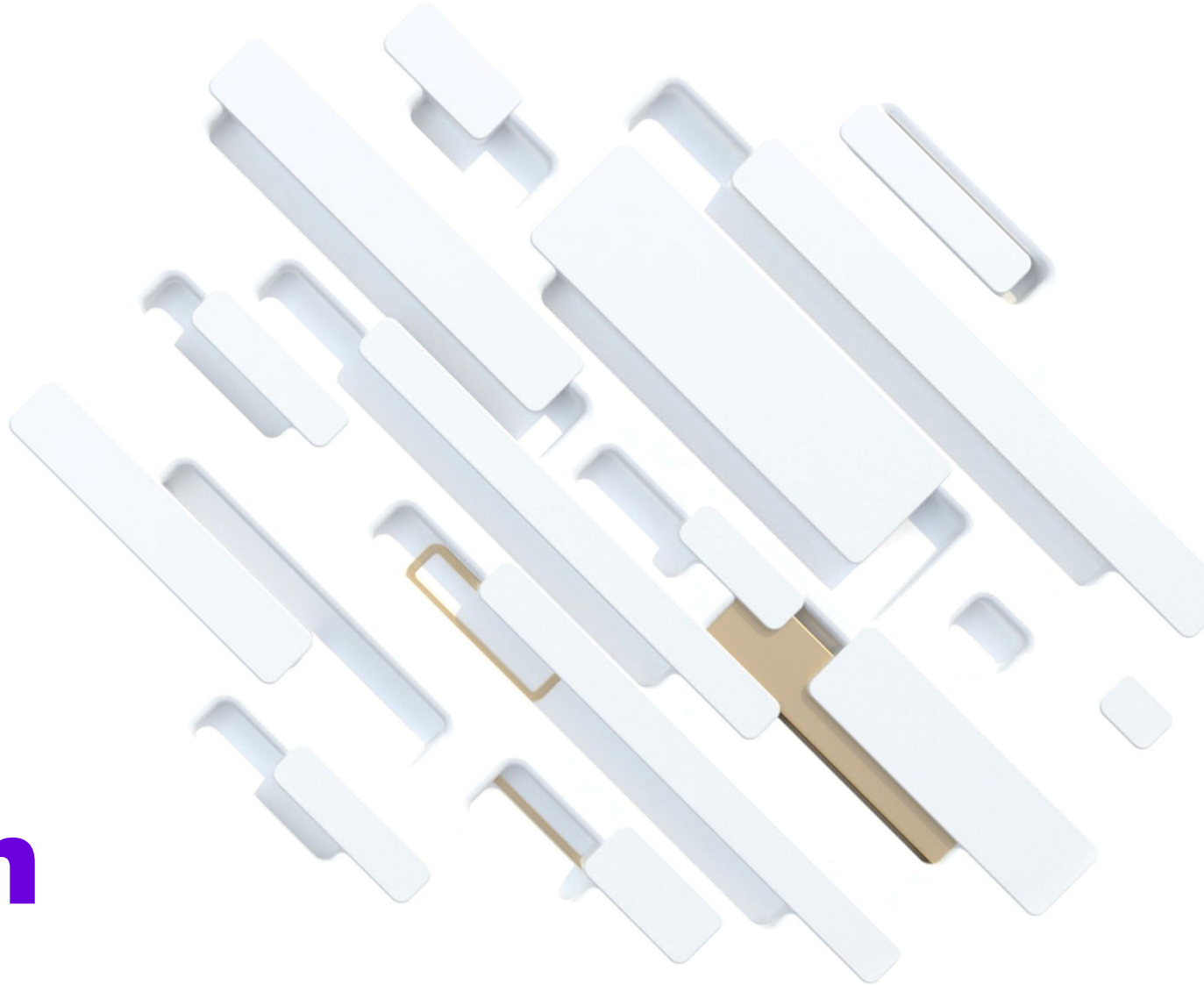
Weka solves the last mile problem

Single storage substrate for the entire data pipeline



03

New Approach



WekaAI for Accelerated DataOps – Small / Medium / Large Bundles

Data Collection: Edge, Core, Cloud

Model Training, Simulation

Inferencing

Lifecycle Mgmt.

TensorFlow

Kubeflow

VALOHAI

Ingest, ETL, Query

Accelerated Libraries

Software-in-Loop, Hardware-in-Loop

Kubernetes / Containers

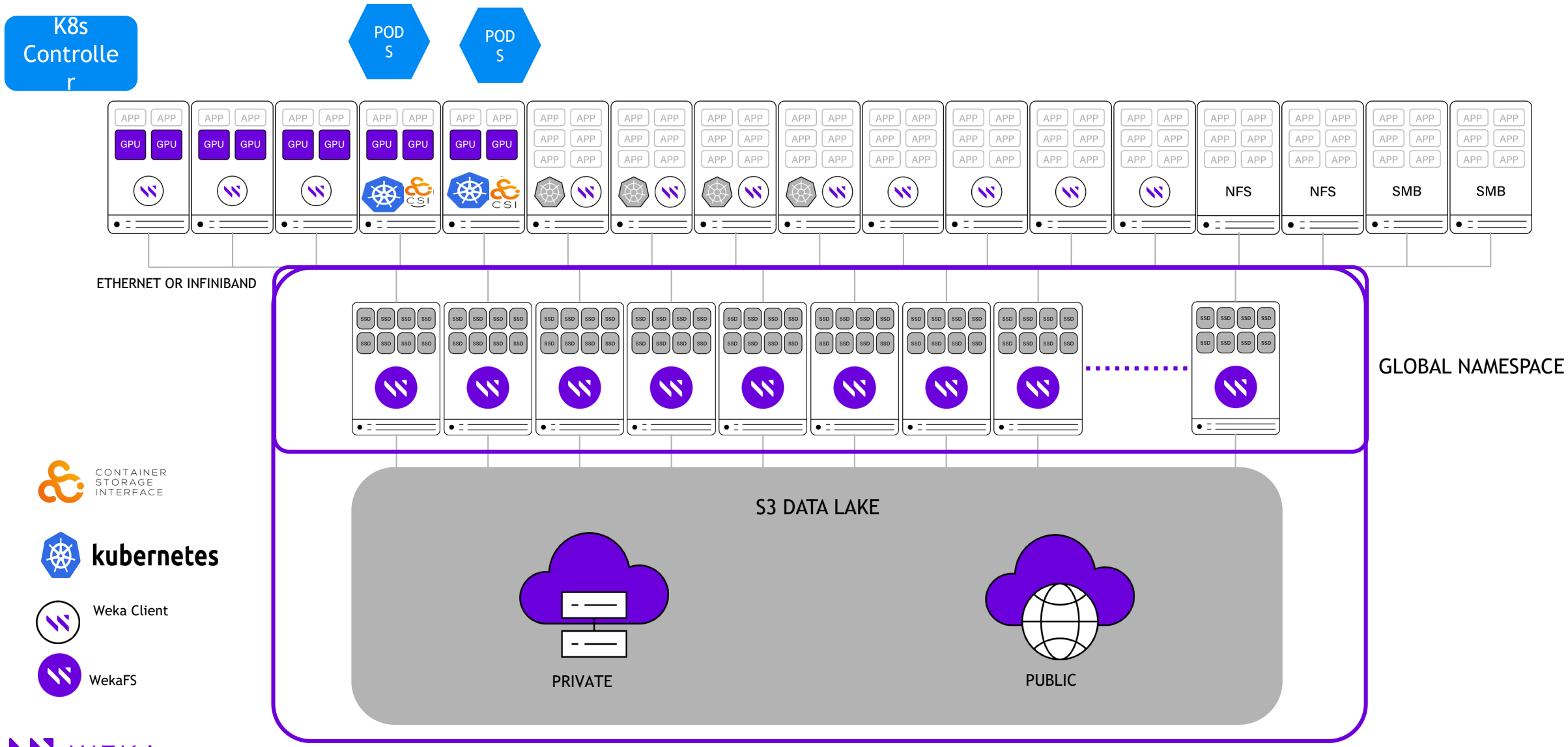
Accelerated Compute: GPUs, Graph Processors, FPGA, Accelerators

WEKA on Storage Servers

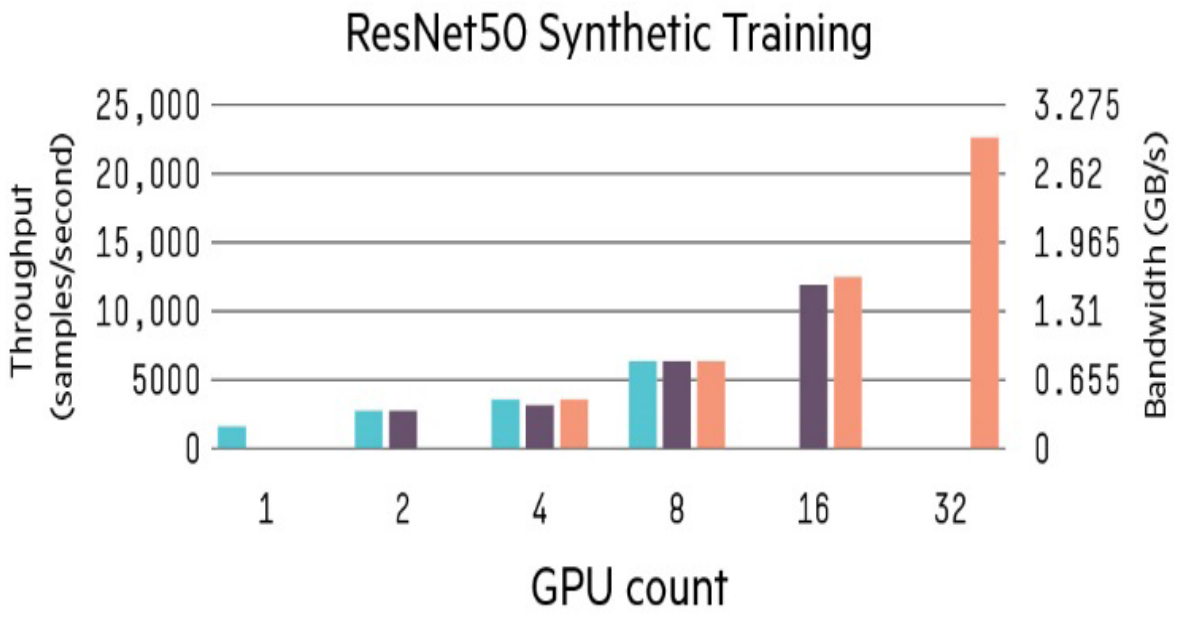
S3 Private Cloud

S3 Public Cloud

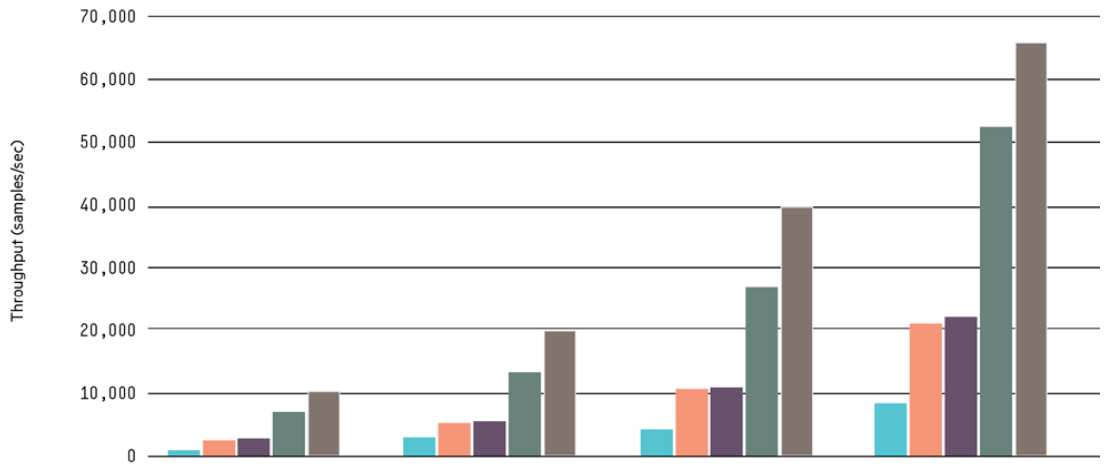
Weka AI deployment



Weka AI, 4 HPE GPU servers - Training and Inference samples/sec



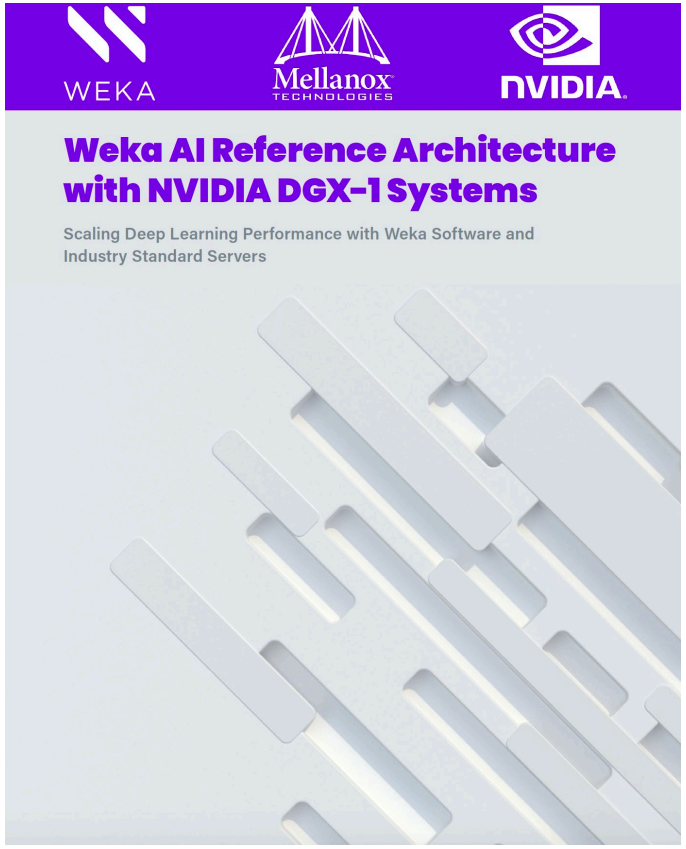
WekaIO Inference Throughput by GPU

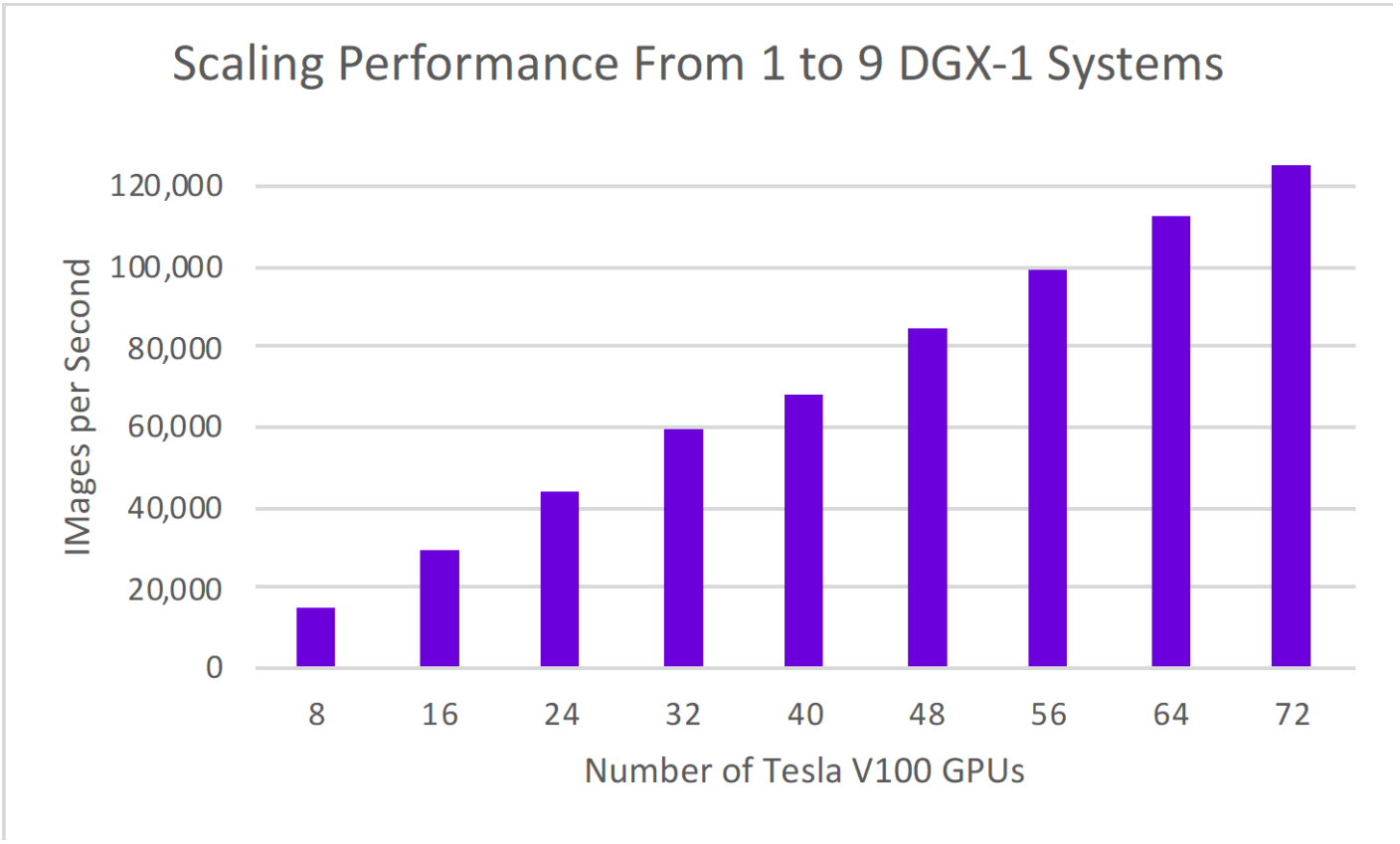


Number of GPUs	1	2	4	8
Inceptionv4	1048	2094	4185	8375
ResNet152	2605	5210	10,421	20,846
VGG16	2738	5486	10,933	21,900
ResNet50	6682	13,312	26,675	52,368
GoogleNet	10,063	20,101	39,919	65,534

Imagenet dataset used TFRRecord on WekaFS, 4x Apollo 6500 running Resnet50 CNN for training and inference. Horovod and NCCL used for intercommunication https://www.weka.io/wp-content/uploads/files/2020/04/HPE-accelerate-performance-for-production-ai-tech-whitepaper-1.pdf

Weka AI Reference Architecture

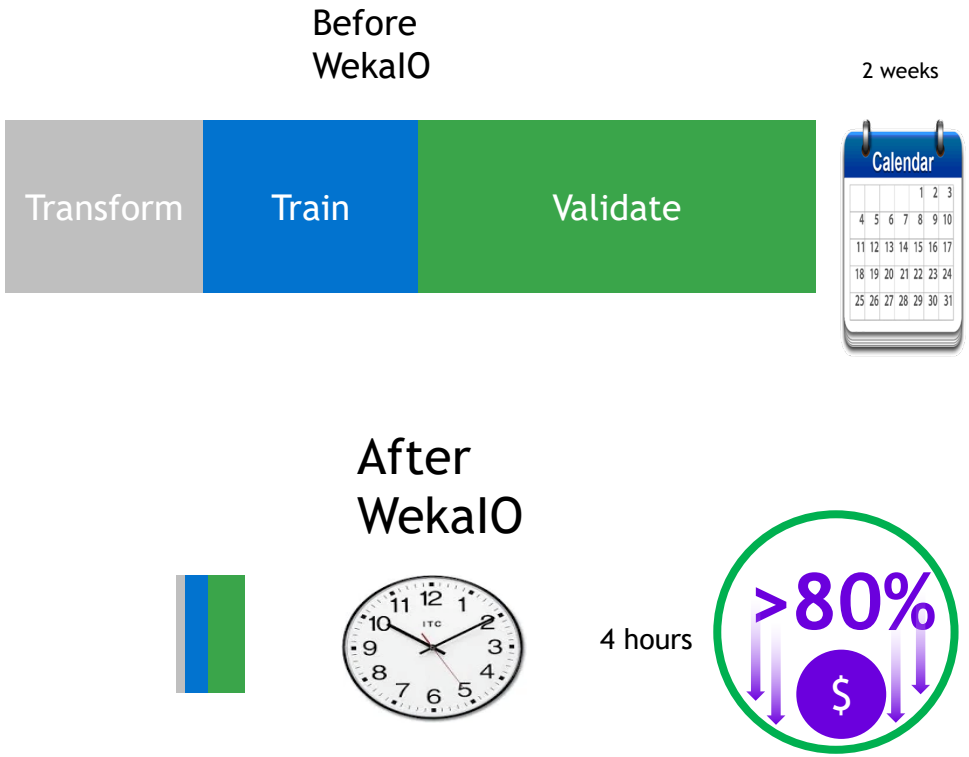




- Fully saturate 100Gbit Network link
- 3x faster than local drive Storage
- 10x faster than all flash NAS
- Perfect linear scaling as cluster expands

Weka AI for Datascientists, CDO's and CAO's

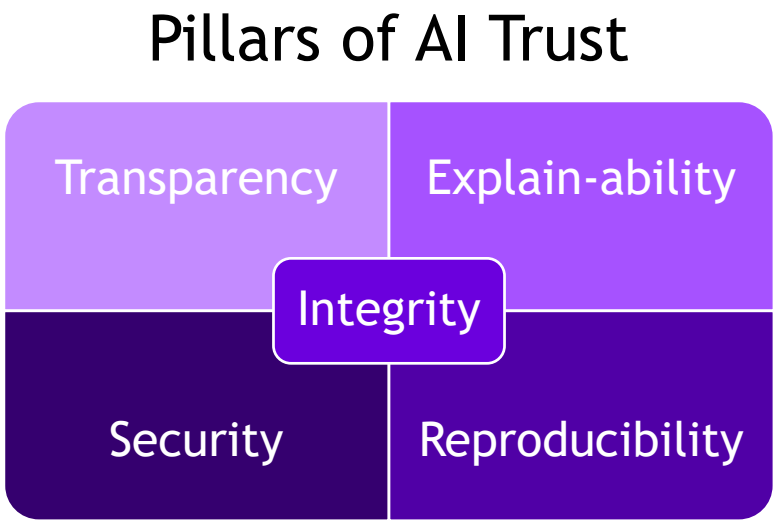
- Improve productivity and faster time to market and value
  - accelerate large scale data pipelines with reduced epoch times, fastest inferencing and highest images / secs benchmarks
  - run entire pipeline on the same storage backend
  - Faster than local storage



30% better utilization results into $1.13M in savings for 10 node GPU cluster with 3 Data scientists, over 3 years

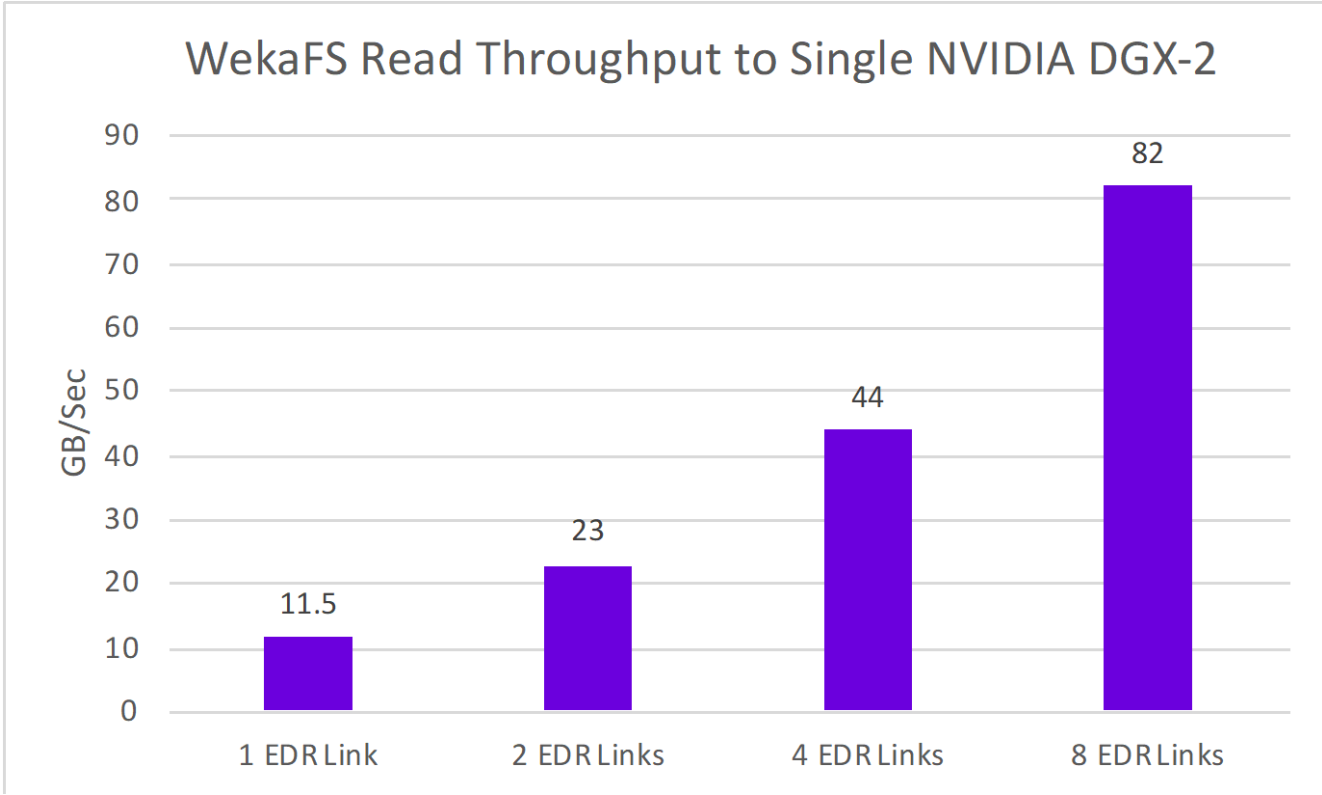
Weka AI for Data scientists

- Data compliance and security
  - in-line encryption support enables compliance
- Explain-ability and Reproducibility for experiments
  - instant space efficient snapshots make it easy to maintain versions
  - Snap2object retains versions for reproducibility and explain-ability
- Hybrid workflows
  - Dev and Test experiments in the public cloud, data mobility and rehydration on-premise for production



WekaFS Delivered 82GB/sec to a single NVIDIA DGX-2

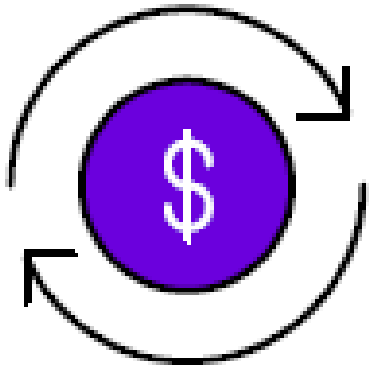
- WekaFS with GPUDirect Storage fully saturates an InfiniBand EDR link, delivering full bandwidth to data hungry applications
- Performance scaled linearly as more EDR links were added
- WekaFS achieved 82GB/sec to a single NVIDIA DGX-2 across 8 EDR links using 2 Supermicro BigTwin servers



NVIDIA and Weka are currently running joint GPUDirect Storage beta program with customers and partners!

LOSF – Lots of Small Files

- Eliminating silos and multiple copies, but providing single storage platform for entire data pipeline
- Best scalability up-to EB's of storage with trillion of files across directories and single directory
- Best agility with data management across the edge, core and cloud
- Best TCO, with performance at scale
  - leverage NVMe flash for performance and object store for capacity
  - built in data protection eliminates need for another solution



Marquee customers

cerence

"We looked at our legacy architecture and instead of taking an evolutionary step and upgrading every component, we took the revolutionary approach. Weka cost-effectively enables both the use of POSIX and object storage with performance and latency that is far superior to any other solution."

Bridget Collins, Chief Information Officer

Electric Car Company

*WekaIO demonstrated that it was the only file system that could **fully saturate the GPU cluster**. With WekaIO, the data scientists were able to significantly improve productivity by removing time consuming data copy tasks into local disks. In addition WekaIO provided **seamless integration to our massive training system data lake**.*

Thaddeus Fortenberry, Autopilot Infrastructure Architect

tusimple

*WekaIO was the clear choice for our DNN training...**standard NAS would not scale** and Matrix [was] the **most performant of all the parallel file systems** we evaluated...we really liked that it was hardware-independent allowing us better **control over our infrastructure costs**.*

Dr. Xiaodi Hou, Co-founder and CTO

WeRide 文远知行

We built a GPU farm, and we needed a high-speed data pipe to feed it. We evaluated open source solutions, HDFS, and the public cloud. We chose Weka for its ability to provide cost-effective, high-bandwidth I/O to our GPUs, product maturity, customer references, and stellar on-demand support

Paul Liu, Engineering Operations Lead

INNOVIZ TECHNOLOGIES

Weka's storage scalability and the ability to grow the infrastructure without losing performance, was a key factor in the decision to select the Weka file system.

Oren Ben Ibghei, Infrastructure IT


WEKA

G2M RESEARCH

Panel Questions and Audience Surveys

Panel Question # 1

- Clearly, the training problem is as much a storage networking and access issue as it is a volume issue. What sort of architectures are best for performing autonomous vehicle training?
  - NVIDIA
  - Weka
  - b-plus

Audience Survey Question #1

- Is your organization more concerned about storage for data acquisition, training, or operation of AVs? (check one):
  - Our primary concern is on-vehicle storage for data acquisition: 13%
  - Our primary concern is on-vehicle storage for data during AV operation: 7%
  - Our primary concern is on-vehicle storage for both data acquisition and AV operation: 30%
  - Our primary concern is storage in the lab/datacenter for AI/ML training: 23%
  - We are not concerned/no opinion: 23%
  - Other: 3%

Panel Question #2

- Test vehicles have many of the same data storage challenges that training configurations do. What storage architectures and technologies work well for these applications?
  - b-plus
  - Weka
  - NVIDIA

Audience Survey Question #2

- When you consider solutions for AI training and data acquisition, what factor is most important to your organization? (check one):
  - Cost: 23%
  - Storage capacity: 13%
  - Storage performance: 45%
  - Storage networking: 10%
  - Future expandability: 3%
  - Other issues: 6%

Panel Question # 3

- When optimizing storage performance for AI training and validation, what factors should be considered?
  - Weka
  - NVIDIA
  - b-plus

Audience Q&A

G2M RESEARCH

Thank You For Attending

G2M

RESEARCH