IBM SpectrumAI with Nvidia for Autonomous Driving (AD)

Frank Kraemer
IBM Systems Architect
mailto:kraemerf@de.ibm.com

Spectrum Scale UG
11th March 2019 Singapore

IBM

Gardens by the Bay
Sat 9th March 2019

https://www.opengovasia.com/asias-first-fully-operational-autonomous-vehicle-running-at-gardens-by-the-bay-singapore/

Continental

International EN Search

Press Release 2018-11-15

Continental and EasyMile inaugurate new autonomous driving R&D team in Singapore

- **A signed Memorandum of Understanding enables intense collaboration by experts from both companies to prepare the test of self-driving vehicles on public roads of the city-state**
- **Strongly interconnected with Continental's and EasyMile's global Autonomous Driving teams, the Singapore R&D focus will be pointing towards environmental perception, driving decision strategies, artificial intelligence, vehicle safety as well as system integration**

https://www.continental-corporation.com/en/press/press-releases/cube-singapore-150546

NTU and Volvo launch world's first full-sized driverless electric bus for trial

SINGAPORE: Nanyang Technological University (NTU) and Volvo Buses on Tuesday (Mar 5) 2019 launched a 12-metre autonomous electric bus in what they said was a world first.

https://www.channelnewsasia.com/news/singapore/driverless-electric-bus-launched-by-ntu-and-volvo-in-world-first-11311838

Grab and NUS open AI Lab to Transform Cities and Transportation in Southeast Asia

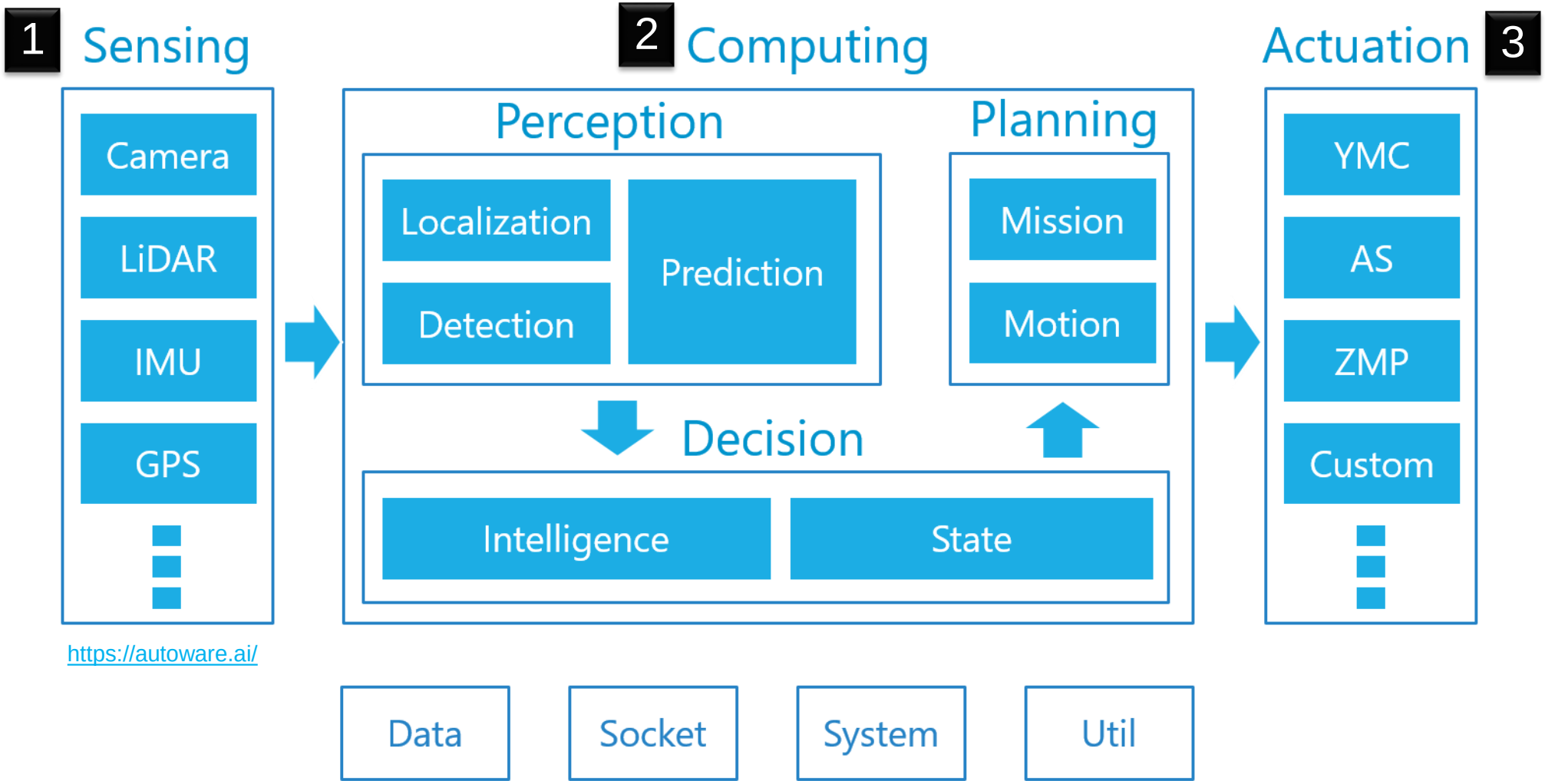
Singapore, 18 July 2018 – Grab, one of the most frequently used online-to-offline (O2O) mobile platforms in Southeast Asia, and the National University of Singapore (NUS) today launched an artificial intelligence (AI) laboratory to develop solutions to transform urban transportation and pave the way for smarter cities in Southeast Asia. The Grab-NUS AI Lab, which has been set up with a joint initial investment of S$6 million, is Grab's first major AI laboratory and NUS' first AI laboratory with a commercial partner.

Suntec Conference Center today

This is Ulf !

Autonomous Driving = Sense + Think + Act

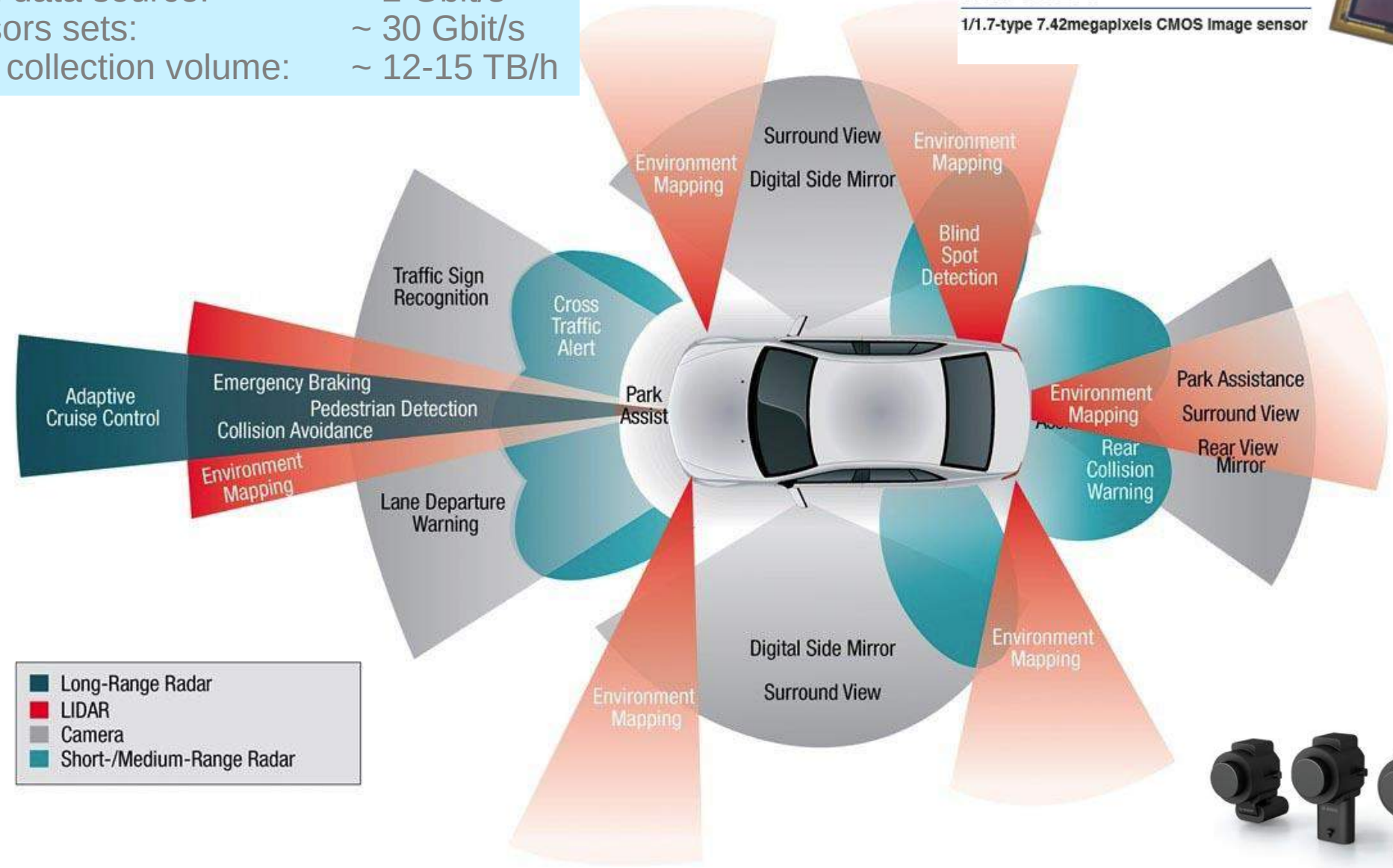
The Automotive Industry has to solve this highly complex problem.



https://autoware.ai/

Automotive Sensor Setup for AD

Each data source: ~ 2 Gbit/s
Sensors sets: ~ 30 Gbit/s
Data collection volume: ~ 12-15 TB/h



http://currencyobserver.com/2017/12/global-automotive-sensors-market-2017-2022/

AD Development and Test Vehicles from BMW

Inference AI Computing inside the car

Example: Car data logger at BMW

https://www.bmwgroup.com/en/innovation/technologies-and-mobility/autonomes-fahren/campus.html

More BMW details:
https://youtu.be/Hbm6IcD78R0

Example of Car Data Logger and Copy Station (by ViGEM)

The global data logger market identifies that the increasing demand for test, measurement, and control solutions will have a positive impact on the market's growth. The implementation of automation demands the need for data loggers that can record and communicate data which are either be connected wirelessly or by Ethernet/USB cables. To help reduce human labour in the work environment which will subsequently reduce errors, there is a growing demand for the adoption of cost-effective automated measurement equipment.

‚In Car' Data Logger

Lab Copy Station (Intel x86 PC)

https://vigem.de/en/content/cca-9002

Portable SSD 16TB

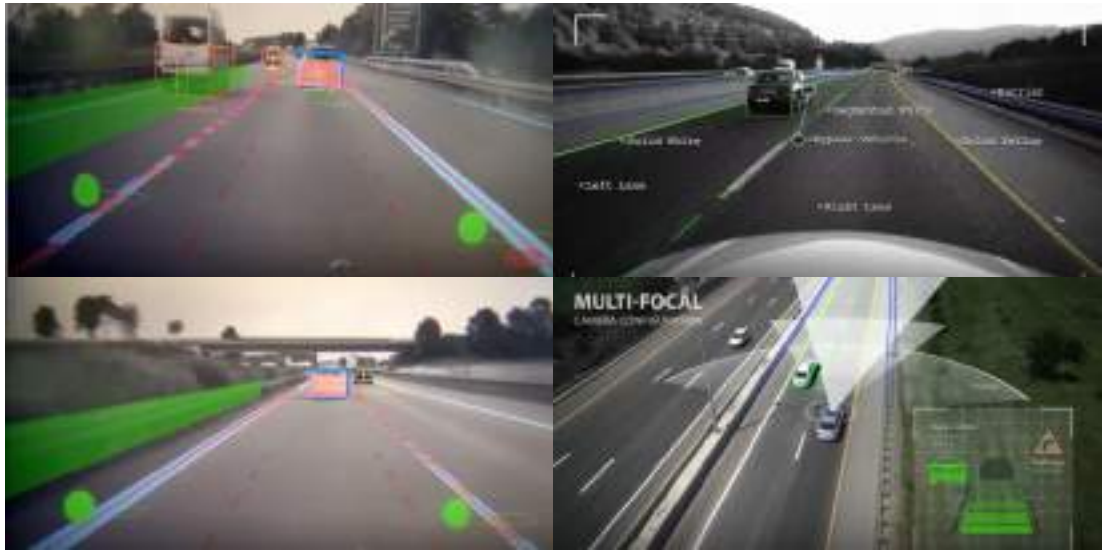
Top players: National Instruments Corporation, Ammonit Measurement GMBH, Dolphin Technology, Omega Engineering Inc, Omron, Testo, Vaisala, Onset HOBO, Rotronic, Dickson, HIOKI, Yokogawa Corporation, Sensitech, Fluke, ELPRO-BUCHS AG, Delta-T Devices, Grant Instruments, CSM GmbH, Kipp & Zonen, Gemini, OTT Hydromet, TTTech Computertechnik AG, Dwyer Instruments, Huato, more....

Lab based Copy Station for Data Logger

pictures by BMW

Automotive Industry generates large amounts of data

- Storage of data (sensor / video) is very **costly**.
- Handling of these data is **difficult** i.e. due to high required bandwidth.
- For testing purposes sensor / **video data** are much more complex in comparison to discrete bus signals, electronic values, etc.



Sources: Images from https://www.youtube.com/watch?v=4jW0fJ80VG8
https://www.youtube.com/watch?v=dhEgD6ZFlQE
https://www.youtube.com/watch?t=21&v=39QMYkx89j0

Sensor / video data must be synchronously captured, stored, modified and executed with other testing data such as CAN, FlexRay, Radar, LiDAR, HiSonic, etc. – most common formats are: **ADTF v2/3** (digitalwerk) **RTMaps** (Intempora) **MDF4** and **ROS/rosbag**.

Data Management for ADAS/AD

Test Drives

- Europe
- USA
- China
- Japan
- Asia
- Africa

50-70 TB / day / car

R&D Labs: tagging

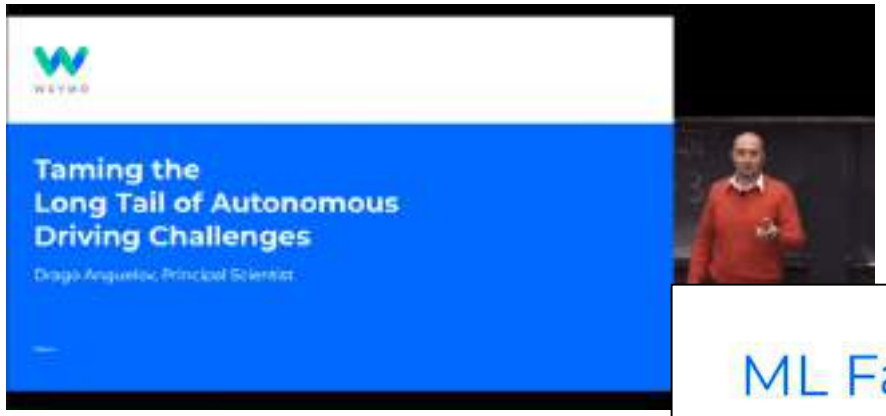
Labeling

Training Data as a Service (TDaaS)

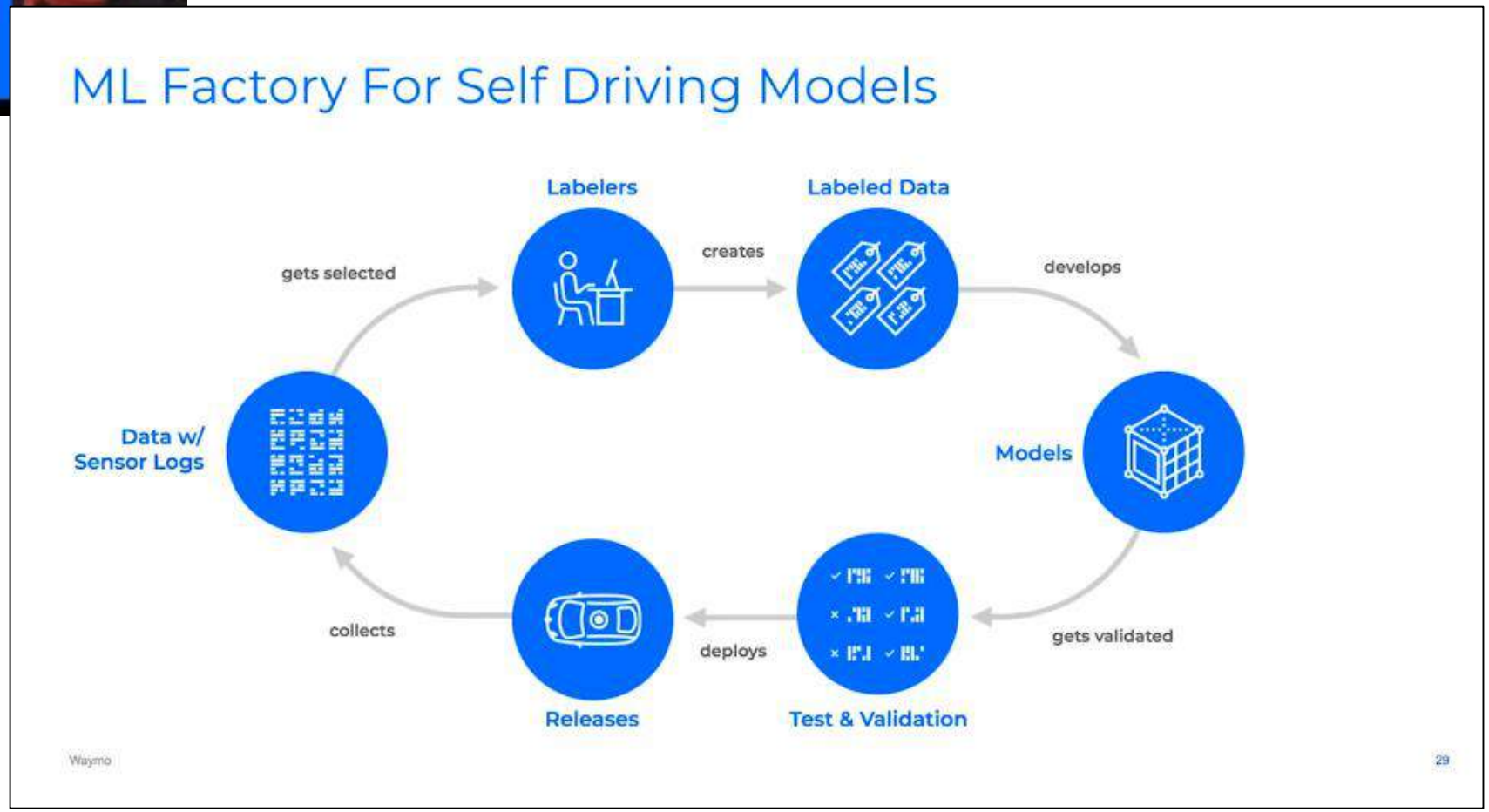
R&D Labs: developing & testing & (re-)simulation & AI training

- 300-500 PB data in total

Drago Anguelov from Waymo



Drago Anguelov (Waymo) - MIT Self-Driving Cars. (2019) YouTube. Retrieved 13 February 2019, from https://www.youtube.com/watch?v=Q0nGo2-y0xY



The AD Data Pipeline

INGEST

CLASSIFY

ANALYZE / TRAIN

Data In

Edge

Insights Out

Trained Model

Transient Storage
SDS/Cloud
Software Defined
Throughput-oriented

Global Ingest
Cloud
Capacity Tier
Throughput-oriented
Globally accessible

Fast Ingest / Real-time Analytics
SSD
High throughput
Performance Tier

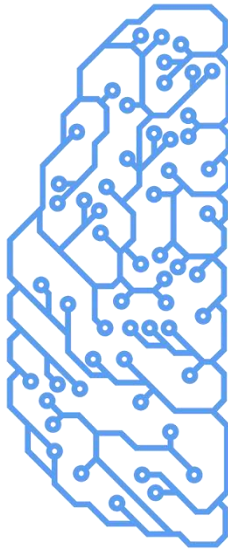
Classification & Metadata Tagging
High volume ingest & index
Automated tagging

ETL / Data Processing
SSD/Hybrid
High throughput
Small & large I/O
Performance & Capacity

Hadoop / Spark Data Lakes
Hybrid/HDD
Throughput-oriented
Performance & Capacity

ML / DL
SSD/NVMe
Hight throughput
Low latency
Small, random I/O

Archive
HDD *Cloud* *Tape*
High scalability
Large I/O
Sequential writes



The SpectrumAI Data Pipeline for AD

INGEST

CLASSIFY

ANALYZE / TRAIN

Data In

Edge

Insights Out

Trained Model

Transient Storage
Spectrum Scale / Elastic Storage Server (SSD models)
Cloud Object Storage

Global Ingest
Cloud Object Storage

Fast Ingest / Real-time Analytics
Elastic Storage Server (SSD models)

Classification & Metadata Tagging
Spectrum Discover

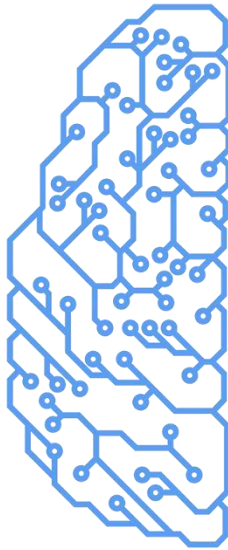
ETL / Data Processing
Elastic Storage Server (Hybrid models)

Hadoop / Spark Data Lakes
Elastic Storage Server (Hybrid models)

ML / DL
Elastic Storage Server (SSD / NVMe models)

Archive
Elastic Storage Server (HDD models)
Cloud Object Storage
Spectrum Archive (Tape)

IBM Spectrum Scale
IBM Cloud Object Storage
IBM Spectrum Discover
IBM Spectrum Archive



Converged Solution for Data Science Productivity

Introducing IBM SpectrumAI with NVIDIA DGX-1/2

A Scalable, software-defined infrastructure powered by IBM Spectrum Scale and NVIDIA DGX-1 systems. IBM SpectrumAI with NVIDIA DGX is the perfect engine for your data pipeline.

The workhorse of an AI data infrastructure on which companies can build their shared data service.

High-Performance to feed the GPUs

- NVMe throughput of 120GB/s in a rack
- Over 40GB/s sustained random read per 2U

Composable to grow as needed

- Up to 9 DGX-1 servers (72 GPUs) in a rack
- Storage scale-out from a single 300TB node to 8 Exabytes and a Yottabyte of files

Extensible for the AI Data Pipeline

- Support for any tiered storage, including Cloud and Tape

IBM SpectrumAI with NVIDIA DGX

Reference Architecture Building Blocks Description

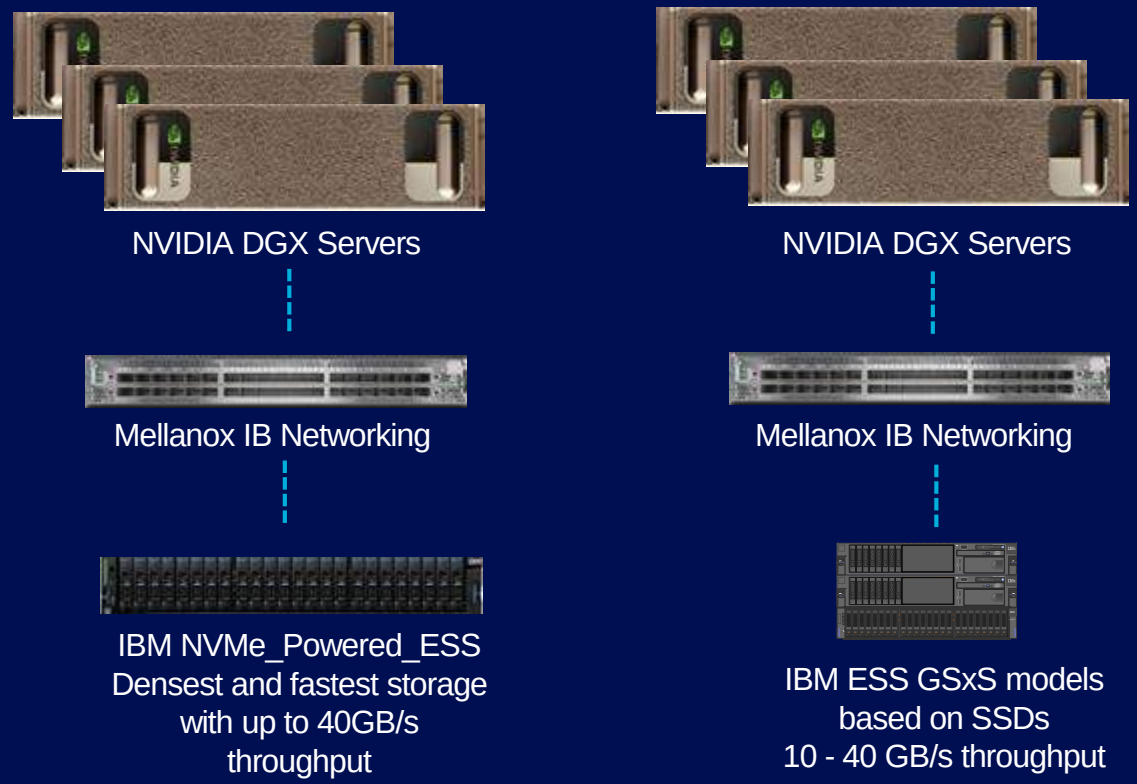
DGX-1 or DGX-2 Servers – purpose-built solutions for AI and machine learning, integrating eight of the world's most advanced data center accelerator – the NVIDIA Tesla V100 Tensor Core GPU

The NVIDIA DGX software stack, optimized for maximized GPU-accelerated training performance, including the new RAPIDS framework to accelerate data science workflow

IBM Spectrum Scale v5, the leading software-defined file storage, architected specifically for AI workloads with enhanced small file, metadata and random IO performance.

NVMe all-Flash storage for extremely low latency power efficiency and data density. Using IBM Spectrum Scale distributed data protection it delivers over 300TB in every 2U building block and 120GB/s of data throughput in a rack. (GA 2019)

Seamless data pipeline connectivity across multiple racks, other IBM SpectrumAI configurations, and workstations to provide the Data Scientists with a unified view of their AI data pipeline.



IBM SpectrumAI Reference Architecture

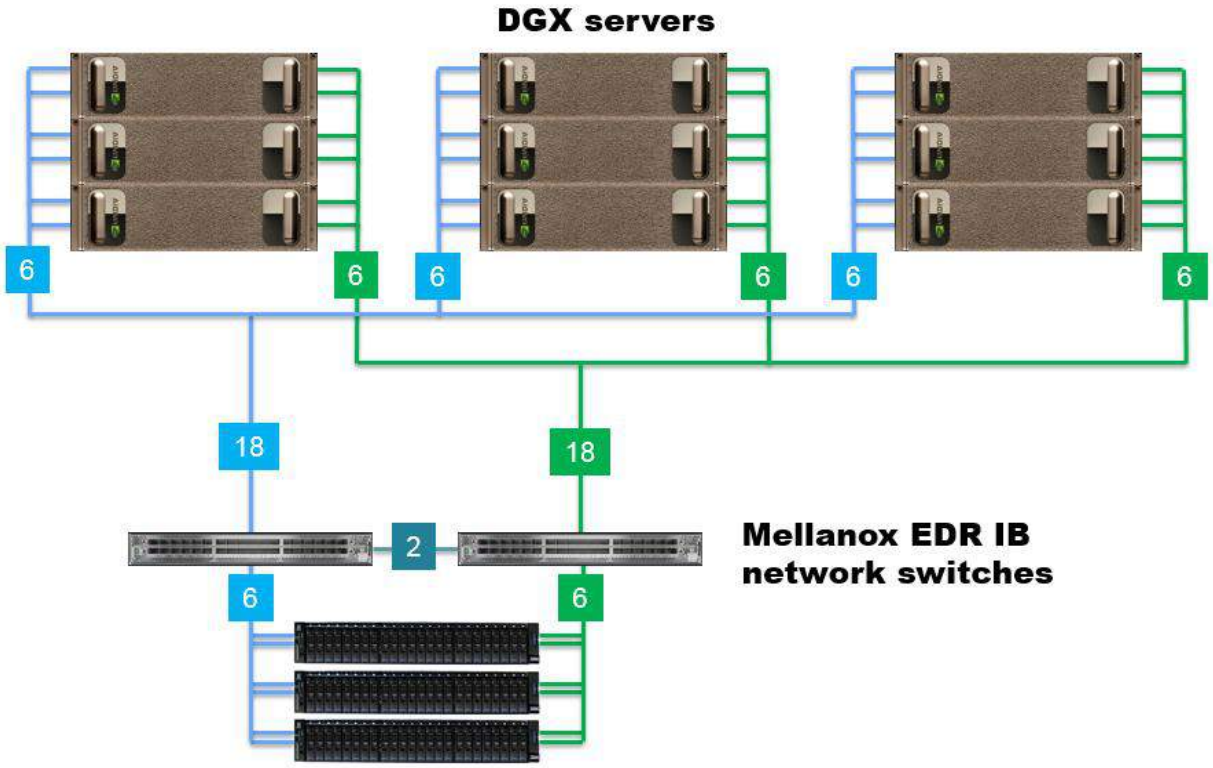
IBM Storage & SDI Solutions
Reference Architecture Paper
November 2018

IBM SpectrumAI with NVIDIA - Proven Infrastructure Solution for ML/DL workloads

High-Performance IBM Storage with NVIDIA DGX-1 servers ready to Power your AI data pipeline

IBM

NVIDIA



IBM Spectrum Scale

https://www.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=86022386USEN&

Reference Materials

AI in action: Autonomous vehicles
https://www.ibm.com/blogs/systems/ai-in-action-autonomous-vehicles/

IBM Storage Solutions for ADAS and Autonomous Driving (AD)
https://www.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=34019934USEN

IBM Big Data for Autonomous Driving
https://www.youtube.com/watch?v=eGhiIHDJaqI

IBM SpectrumAI Information
- Solution Brief
- Reference Architecture
- Benchmark Results

https://www.ibm.com/it-infrastructure/storage/ai

IBM Systems
Solution Brief

IBM

Highlights

- Leverage no-bottleneck architecture to scale performance for extreme throughput and low-latency access.
- Best-in-class storage infrastructure that grows and shares while automatically and quickly moving file and object data to optimal storage tiers.
- Enable data-anywhere access that spans storage and locations to accelerate ADAS data management across the data center or around the world.
- Leverage built-in authentication, encryption, security and replication options to meet ADAS data management and regulatory requirements.
- Cut costs up to 90% with intelligent archiving and automatic policy-based storage tiering from flash to disk to tape.
- Eliminate costly replication, significantly lowering the total cost of ownership for storage systems at the petabyte level and beyond.

IBM

IBM Storage Solutions for ADAS and Autonomous Driving

Extremely Scalable, Cost-Optimized ADAS Data Management and Autonomous Driving Development Infrastructure

The automobile is quickly morphing from an isolated, largely mechanical piece of equipment to one of the most technically sophisticated and connected platforms on the planet. From entertainment and navigation to driver assistance and crash avoidance, today's car is vastly different from those of a few years ago. The huge opportunity to share the future of that connected car, especially around autonomous driving, is drawing the interest of both technology and auto manufacturing companies.

The one thing these initiatives all have in common is data – miles and miles of data. Each sensor and system on connected cars generates a steady stream of information. The research and development behind future systems requires analysis of massive files and data sets. Dealing with the volume, velocity and variety of all this data creates a unique challenge.

The automotive industry is entering a new, highly competitive, transitional period where demand for new conveniences, safety capabilities and selling models are driving dramatic change. Once an industry consisting of pure hardware and adrenalin, automotive design is increasingly differentiated by software – with many visits to the dealership replaced by over-the-air bug fixes. At the forefront is Advanced Driver Assistance Systems (ADAS), which introduce disruptive requirements on engineering IT infrastructure – particularly storage, where even entry-level capacities are measured in petabytes.

Extreme scalability demands for autonomous driving (AD) development

Autonomous vehicle development requires a lot of data which is generated by the vehicle's hardware, including: a camera that generates 20-60 MB/s, sonar at 10-100 KB/s, radar upwards of 10 KB/s, LiDAR systems that range between 10-70 MB/s, and GPS that runs at 50 KB/s.

To put that figure into perspective, self-driving cars will consume and generate approximately 50-70 terabytes (TB) of data for every eight hours of driving.