Data-ingestion: Enabler for successfully developing autonomous vehicles

Martin Schleicher, Executive Vice President Business Management
Autonomous Driving Ingolstadt #3, December 11, 2018
Agenda

1. About Elektrobit
2. The need for large amounts of data
3. Acquiring data with suitable logging devices
4. Ingesting data into the development environment
5. Data processing
6. Summary
Elektrobit (EB)

**Technical competencies**
EB’s technical core competencies are development of automotive-grade (software) products and engineering services.

**Employees**
More than 2,200 employees* worldwide. Spans three continents and ten countries.

**Consistent growth**
Average growth (CAGR) > 10 %

**Global presence**
Development and business offices in Austria, China, Finland, France, Germany, India, Israel, Japan, Romania and USA.

**Continental AG**
Wholly owned, independent subsidiary of Continental AG.

**100+ million**
Over 100 million vehicles on the road and 1 billion embedded devices.

*May 2018, incl. Argus, excl. e.solutions.
1. About EB

Products and solutions

Automated driving
- Hardware and software products for development, test, visualization, and validation.
- Key software components to bring automated driving functions and systems to serial production.

Vehicle infrastructure
- AUTOSAR standard
- Single- & multi-core OS
- Functional Safety OS
- Embedded Security
- Automotive networks, e.g. Ethernet

Connected car
- Intelligent big data analytics & online diagnostics
- Scalable backend infrastructures
- Cyber security solutions plus modular add-ons by Argus
- Software updates over-the-air

User experience
- Navigation client for connected use cases
- Electronic horizon provider enabling map-based ADAS functions
- Model-based development of multimodal user interfaces
- Augmented reality solutions

Consulting services
- Consulting services for Functional Safety and Software Architectures
- Lean Software Development
- Established agile processes

Verification and validation
- End-to-end testing of complex embedded software systems
- Test concept development
- Independent verification and validation of software systems
The EB Assist product line at a glance

Hardware products

EB Assist CAR Box
- High-performant and reliable automotive-grade PC systems for testing and validation
- Data-logging, replaying, and simulation of real and virtual driving scenes

EB Assist bus tools
- Modular I/O slot cards, I/O interface modules, and simulation tools
- Built for highly precise data-logging, replaying, and simulation (rest bus simulation)

Software products

EB Assist Busmirror
- Tool for testing ECU software during implementation stage, both on hardware and on PC
- Supports all established bus systems, including Ethernet/BroadR-Reach, FlexRay, CAN, and LIN

EB Assist ADTF
- Tool for the development, testing, validation, and visualization of ADAS and AD systems
- Wide range of toolboxes available to extend its functionality

Test lab by EB
- Comprehensive driving scene database and management
2. The need for large amounts of data

Huge volumes of data for developing autonomous vehicles

• The prevalence of “data collection” vehicles which carmakers, Tier 1’s, suppliers, and technology providers are driving around to build up their data libraries are needed to:
  – Test algorithms
  – Train AI systems
  – Validate vehicles and components
  – Create higher fidelity simulation environments.

• The amount of data which is being collected can be mind-boggling
Growing complexity in sensor setup and supported functions

2. The need for large amounts of data

- Camera
  - Active park assist
  - Cross traffic assist
  - Adaptive cruise control
  - Traffic jam assist
  - Drowsiness sensors
- Short-range radar
  - Steering angle sensor
  - Adaptive cruise control
  - Traffic jam assist
  - Emergency braking
  - Long-range radar
  - Collision sensor
- Ultra sound
  - Tire pressure sensor
  - Wheel speed sensor
- Lidar
  - Long-range radar
2. The need for large amounts of data

Verification and validation requires a vast test coverage

- Complexity
- Unusual situations
- Hazards
- Difficult conditions
2. The need for large amounts of data

Assume 240 million kilometers without accidents

5% done with real driving equals 12 million test kilometers

95% done in simulation equals 228 million virtual test kilometers

- An average of 200 test cars with each recording 275 km/day
- For one year (220 working days)
- Producing 10 TB per day per car
- The whole fleet generating 2 petabytes of data per day

- 114 000 scenarios with 1 km length each
- 2 000 variants per scenario at an average of 60 km/h
- 158 000 days of simulation on one high performance computer
- 10 000 high performance computers running software and simulation take 15.8 days

At the same time, development and update cycles must get shorter. Virtualization enables parallel replay and simulation for accelerated testing.
Recording driving scenes: basic element for validation

Cloud-computing platform (public or on-premise)

Test data platform

Database

- Raw data
- Label ground truth
- Scenarios
- Test cases
- Test results
- Archive

Data enhancement

Test case management

Test execution

Test post-processing

Test drive recording (replaying OLT)

Scenario generation (simulation OLT / CLT)
3. Acquiring data with suitable logging devices

EB Assist CAR Box – product variant EB 9200

High-performant and reliable device for data logging, replaying, and simulation

- Automotive grade and robust PC system
- Highly flexible and customizable I/O configuration with EB Assist bus tools
- RAID storage system for up to 16 TB SSD
- Logging of sensor data and vehicle bus communication data during test drives
- Shock and temperature resistant
- Ready for mass production projects
- UPS optional
3. Acquiring data with suitable logging devices

Use-case: data logging solution for the test fleet of a Tier 1

Global test drive program with a three-digit fleet of vehicles requiring a data logging solution for ADAS/automated driving product validation.

**Constraints:**
- Developing and delivering operational solution in limited time
- A robust solution for in-vehicle usage covering all types of weather and road conditions
- Easy integration into existing test vehicles
- Ensuring worldwide delivery and usage

**Our solution:**
- EB Assist CAR Box variant EB 9200: robust, high-performant, and reliable automotive-grade data logging systems
- Leverages EB Assist ADTF, our tool for automated driving development.
4. Ingesting data into the development environment

Making the recorded data available to your developers
4. Ingesting data into the development environment

Common data ingestion methods

Uploading data via company network

- This method is an option for every company doing any autonomous driving solutions
- Limited upload bandwidth possibilities and has a higher staff cost to manage the logistics
- Data must be directly accessible by the development team
- Developers need to be able to share the same content simultaneously, which is critical from a performance and availability perspective
- When not needed anymore, data still needs to be preserved on lower cost storage, e.g. regarding regulatory topics
Common data ingestion methods

Using disk mail service

- Rented secure and private disks to be used during the vehicle recording process
- Devices for various needs, disk space ranging from several terabytes to a petabytes of data
- Service to have disks shipped back to a data facility to upload recorded data to the (on-premise) cloud
- Longer logistics lead times due to the drive ordering and drive return shipment times to/from the storage facility.
Common data ingestion methods

Using high speed leased data center facilities

• Highest bandwidth by allowing the recorded data to be directly uploaded on a private and secure high speed data connection to your cloud service

• Usually has highest cost of the presented options due to the leased data center services and staff needed to handle the logistics

• Enables your data to be available for your cloud services with the least amount of logistic lead time
Data is ingested. What’s next?

Various possibilities once the data is made available

• Sensor validation can be performed on each sensor as the data is brought into the cloud
• Raw sensor values can be abstracted into a scenario database for future test case scenario generation
• Data can be archived in the cloud along with locale attributes such as weather and driving conditions
• AI algorithms can be trained against the processed data
• Metrics can be made on the entire fleet being recorded
What does the workflow with the data look like?

5. Data processing

- Data recording
- Scenario generation
- Data enhancement
- Test case management
- Test execution
- Test post-processing

Dashboard
Recorded Time
> 3 years

© Elektrobit (EB) 2018 | Confidential
In this EB robinos example, you can see the capabilities of the Grid Validation Framework.
Recipe for enabling test and validation of AD systems

In Summary

• Getting ahead of the data challenges is the key to evade potential issues in development
• Decide on a path to follow when it comes to collecting, ingesting, and processing data
• Leverage proven-in-use software and hardware tools
• Cloud-based validation to efficiently manage the required millions of miles of driving scenes and petabytes of simulation data