AVL List GmbH (Headquarters)

# From Silicon Valley to the Test Bed: Bringing Big-Data Technologies into ODS

ASAM General Assembly 2018

Open Technical Seminar

**Dr. Sandi Pohorec**

# Agenda

- **Motivation**
- **ASAM ODS – an industry standard**
- **New challenges**
- **New technologies**
- **What ODS can offer in the area of Big Data**
- **ODS Big Data roadmap**
- **ODS Big Data goals for 2018**
- **ODS BD enables multi-vendor solutions**
- **ODS BD standardizes the best practices from end-to-end solutions**
- **ODS BD from a user view**
- **Summary**

# Motivation

> “Information is the oil of the 21st century, and analytics is the combustion engine”
>
> (Peter Sondergaard, Senior Vice President, Gartner)


**ASAM Open Data Services will enable scalable analytics by using the Big Data technology stack**

# ASAM ODS – an industry standard

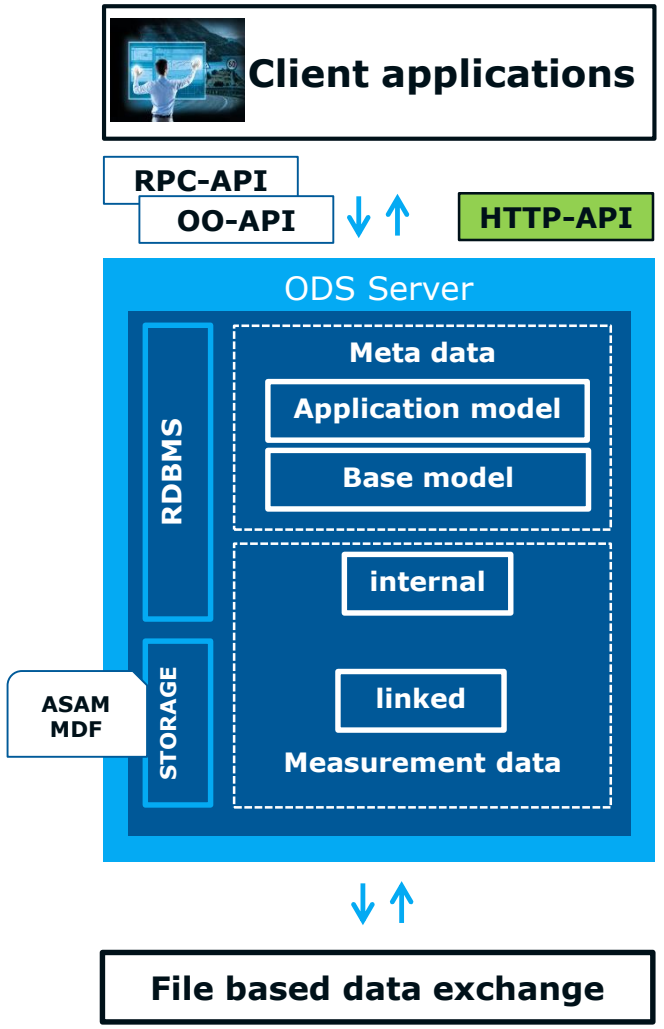
ASAM Open Data Services (ODS) focuses on **persistent storage** and **retrieval** of **testing meta** and **measurement data**

**Important features:**

- IT architecture independent data access
- Well designed balance between
  - High adaptability of the database
  - Support of well-defined different application scenarios
- API's and file based data exchange support

**Major advantages:**

- *Truly an **open** data service:*
  - *Vendor lock-in not possible*
- *Complete solution for test data management*



# New challenges


**Real driving emissions**

- More testing:
  - data **volume** grows **exponentially**
- More applications:
  - data **variety** and **variability** increase


**Electrification**

- Cost optimization:
  - **fusion** of **simulation** and **real-world** testing
- Responsive real-time analytics:
  - **velocity** goes from batch to **real-time**


**Autonomous driving**

**ODS needs to successfuly manage all challenges and enable horizontably scalable storage, search, processing, advanced analytics, machine learning and vizualization capability for all data in its domain**

# ODS Big Data roadmap

## Participants in 2017

OEMs

TIER 1

ODS BIG DATA TECHNICAL EXPERT GROUP

Tool Vendors / Service Providers

PLANNED

**Define** the **problem and collect** the **information on:**
- Big data use cases
- Scalable meta storage
- Scalable mass storage
- Near data processing

**Analyze** the **work packages** and **define** the **solution:**
- General approach and architecture
- Meta and mass storage
- Scalable near data processing

**Create** and **release** the **standard:**
- Denormalized meta data representation for indexing services
- Big data formats for ODS: PARQUET and AVRO
- Impact on and adaption of ODS 6

**Implement** and **improve:**
- Vendors start to implement the initial release of the standard
- Standard workgroup gathers feedback and plans refinements of the standard
- Standard workgroup works on extensions of the standard: streaming, complex data types, …

**2016** | **2017** | **2018** | **2019**

ACHIEVED

**Workpackages** have been **defined**

**RFQ** for **technical service provider** is published:
- extensive knowledge of both ODS and Big Data required to formulate a successful standard

**AVL, MHP and National Instruments** provide the technical expertise to the workgroup

Overall **architecture** encompassing all relevant aspects has been defined and **demonstrated** on AVL Santorin MX and AVL Concerto 5

Workgroups (meta data, mass data) are **active** -still possible to join and contribute.

# What ODS can offer in the area of Big Data

**Define interfaces to** existing Big Data technologies while **considering:**

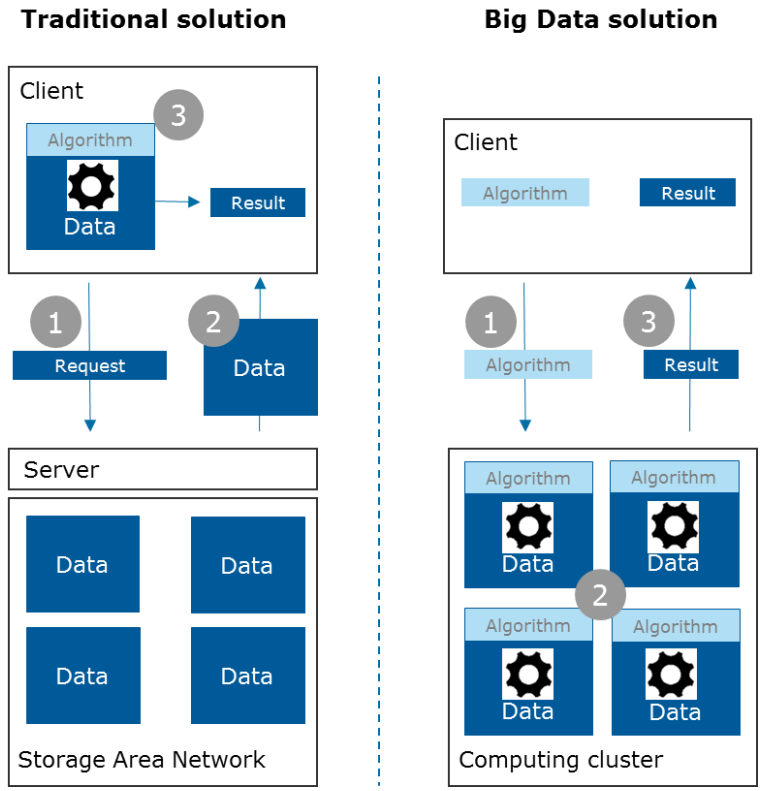
- Approaches for **scalable performance** for ODS: HPC vs Big Data
- Big data **ecosystem/map** and **system architecture**: architecture layers, Hadoop distributions and components stacks
- Current and future **data sources** in the ODS ecosystem: not "just" **time series** but also **complex object** data in many **formats**
- Current and future **use cases**: autonomous driving, …

**Provide** industry wide Big Data standard for ODS that brings consistent **interoperability** and **independence** from vendor lock-in

**Enable** the usage of powerful Big Data components and frameworks for:

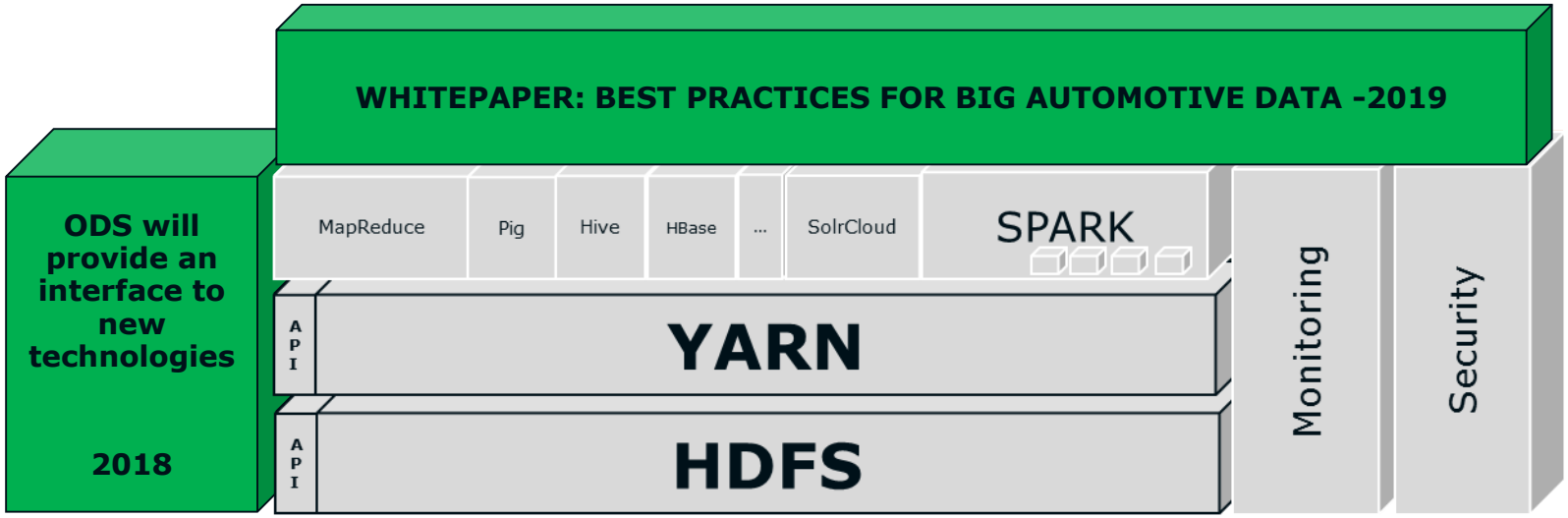
- Near data processing,
- Scalable data analytics,
- Machine learning,
- Advanced and highly interactive visualizations,
- Data analytics partnerships (OEM2OEM, OEM2Tier1)

**while keeping the benefits of ODS!**



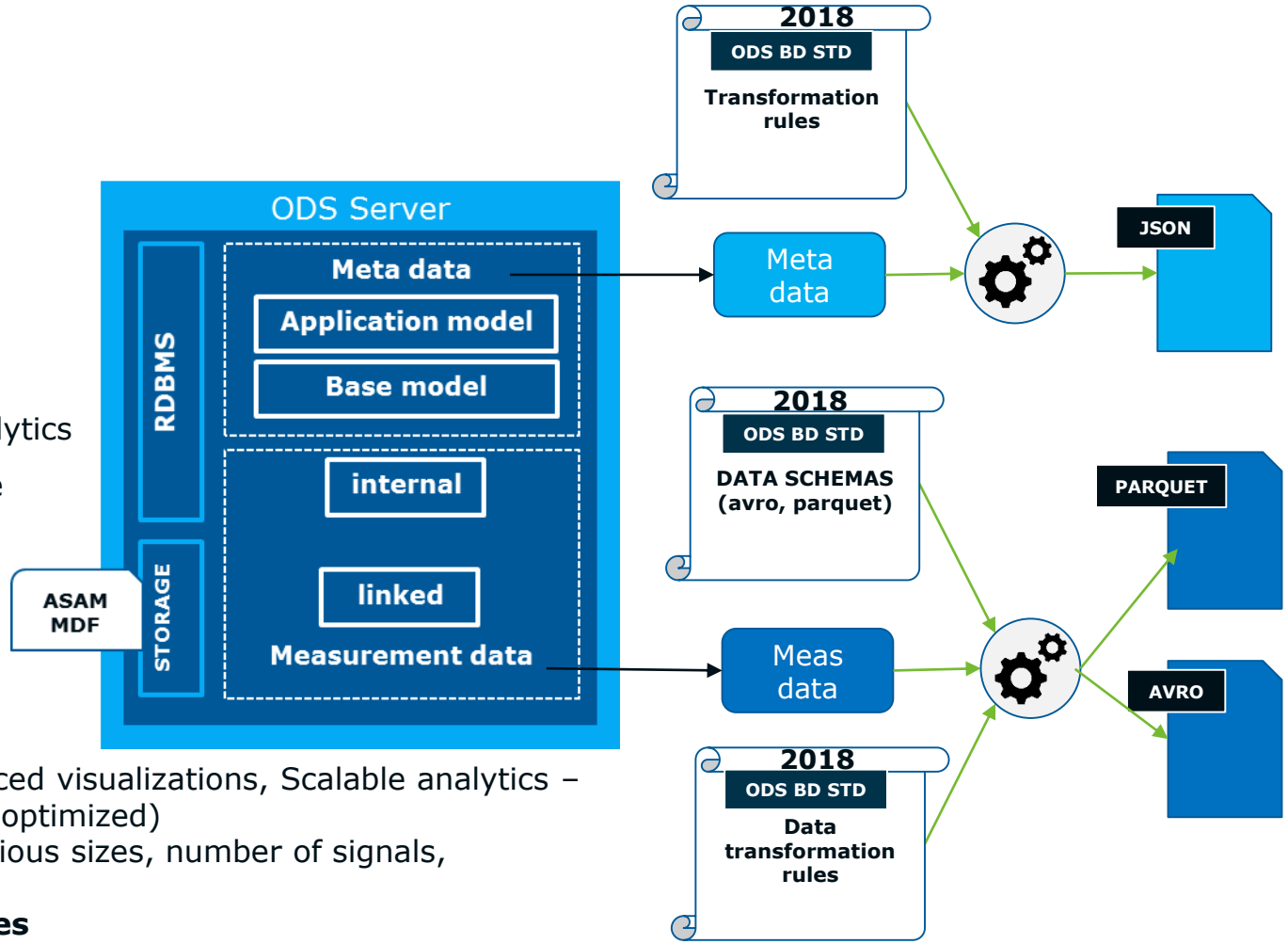
# New technologies: Hadoop

- Hadoop is a platform solution to manage the scalable storage and processing of very large data
  - Scalable storage (fault tolerant)
  - Distribution of data (also hot-cold)
  - Parallelization of the computation,
  - Assigning cluster resources, managing job execution
  - Batch data processing
  - Interface towards clients (API)
  - Numerous purpose-build frameworks
  - Scalable search
  - In-memory data processing (Spark):
    - Batch,
    - Streaming,
    - Graph processing,
    - Machine learning.
  - Enterprise ready with full support for security and monitoring (data lineage)



# ODS Big Data goals for 2018

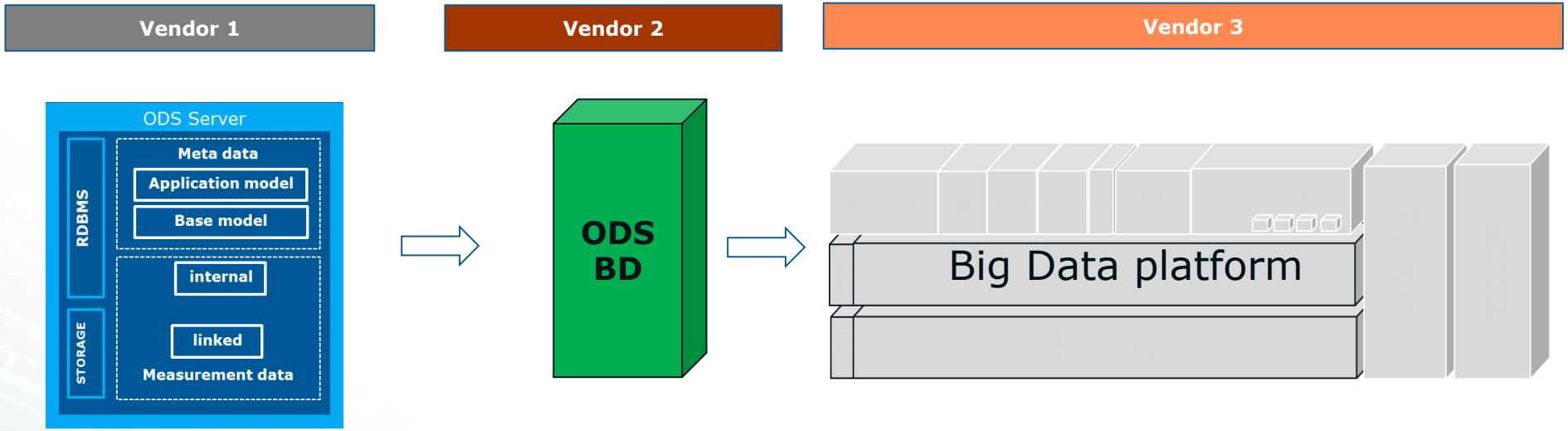
- **Searchable representation of meta data**
  - **Standardize**
    - Transformation rules for the denormalized representation
    - Industry standard JSON format
  - **Enable**
    - Enable utilization by non-ODS clients/tools
    - Choice of indexing technology/platform
    - Hor. scalabilty and HA – meta queries should not block analytics
    - Powerful matching capabilities with near real time response
- **Big Data format for measurement (time series) data**
  - **Standardize**
    - **Formats**
      - AVRO (row major format)
      - PARQUET (column major format)
    - For each format multiple **data schemas**
      - Different use cases (Random access for advanced visualizations, Scalable analytics – parallelism per file/per channel/…, Write/Read optimized)
      - Different properties of measurement data (Various sizes, number of signals, One/multiple sampling frequencies)
      - For each schema defined **transformation rules**
  - **Enable**
    - Horizontally scalable storage and near data processing of ODS measurement data on Big Data clusters



# ODS enables multi-vendor solutions

**Components from different vendors can be used for each step**

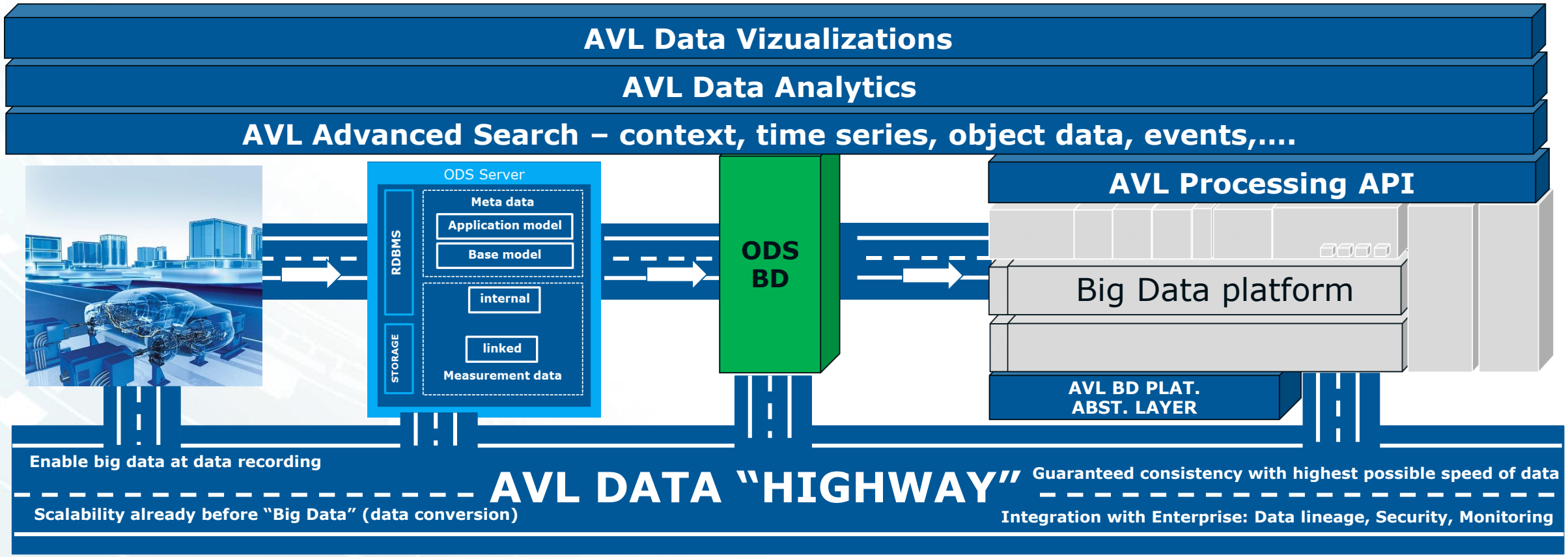
Standard only defines the **interfaces** and **data formats/schemas** – vendors **are free to use different approaches** how to best satisfy the NFRs: **performance**, **scalability**, **reliability**, **availability**, **data integrity** and **security**




Enable big data at data recording

# ODS standardizes the best practices from end-to-end solutions

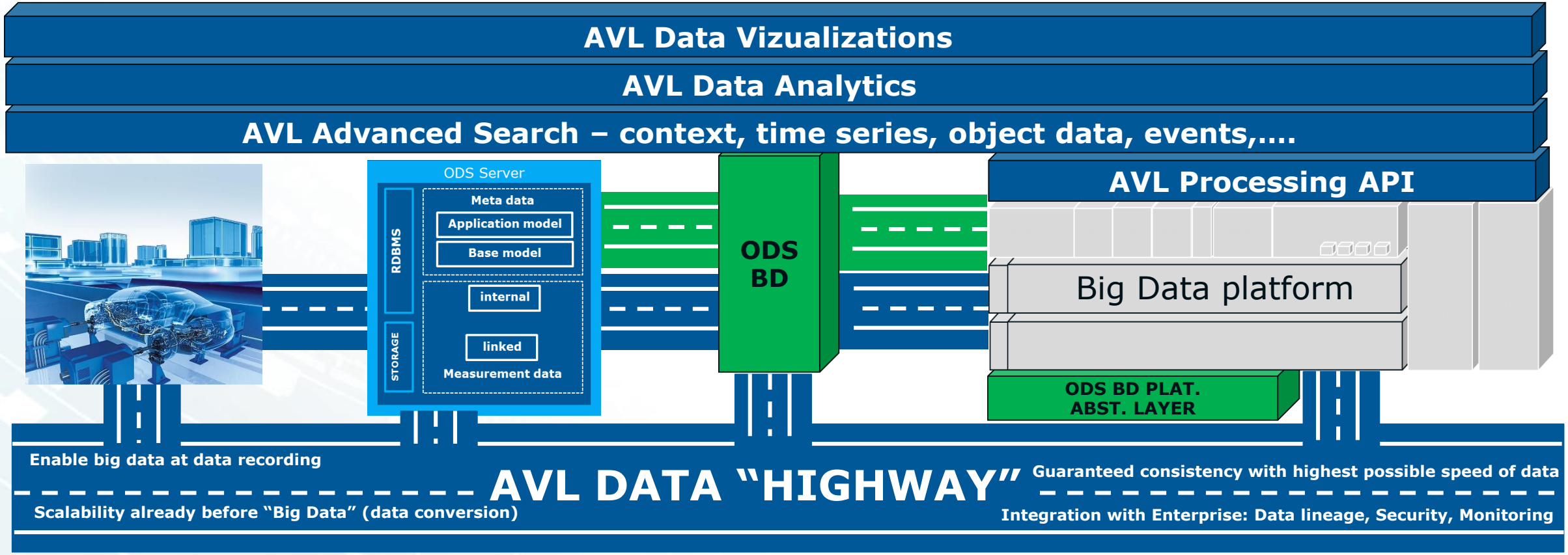
**AVL products form an end-to-end solution which provides more value that a combination of individual components: highest consistency, speed, integration, enterprise readiness, advanced data search of all data types, advanced analytics capability with insightful data visualization in near-real time.**



# ODS standardizes the best practices from end-to-end solutions

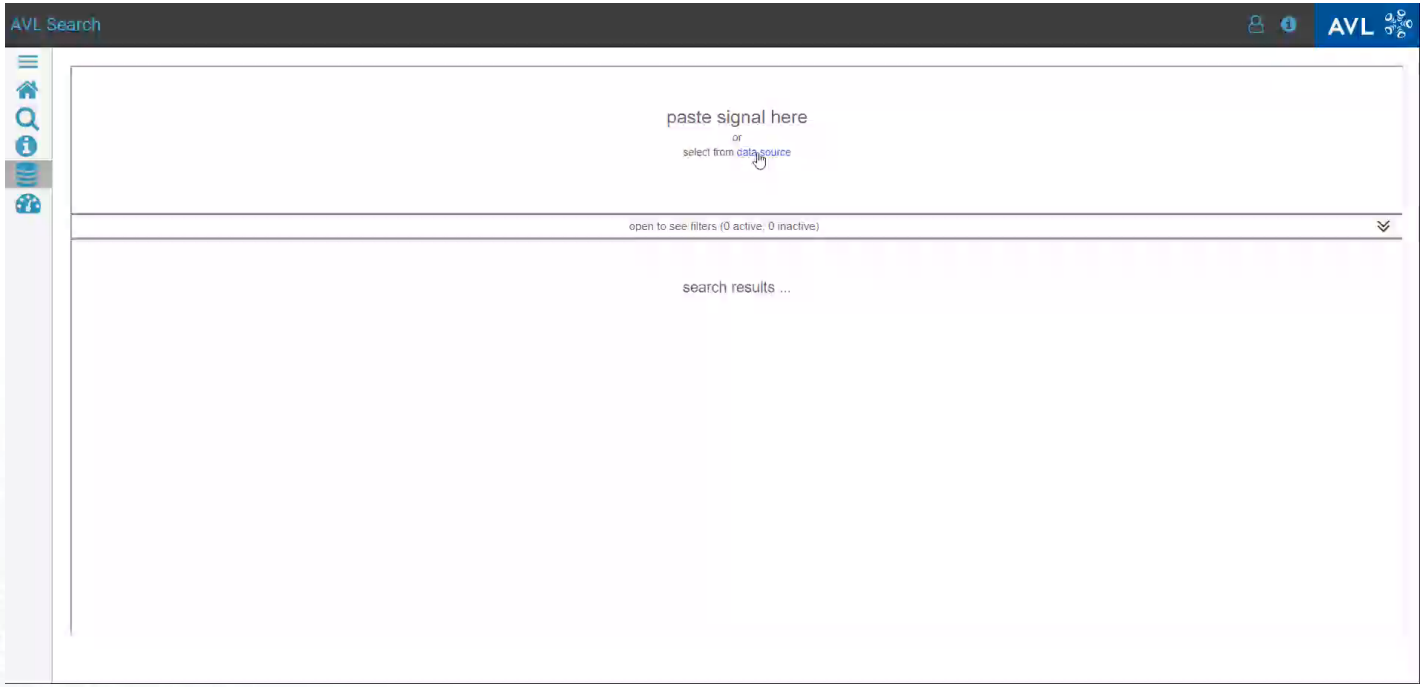
**AVL products form an end-to-end solution which provides more value that a combination of individual components: highest consistency, speed, integration, enterprise readiness, advanced data search of all data types, advanced analytics capability with insightful data visualization in near-real time.**

AVL will bring the knowledge from its products and projects into the ODS Big Data standard



# User view on the ODS Big Data

**Search for measurements simply by selecting an interesting part of a signal**



# Summary

- **Big-Data is a top-level requirement for ODS**
- Workgroup for ODS Big Data was formed in **2016**
- In 2017 workgroup got to a common understanding
  - Major considerations of scalability
  - Focus points for standardization
  - **Result: Architecture of an ODS solution (ODS server + Big Data platform) and the interfaces**
- In 2018 the workgroup is working on specifying the first artefacts:
  - **Meta data transformation and denormalization** to industry standard JSON form
  - **Mass data transformation and conversion** to the most important formats in the Big Data ecosystem: avro and parquet
- **The work does not stop in 2018**, the standardization will **continuously cover more and more aspects** of the automotive big data world
  - Based on best practice from vendors
  - AVL was leading the technical supplier group in 2017 and will continue to bring its knowledge and experience to the standard


Thank You
AVL
www.avl.com