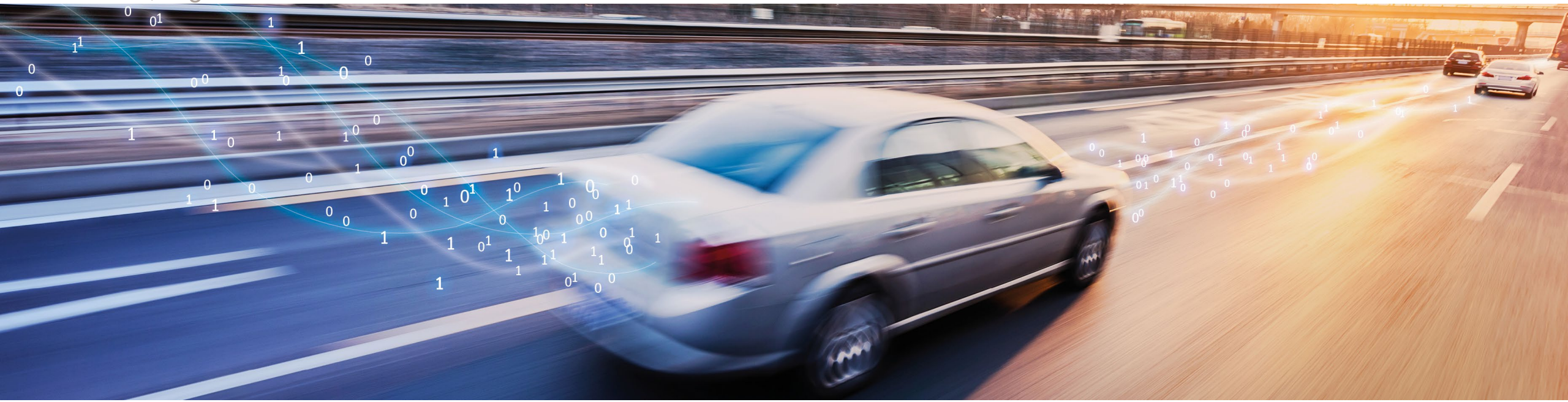
The Road to ASAM ODS 6.1

Introducing Big-Data features into ODS

Dr. Ralf Nörenberg
Board Member, ASAM e.V.
CEO, HighQSoft GmbH

Novi
October 2019



ASAM
Association for Standardization of Automation and Measuring Systems

ASAM ODS

A brief Introduction (of ~1000 pages of Standardization results)

ASAM Standards

Domains and Portfolio

Data Management & Analysis

CEA | ODS

Measurement & Calibration

CDF | CPX | MCD-1 CCP / XCP | MCD-1 POD | MCD-2 MC | MDF | MCD-2 CERP

Test Automation

ACI | ASAP 3 | ATX | GDI | MCD-3 D | MCD-3 MC | XIL

Diagnostics

MCD-2 D | MCD-3 D | OTX Extensions

Software Development

CC | FSX | ISSUE | LXF | MBFS | MDX

ECU Networks

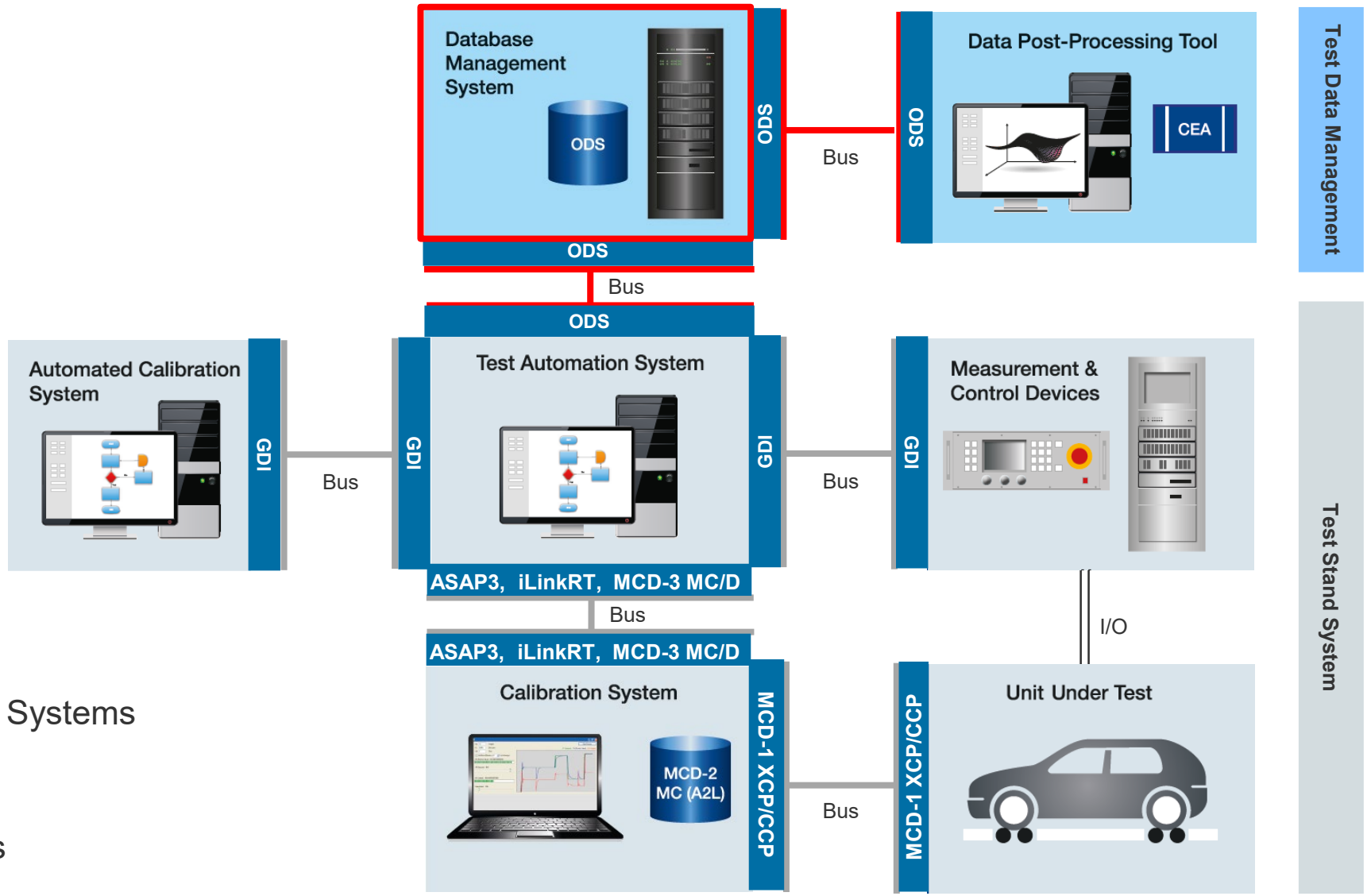
MCD-2 NET

Simulation

Open* standards

ASAM ODS in Test Systems

Application in a Test System

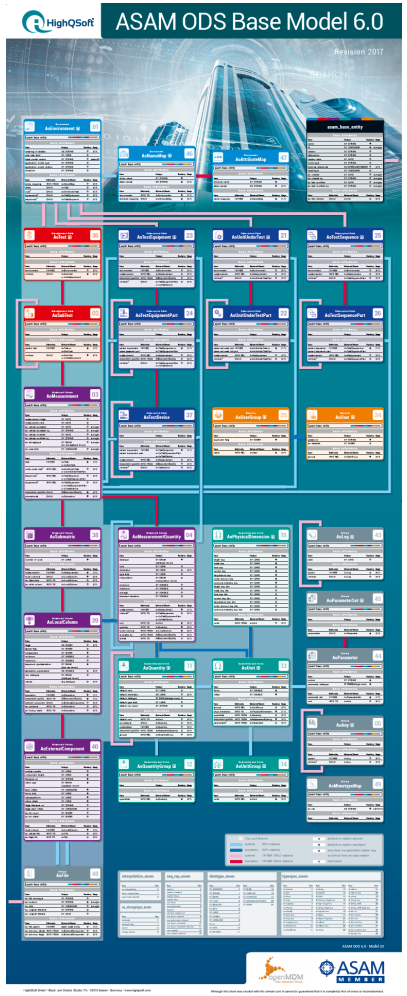


Data Sources are

- Test Stands / Automation Systems
- Vehicles
- Sensors and Devices
- Produced files or streams

ASAM ODS in Scope

What does the standard define?



Content: defines a common basis for data generation, storage and analysis

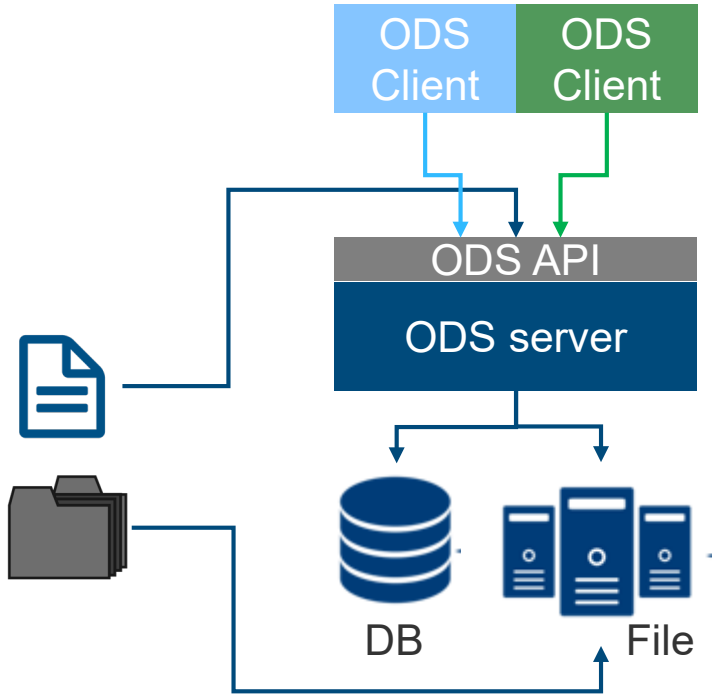
- …is a world-wide standard for more than 25 years
- …is supported & utilized by most automotive OEMs world-wide

Definitions: The ASAM ODS Standard defines

1. Physical Data Storage (ASAM ODS Base Model)
2. Standardized Access Layer (ODS API)
3. Data Exchange Format (ATFx File Format)

ASAM ODS data is categorized into two fields:

- …meta data / describing data [database, kilobyte]
- …mass data / channel data [file, terrabyte]



ASAM ODS 5.3.1 – Why moving forward?

Year 2016

ASAM

ASAM ODS 5.3.1 – Why moving forward?

Acknowledged restrictions

ASAM ODS access is realized with an **API** (access method) and a **serialization protocol** (mass data transfer).

1) Serialization Upgrade required: “CORBA dependency”

- Technology that is no longer further developed
- Firewall problems with enterprise situations
- “Out-of-technology” (less and less developers know the problem)

2) API Upgrade: New and alternative technologies on the market

- WEB and REST are far more common and flexible for interfaces and data retrieval

3) “Big Data Integration” vs. the defined physical storage (SQL and file)

- SQL and file formats as a potential limitation: New and alternative technologies are available
- “Big Data” may require alternative data sources and data access methods for data analysis

>>> Further development is required! <<<

ASAM ODS 6.0.1: A new development

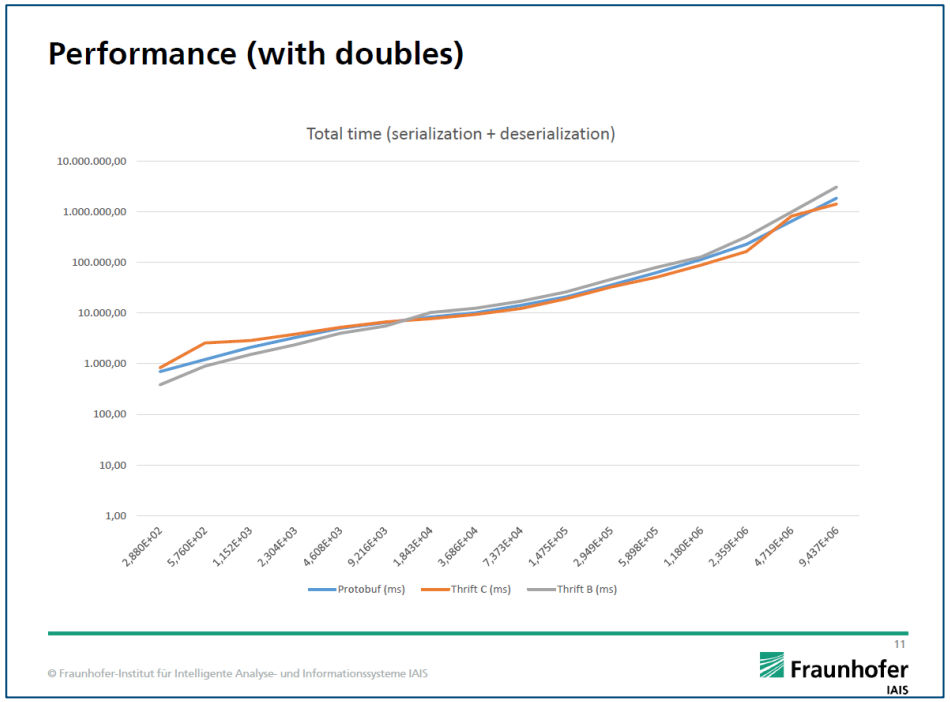
Developing a major upgrade (2016 - 2018)

ASAM

ASAM ODS 6.0.1: Choosing the technology - Step 1

General investigation with two favorite candidates

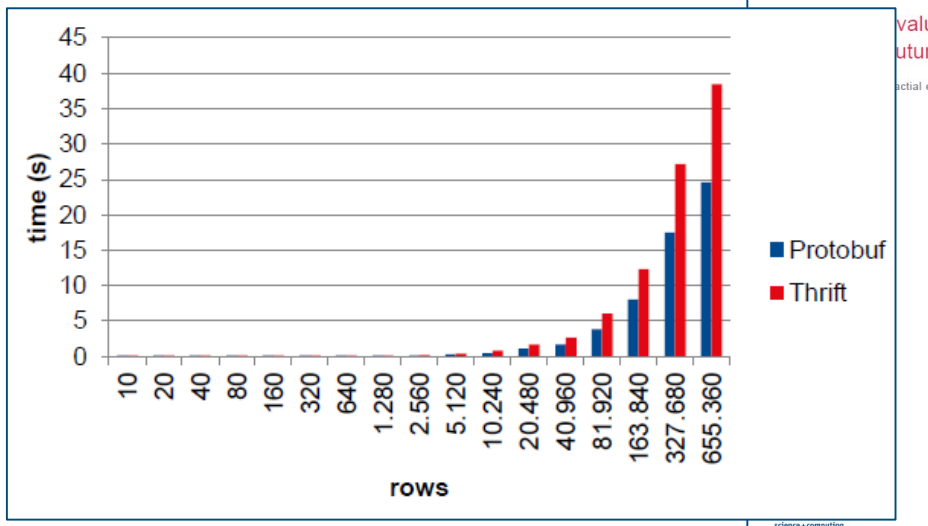
- **Technology Study by Fraunhofer (API / Serialization)**
  - Candidate 1: “Thrift / Thrift”
  - Candidate 2: “REST (HTTP) / Protobuf”
- **Criteria**
  - Language support
  - Encoding support
  - API / RPC features
  - Multiple services
  - Performance (see right)
  - Security
  - Extensibility
  - Entry level for suppliers
  - Alignment with web technologies
- **Decision**
  - Although Thrift was recommended by Fraunhofer (slightly higher “score”)…

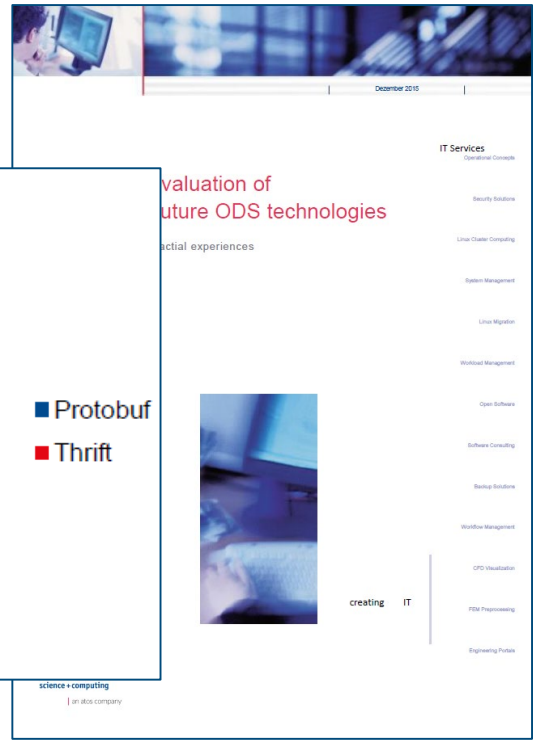


ASAM ODS 6.0.1: Choosing the technology - Step 2

Specific investigation of the two favorite candidates

- **Practical Implementation Study by science + computing**
  - Candidates remain the same: “Thrift / Thrift” vs. “REST (HTTP) / Protobuf”
  - Study is more detailed and with ODS related use-case





- **Criteria (more ASAM ODS related)**
  - Language support
  - API / RPC features
  - Multiple services
  - Performance
  - Extensibility
  - Alignment with web technologies

- **Decision**
  - Although Thrift was recommended by Fraunhofer…
  - …and after a practical implementation experience (by science + computing) and long discussions…
  - REST / Protobuf were chosen

ASAM

REST (API)

Overview and Implementation

REST (Representational State Transfer) API

- can be **used** over nearly any protocol, it usually takes advantage of HTTP
- Client Server: Separation of responsibilities of Client (UI) and Server (Data Management)
- Unique URLs for resource identification
- Processing of resource independent of representation (XML, HTML, JSON, ..)
- Layered System: Other systems can exist between client and server (e.g. for load balancing)

>> Simplified Handling with HTTP-Functions: GET, POST, PUT, DELETE

Summary & Benefits

- ODS 5.3.1 OO-API has: > 200 methods
- ODS 6.0 HTTP-API has: 30 functions
  - **15 common functions** (Connection Handling, Descriptive and Measurement Data, Transaction Handling)
  - **15 specific functions** (Application Model modification, Security Administration, Event Notification, Misc.)

>> Reduction of functions simplifies utilization and reduces error-proneness (e.g. performance)

Protobuf (Serialization)

Overview and Implementation

Targets

- Serialization of structured data
- Simplicity and performance

Protobuf

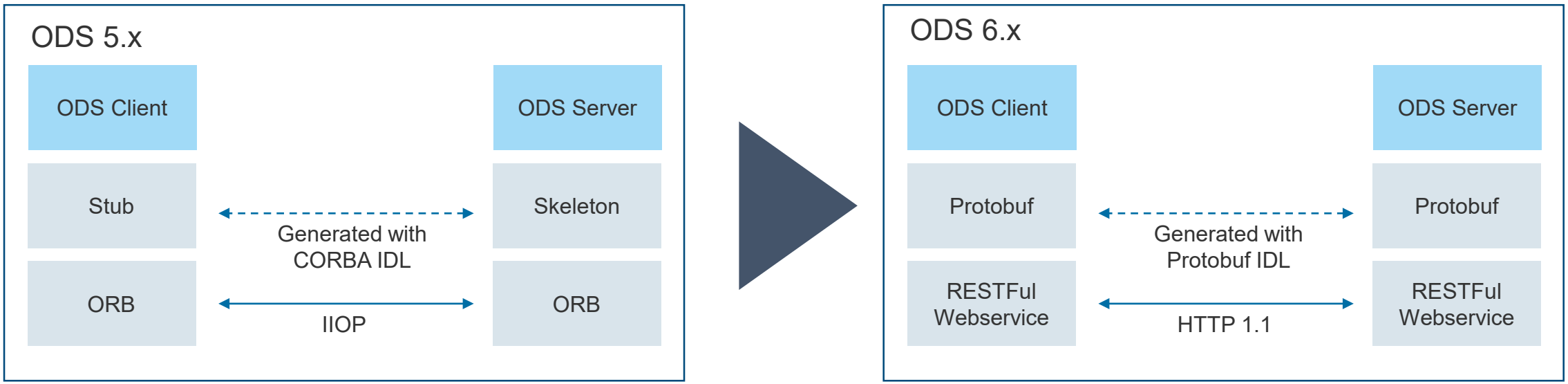
- Development of Google since 2001(starting 2008 Open Source with Apache 2.0 license)
- Platform- and language independent (Java, C++, Python, JavaNano, Ruby, Objective-C, und C#)
- Binary format in contrary to XML und JSON (3-10x smaller messages and 20-100x faster)

Summary & Benefits

- “De-facto standard” by Google
- Allows utilization with modern firewalls (one port required)
- Works as per expectations (performance and implementation evaluations)

ASAM ODS 6.0.1: REST and Protobuf

Overview



ASAM ODS 6.0.1 is fully compatible with the previous ODS 5.3.x standard.

Code examples are available for C#, C++, Python, JAVA, JAVAnano

ASAM ODS Big Data (“BODS”) – Road to ODS 6.1

Starting as a parallel working group (2017)

Proposal, Use-Cases and Implementation

Overview

Phase 1: Ideation and Proposal Description

- Allowing big data technologies to support huge measurements within ODS
- **Approach A:** API definition to access data
- **Approach B:** Storage definitions to fix structure of data

Objective: Being free to choose…

- …indexing services (Solr, OrientDB, …)
- …scheduling and resource management technologies (e.g. Spark, …)
- …data processing languages (e.g. Python, Java, …)

Phase 2 ”Concept Work” and **Phase 3 “Implementation”** Validation and implementation based on Use-Cases and Non-Functional Requirements and

- WP 2 – Mass data storage in HDFS >> **JSON**
- WP 3 – Definition of processing access layer
- WP 4 – Context data storage in HDFS >> **Parquet**, **AVRO**
- WP 5 – Findings and preparation of next steps

Initial Participants in 2017

- Audi
- AVL
- BMW
- Bosch
- Cummins
- ETAS
- Ford Motor Company
- GM
- HighQSoft
- IASYS
- Müller BBM
- National Instruments
- Peak Solution
- PSA
- RD Electronic
- Vector
- White Pine

ASAM ODS 6.1 (“One Working Group”) - Status

Consolidation of Working Groups (2018 - 2020)

Meta Data File Format: JSON

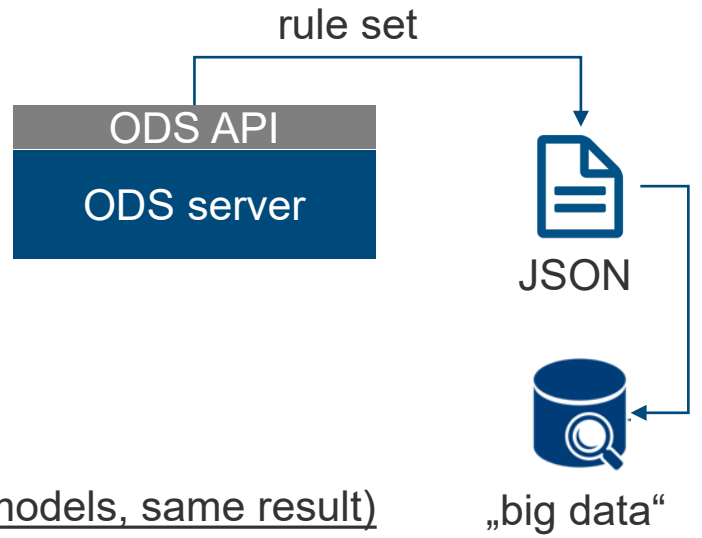
Overview

ODS Instance Data is META data that is stored on an ODS server. The instance data provides semantic knowledge and describes MASS data, which is also managed by the ODS server.

An instance data format is defined for replicating instance data from an ODS server into a target instance data store.

The target instance data store is available in a big data environment

- Instance data retrieval (e.g. in a key-value store, document database)
- Instance data search (Search engine)
- Further instance data processing (Data mining, e.g. Graph DB)
- Instance data consolidation (Data warehouse e.g. HBASE)



Result

- Export rules that define the content of meta information within the file (different models, same result)
- Definition of the JSON file including the META information
- JSON file definition is “ODS oriented” and not “index oriented” and thus may require a further step

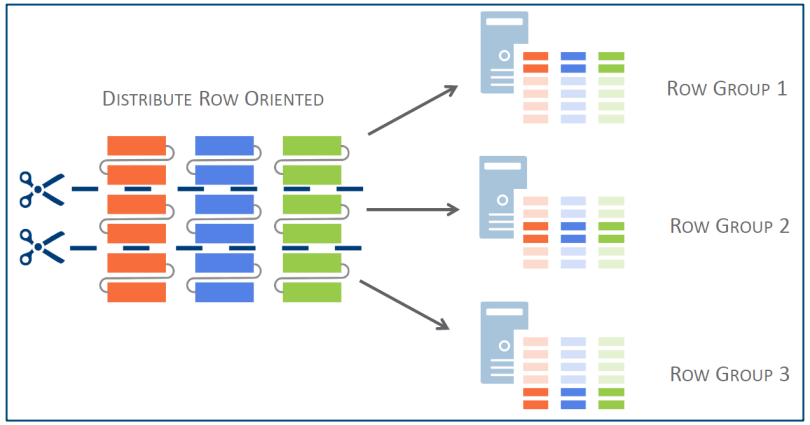
Mass Data File Format: Parquet

Overview

Apache Parquet is a leading column-oriented format for the big data ecosystem. The main purpose is to store the files in Hadoop. Storage of physical data, raw data and flags is possible. No storage of meta data.

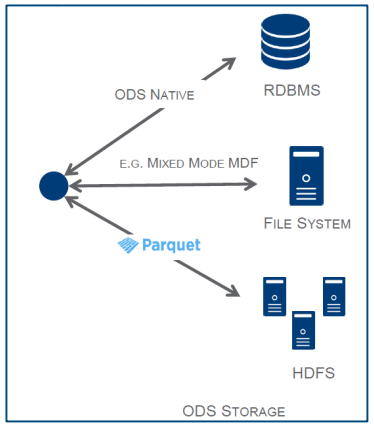
Definition of the Parquet POINT schema (“flexible & distributable”)

- Every “AoMeasurementQuantity” has got an own column, rows are corresponding to one value.
- A measurement can be distributed over multiple files (horizontal & vertical slicing is possible)



Parquet Packed schema (“compact”)

- Fixed structure in the PARQUET file that can be enhanced by customer requirements.
- Multiple values can be stored in one row of “AoMeasurementQuantity”
- Compressed Information is possible



Result

- Flexible mass data storage is defined for utilization in Hadoop systems

Transport File Format: AVRO (Meta and Mass Data)

Overview

Avro has been designed as a language neutral data serialization system. It is not primarily intended for use in data analytics but in data serialization and transport.

ASAM ODS decided to support Avro as it can bridge the gap between the ODS Server and Big Data ecosystems.

AVRO Packed schema

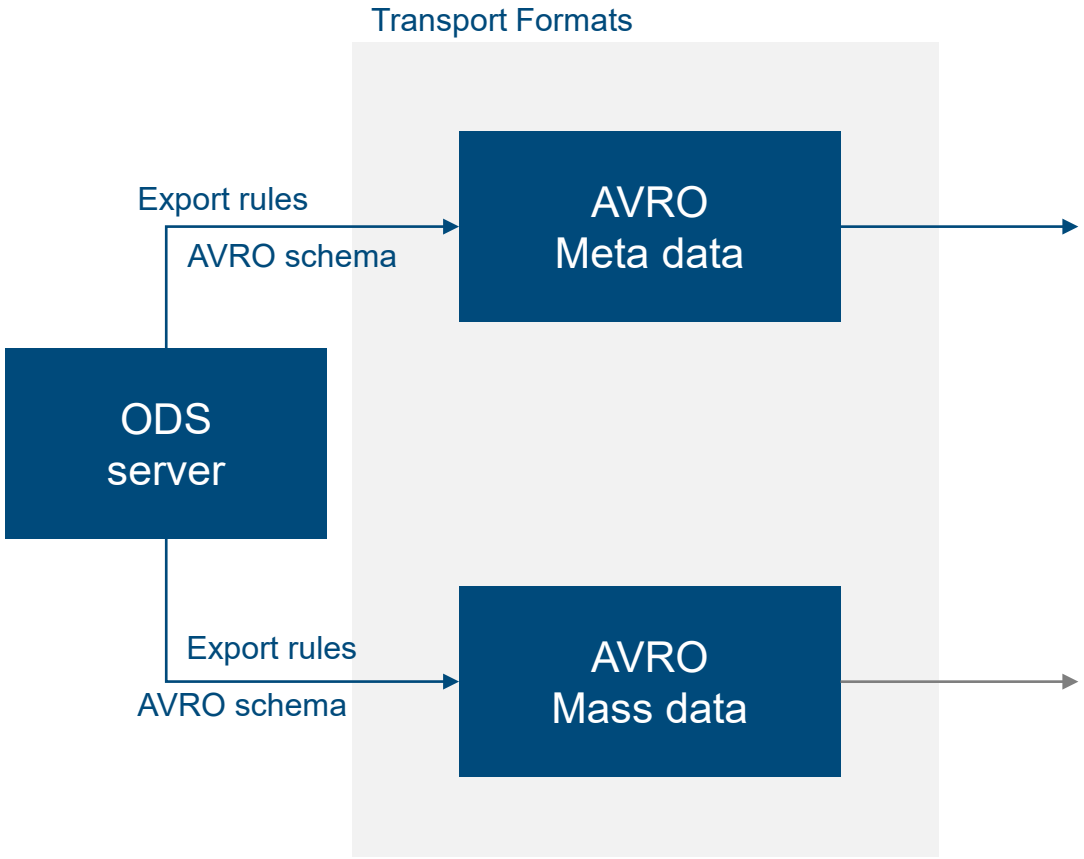
- Contents of the AVRO file are identical to the “Parquet Packed scheme”
- The benefit of the AVRO scheme are:
  - … it is write-optimized; it may be used e.g. for transport of mass data.
  - … its availability on many platforms.
- (The drawback of AVRO is that it is not read-optimized; its performance drops if only a few information items are used for analysis)

Results: Definition of the AVRO schema for

- meta data transport and transformation into JSON (optional)
- mass data transport and transformation into Parquet (optional)

The full ODS 6.1 example: Generating data in urge…

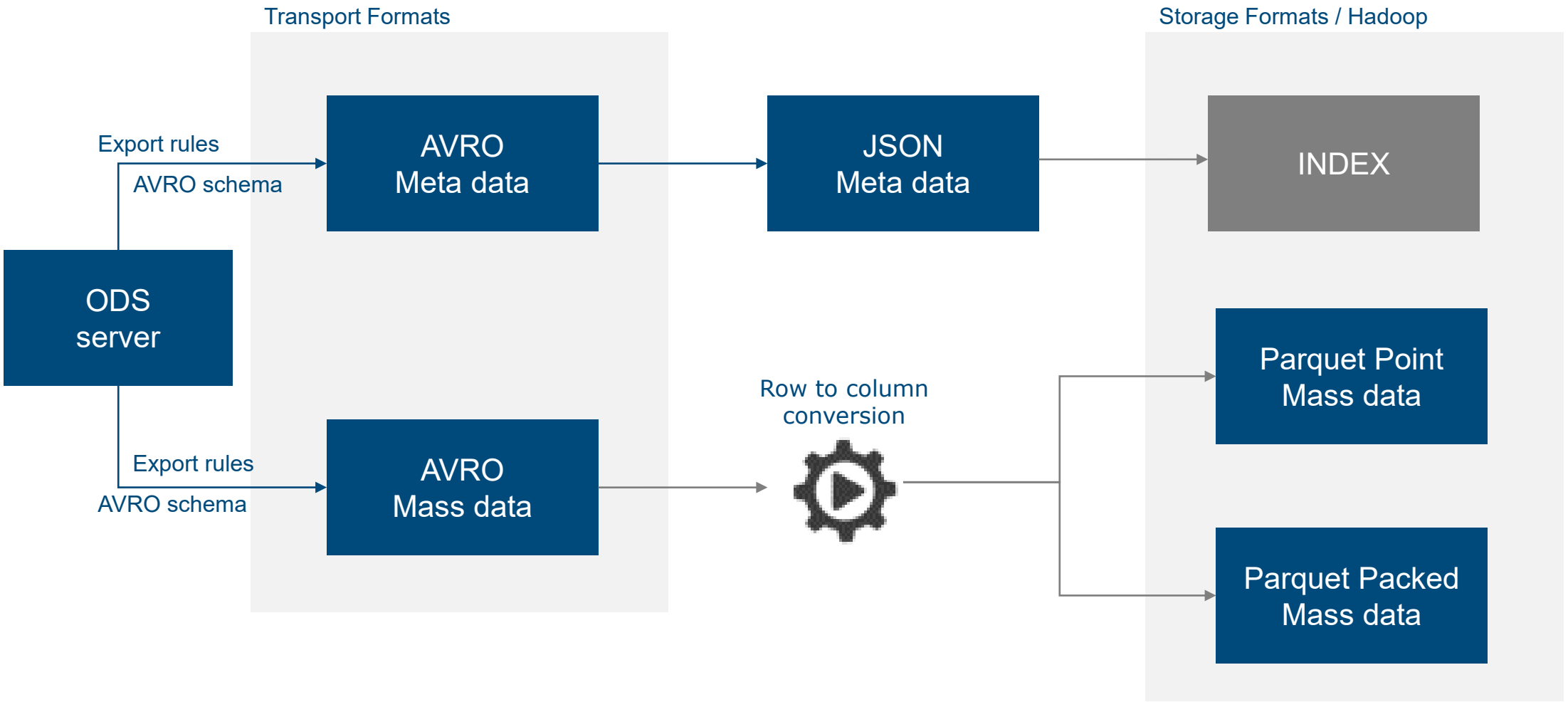
Overview



Producing data quickly in a „must write fast“ scenario…

The full ODS 6.1 example: … to storing it in a ”big data” system

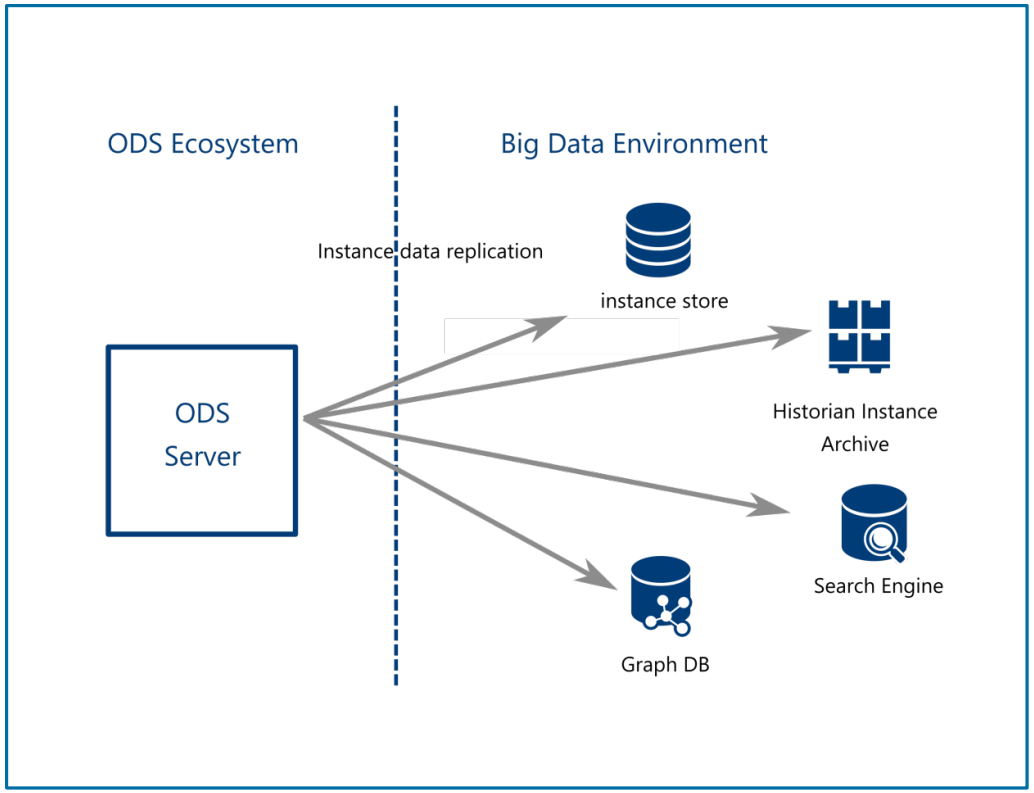
Overview

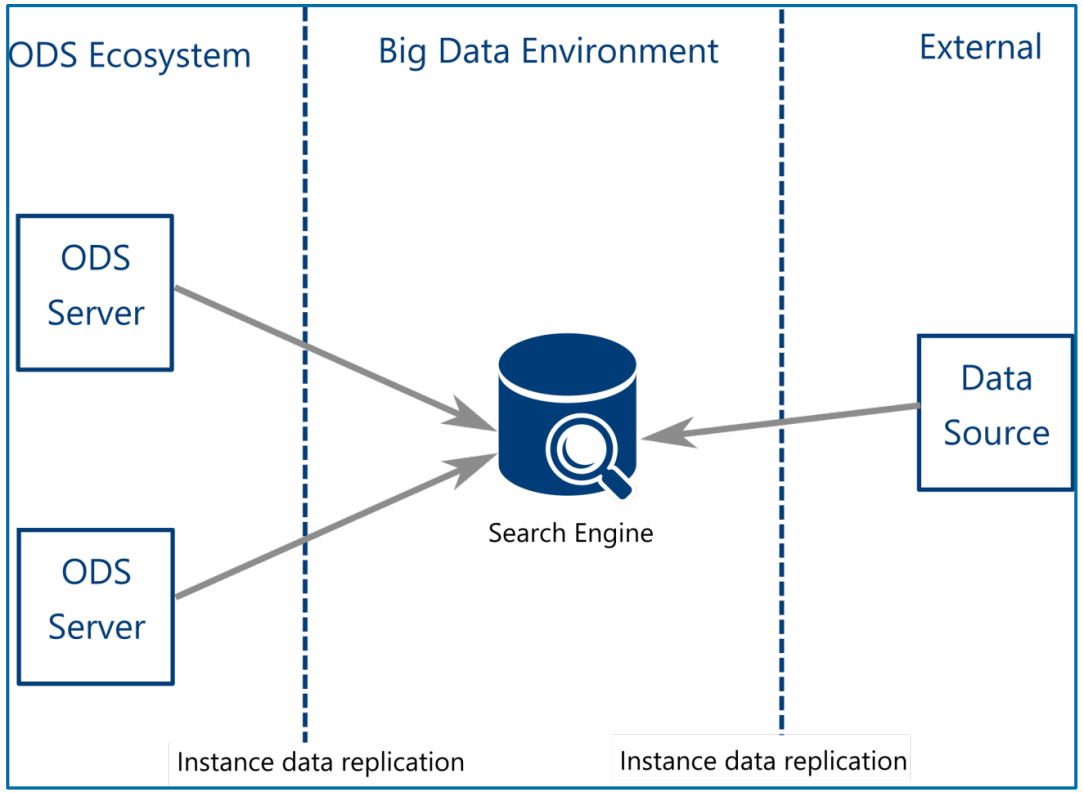


ASAM

Summary & Enabled Use-Cases

ASAM ODS 6.1





Support for most “big data” use-cases is provided.

ASAM ODS 6.1 (“One Working Group”) – next steps

Cross-Test (2019)

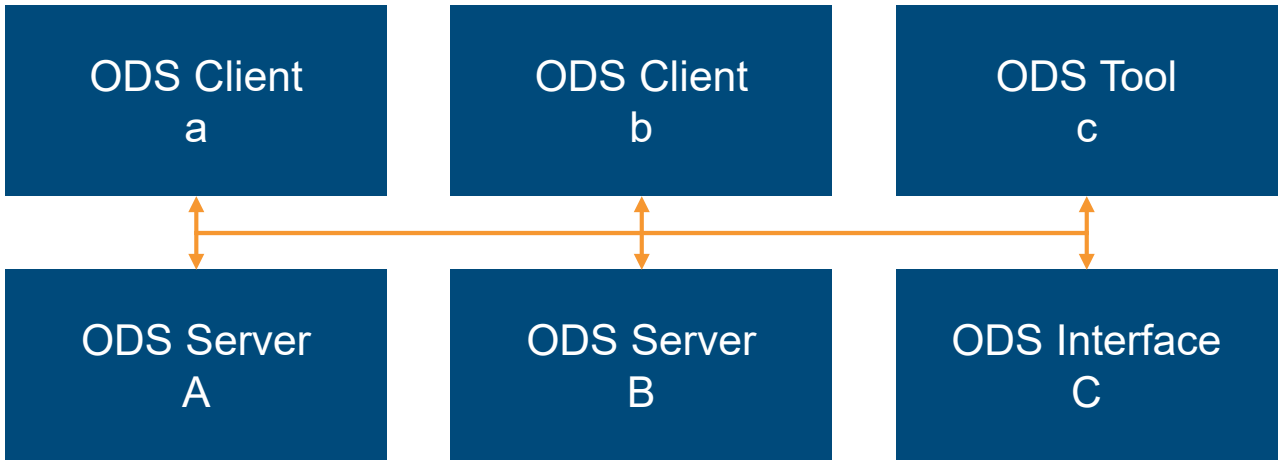
ASAM

ASAM ODS Cross-Test 2019

Overview

At a cross test, the participating companies cooperate to verify and improve their software in real life scenarios. Participants help each other to understand the tiny differences in the formats of data of different programs and to solve the difficulties they might encounter in interpreting these data.

The cross test will offer suppliers of ASAM ODS products the opportunity to test their client applications and servers against a set of real-life example data, provided by OEMs. The aim will be to test exchangeability of ODS data between software products and replaceability of the different existing ODS servers and clients.



https://www.asam.net/conferences-events/detail/asam-ods-cross-test-2019/

ASAM ODS 6.2 (open topics)

An Outlook / Ideation

ASAM

ASAM ODS 6.2 (open topics)

Overview

“Approach A”: The second step of BODS Development (Step 2/2)

- Data Access Interface to mass data
- Data Analysis Interface

Data types and storage (MDF4)

- Consolidation with MDF4 has been executed (MDF4.2)
- MDF4 is used as single data container >> ODS as server infrastructure

openX integration / compatibility

- OpenX data compatibility
  - Geometry data (location and movement)
  - Road Conditions data
  - Scenario data described with ODS (OpenScenario)
  - Labeling of objects (e.g. vehicle) within mass data

} Ideation in process

Thank you!

Dr. Ralf Nörenberg
Board Member, ASAM e.V.
CEO, HighQSoft GmbH

Phone: +49 176 10474402
Email: ralf.noerenberg@asam.net
ralf.noerenberg@highqsoft.de

For more information
on ASAM visit

www.asam.net

ASAM

Enabled Use-Cases

Overview