SAP and Hortonworks Reference Architecture

June 2014
A Modern Data Architecture With SAP
Big Data Reference Architecture

**SOURCES**
- Sensor
- Geo
- Logs
- Text
- Structured
- Weather
- Social
- Other

**Data Ingestion & Provisioning**

**Low Cost Scale-Out Storage**

**HADOOP DATA LAKE**

**SAP HANA**
- Stream Processing
- Accelerated Analytics
- Application Services

**APPLICATIONS**
- Analytic
- Predictive
- Web
- Mobile

**Deep Analytics & Exploration**

**Data Processing Transformation, Rationalization**

Sensor
Logs
Structured
Weather
Social
Other
SAP/Hortonworks Real-Time Big Data Architecture

Transactional, Analytical, Online Applications

Customer Mobile Applications
SAP Mobile Platform

SAP Business Objects BI Suite
Exploration, Reporting, Dashboarding, Predictive, Mobile

SAP HANA
Real-Time Analytics, Interactive Data Exploration & Application Platform

OLAP Engine
Predictive Engine
Spatial Engine
Application Logic & Rendering(XS)

Federated Smart Data Access
Transfer Datasets

Hortonworks Data Platform
Data Reservoir
Large Scale Data Capture, Generate Analytical Datasets, Train and Validate Predictive Models

Real-Time Data Ingestion
Real-Time Recommendation Applications

Real-Time Response
Inline Predictive Analytics for Transactional Applications

Close-Looped Analytics
Smart Mobile Applications
SAP/Hortonworks ETL Rationalization (loading data faster)

- **SAP HANA**
  - Real-Time Analytics, Interactive Data Exploration & Application Platform
  - OLAP Engine
  - Predictive Engine
  - Spatial Engine
  - Application Logic & Rendering(XS)

- **Federated Smart Data Access**

- **Parallel load of valuable data**

- **SAP HANA**
  - Load, then transform at scale: MapReduce, Pig, Java

- **Hortonworks Data Platform**
  - Data Reservoir
  - Large Scale Data Capture, Generate Analytical Datasets, Train and Validate Predictive Models

- **Data orchestration Services**

- **Batch Data Feeds**

- **Transactional Systems, Databases, Flat Files, Batch Data Feeds**

- **Falcon**

- **No Latency ingestion of data from operational systems**
- **Tiered Storage model offers partitioning into Hot-Warm-Cold data during ingestion.**
- **On-the-fly transformation for Hot Data can be performed in memory using HANA**
- **Off-load pre-processing of data to the Hadoop Platform**
Big Data Interactive Data Exploration

- Interactive high performance Analytics and Visualization
- Agile modeling and shorter turn-around on reports & dashboards
- Exploration of Data in –memory and interactively with Hadoop.
- Uniform Data Science Experience on in-memory and multi-terabyte data sets
SAP/Hortonworks Real-Time Stream Processing

Real-time ingestion from operational systems, sensors and smart devices
- Pattern detection, anomaly detection and streaming analytics on data in flight.
- Scalable storage for offline model tuning and data science.
- Instant visibility across operations and corporate functions
SAP/Hortonworks Real-Time Insights and Models

SAP HANA
Real-Time Analytics, Interactive Data Exploration & Application Platform

- OLAP Engine
- Predictive Engine
- Spatial Engine
- Application Logic & Rendering(XS)

Federated Smart Data Access

Online apps
Mobile apps
Visualization and Reporting

Hortonworks Data Platform Data Reservoir
Large Scale Data Capture, Generate Analytical Datasets, Train and Validate Predictive Models

Real-Time Data Ingestion
- Streaming Data Events, Replicate Data Tables from Transactional Applications
- Real-Time Data Acquisition
- SAP SLT
- Sybase Event Stream Processor
- Sybase Replication Server

SAP/Hortonworks Real-Time Insights and Models
- Real-Time Data Acquisition
- Real-Time Data Ingestion
- Real-Time Response
- Close-Looped Analytics
- Real-Time Recommendation Applications
- Inline Predictive Analytics for Transactional Applications
- Smart Mobile Applications
SAP/Hortonworks Real-Time Big Data Architecture

- **Online apps**
- **Mobile apps**
- **Visualization and Reporting**

**SAP HANA**
- Real-Time Analytics, Interactive Data Exploration & Application Platform
  - OLAP Engine
  - Predictive Engine
  - Spatial Engine
  - Application Logic & Rendering(XS)
  - Federated Smart Data Access

**Hortonworks Data Platform**
- Data Reservoir
  - Large Scale Data Capture, Generate Analytical Datasets, Train and Validate Predictive Models

**Real-Time Data Ingestion**
- Real-Time Data Acquisition
- Streaming Data Events, Replicate Data Tables from Transactional Applications

**Real-Time Recommendation Applications**
- Real-Time Recommendation Applications
- Inline Predictive Analytics for Transactional Applications
- Smart Mobile Applications

**Data orchestration Services**
- Batch Data Feeds

**Transactional Systems, Databases, Flat Files, Batch Data Feeds**