



Intelligent Data Warehousing

BW/4HANA + Data Hub

Thomas Zurek, SAP – @TFXZ
March 2019

PUBLIC

Disclaimer

The information in this presentation is confidential and proprietary to SAP and may not be disclosed without the permission of SAP. Except for your obligation to protect confidential information, this presentation is not subject to your license agreement or any other service or subscription agreement with SAP. SAP has no obligation to pursue any course of business outlined in this presentation or any related document, or to develop or release any functionality mentioned therein.

This presentation, or any related document and SAP's strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information in this presentation is not a commitment, promise or legal obligation to deliver any material, code or functionality. This presentation is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This presentation is for informational purposes and may not be incorporated into a contract. SAP assumes no responsibility for errors or omissions in this presentation, except if such damages were caused by SAP's intentional or gross negligence.

All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.

Agenda

Modern Data Warehouses

BW/4HANA + Data Hub

Conclusions

Key Trends: New Categories of Data

- process mediated data
 - OLTP systems
 - LOB applications
 - prescriptive

traditional

- machine generated data
 - IoT
 - logs
 - descriptive
- human sourced information
 - text, tweets
 - photos, videos, sound
 - requires interpretation

new

➔ **big data: volume, types, origins, real-time**
(aka: **volume, variety, veracity, velocity**)

Data Variety

	A	B	C	D	E	F
1	Year	Month	Month2	Converter	Owner	Yield [kWh]
2	2019	201901	01	SB 1200	SMITH	25,534
3	2019	201901	01	SB 1200	MILLER	25,270
4	2019	201901	01	SB 1200	SMITH	103,551
5	2019	201901	01	SB 1200	MILLER	107,579
6	2019	201901	01	SB 1200	SMITH	89,807
7	2019	201902	01	SB 1200	MILLER	90,789
8	2019	201902	01	SB 1200	SMITH	351,316
9	2019	201902	01	SB 1200	MILLER	360,128
10	2018	201801	01	SB 1200	SMITH	21,700
11	2018	201801	01	SB 1200	MILLER	21,654
12	2018	201801	01	SB 3800	SMITH	85,435
13	2018	201801	01	SB 3800	MILLER	87,731

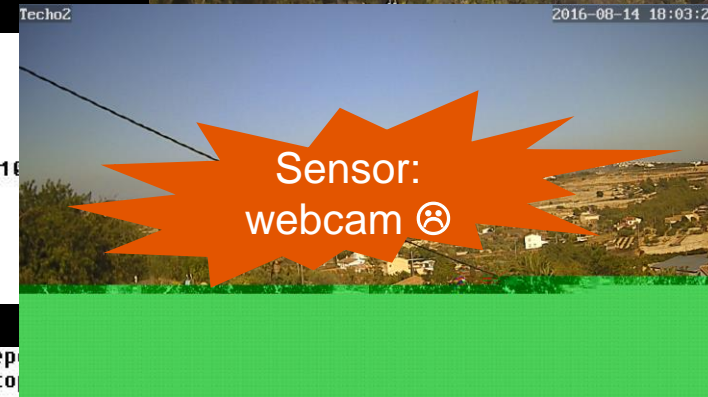
tabular /
traditional



Sensor:
webcam

```
10.7.83.91 - - [06/Jul/2011:22:26:51 +0200] "GET /security/Analysis.htm HTTP/1.1" 200 8326
10.7.83.91 - - [06/Jul/2011:22:26:52 +0200] "GET /favicon.ico HTTP/1.1" 404 280
10.7.83.91 - - [06/Jul/2011:22:27:29 +0200] "GET /security/70/RKT4GSC_WDF.pr HTTP/1.1" 200 686592
10.17.75.27 - - [07/Jul/2011:03:35:11 +0200] "GET /logs/acc1.log HTTP/1.1" 200 10
10.59.145.214 - - [07/Jul/2011:04:51:26 +0200] "GET /BW_Architecture/ HTTP/1.1" 200 10
10.59.145.214 - - [07/Jul/2011:04:51:27 +0200] "GET /BW_Architec HTTP/1.1" 200 10
10.59.145.214 - - [07/Jul/2011:04:51:27 +0200] "GET /favicon.ico HTTP/1.1" 404 280
10.59.145.214 - - [07/Jul/2011:04:51:29 +0200] "GET /favicon.ico HTTP/1.1" 404 280
10.17.75.27 - - [07/Jul/2011:05:00:28 +0200] "GET /misc/access.log HTTP/1.1" 200 5341
10.26.71.232 - - [07/Jul/2011:08:26:54 +0200] "GET /BW_Architecture/ HTTP/1.1" 200 1209
10.26.71.232 - - [07/Jul/2011:08:26:55 +0200] "GET /favicon.ico HTTP/1.1" 404 280
```

Sensor:
weblog

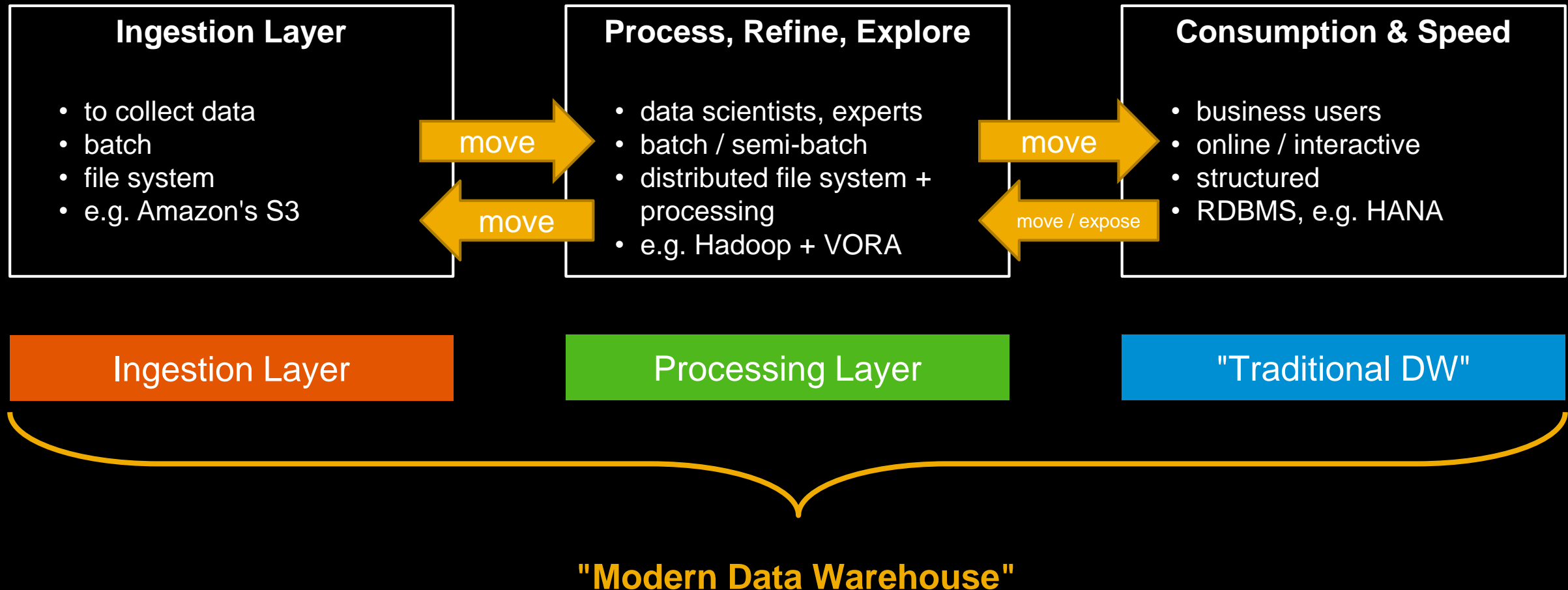


Sensor:
webcam ☹

```
<?xml version="1.0" encoding="UTF-8"?><gpx version="1.1" creator="Runkeep"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="http://www.topografix.com/GPX/1/1"
xsi:schemaLocation="http://www.topografix.com/GPX/1/1 http://www.topografix.com/GPX/1/1/gpx.xsd"
xmlns:gpstpx="http://www.garmin.com/xmlschemas/TrackPointExtension/v1"><trk> <name><![CDATA[Running 2/16/19 3:56 pm]]></name>
<time>2019-02-16T15:56:12Z</time><trkseg><trkpt lat="49.304546000" lon="8.655598000"><ele>101.5</ele><time>2019-02-
16T15:56:12Z</time></trkpt><trkpt lat="49.304523000" lon="8.655726000"><ele>101.4</ele><time>2019-02-16T15:56:16Z</time></trkpt><trkpt
lat="49.304534000" lon="8.655817000"><ele>101.4</ele><time>2019-02-16T15:56:17Z</time></trkpt><trkpt lat=
lon="8.655866000"><ele>101.3</ele><time>2019-02-16T15:56:18Z</time></trkpt><trkpt lat="49.304561000"
lon="8.655889000"><ele>101.3</ele><time>2019-02-16T15:56:19Z</time></trkpt><trkpt lat="49.304561000"
lon="8.655912000"><ele>101.3</ele><time>2019-02-16T15:56:20Z</time></trkpt><trkpt lat="49.304565000"
lon="8.655943000"><ele>101.1</ele><time>2019-02-16T15:56:21Z</time></trkpt><trkpt lat="49.304561000"
lon="8.655979000"><ele>101.0</ele><time>2019-02-16T15:56:22Z</time></trkpt><trkpt lat="49.304561000"
lon="8.656010000"><ele>101.0</ele><time>2019-02-16T15:56:23Z</time></trkpt><trkpt lat="49.304565000"
```

Sensor:
GPX

Storage Tiers in a Modern Data Warehouse



Netflix Example

"Traditonal DW"

Teradata DW

DB

"Hadoop DW"

ETL: Enterprise Scheduler,
Ursula, Aegisthus

Viz
tools

Web
UIs

Hadoop Platform as a Service (Genie)

Job Execution

Resource Configuration &
Management

Traditional Gateways

CLIs

Tools



Hadoop (EMR) Clusters



"Cloud DW"

Cloud Data Warehouse

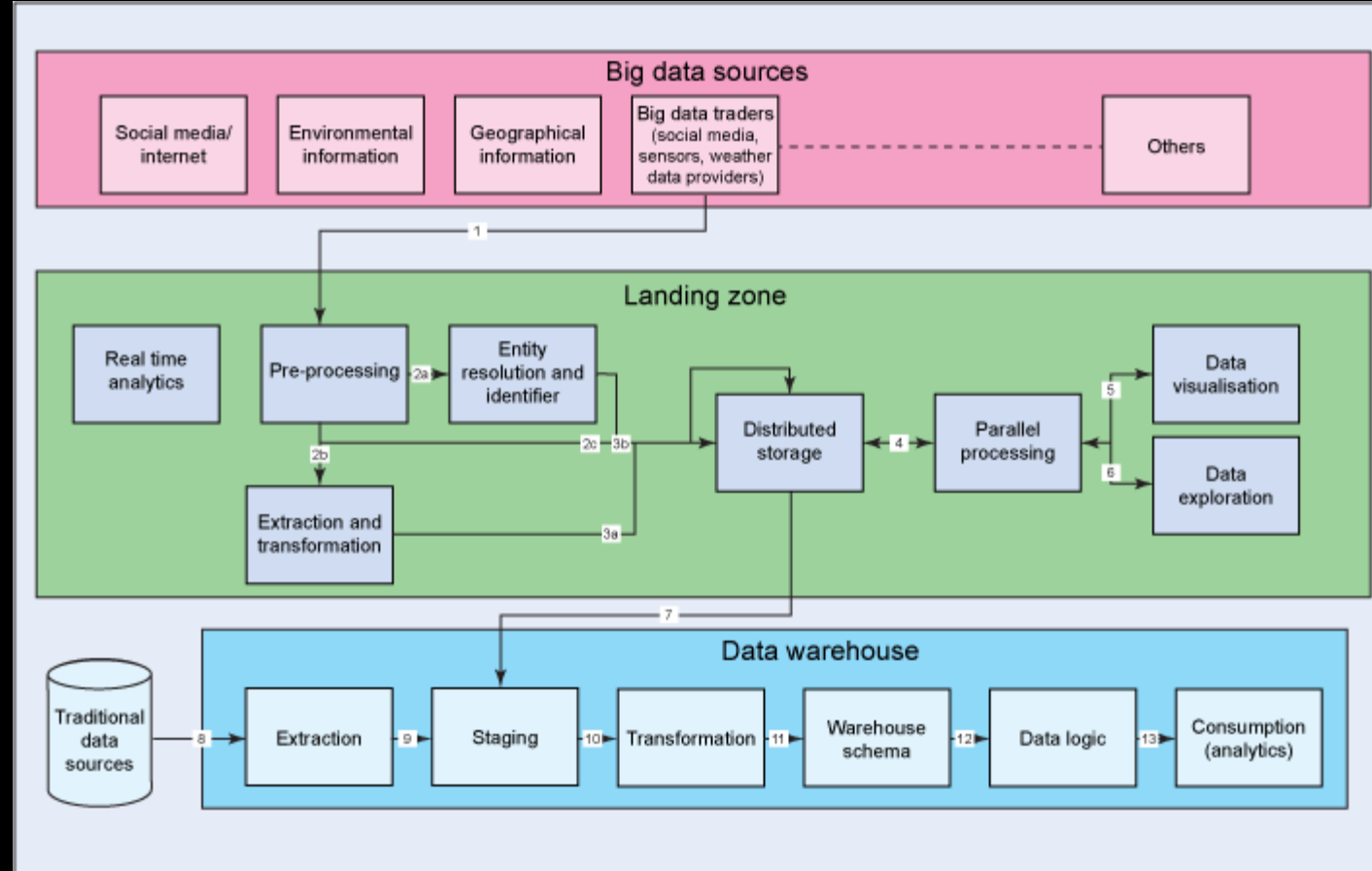


Process

Ingest

<http://techblog.netflix.com/2013/01/hadoop-platform-as-service-in-cloud.html>

IBM Example



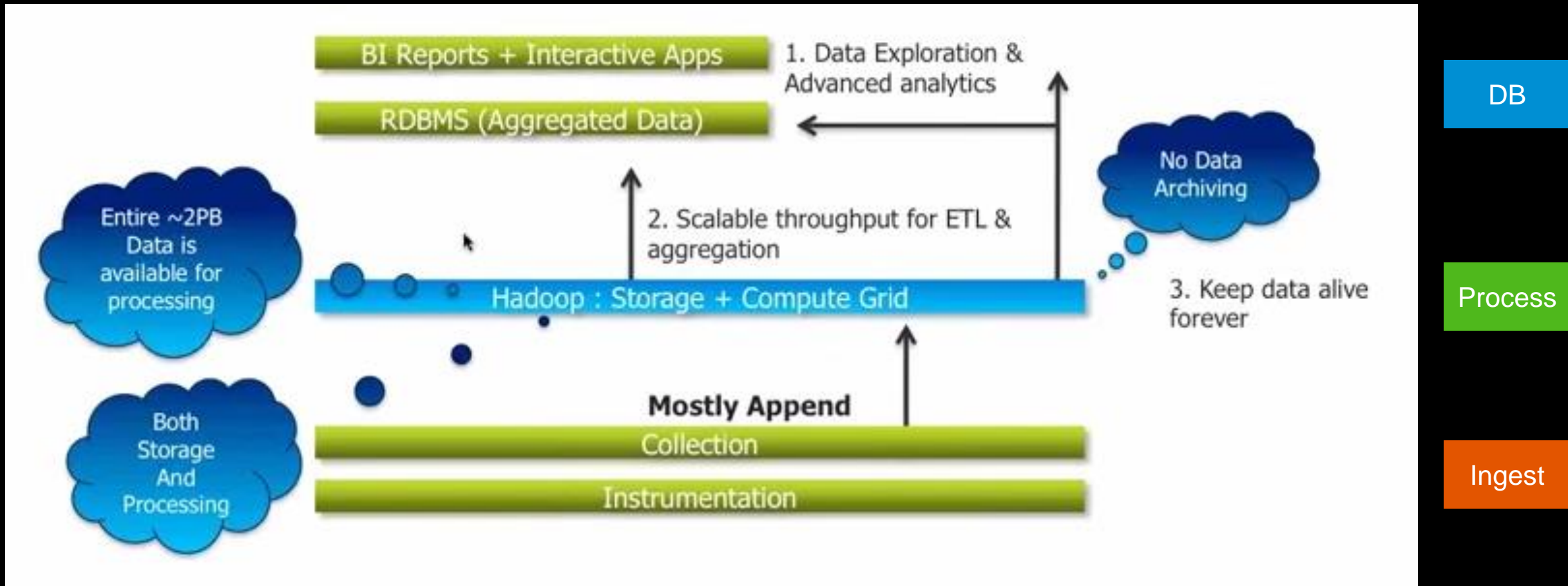
Ingest

Process

DB

<http://www.ibm.com/developerworks/library/ba-augment-data-warehouse2/index.html>

Sears Example



<http://www.edureka.co/blog/big-data-applications-sears-case-study/>

Agenda

Modern Data Warehouses

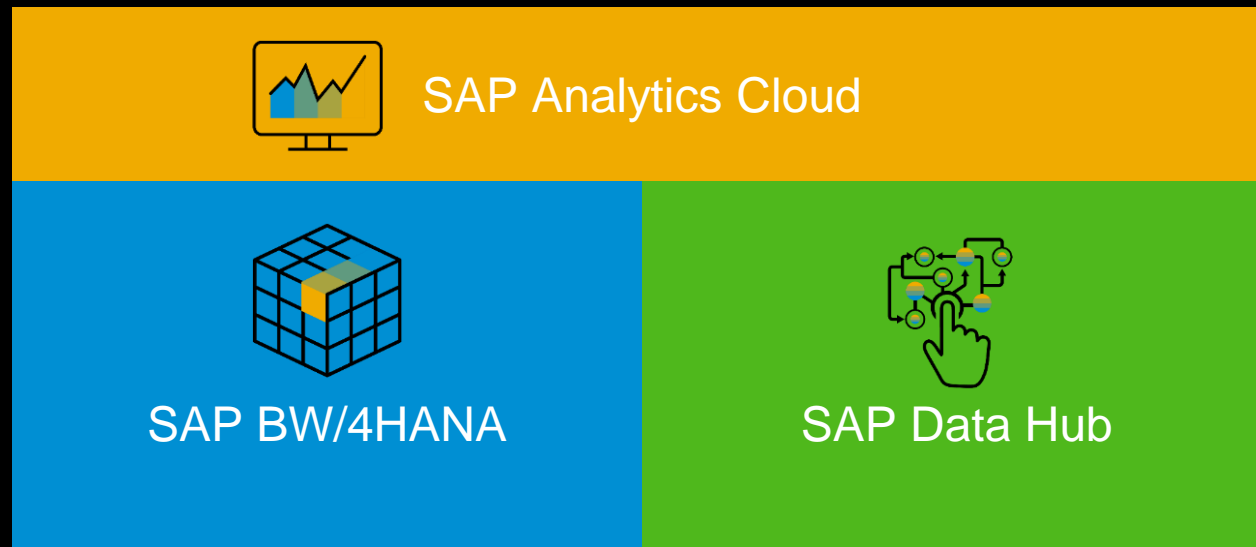
BW/4HANA + Data Hub

Conclusions

Intelligent Data Warehousing

Intelligent data warehousing is not only about **size** and **scalability**, but about **agility** and **integration** capabilities in a distributed, more and more complex customer system landscape of **cloud** and **on-prem** systems, classic **structured** and new **non/semi-structured** data sources.

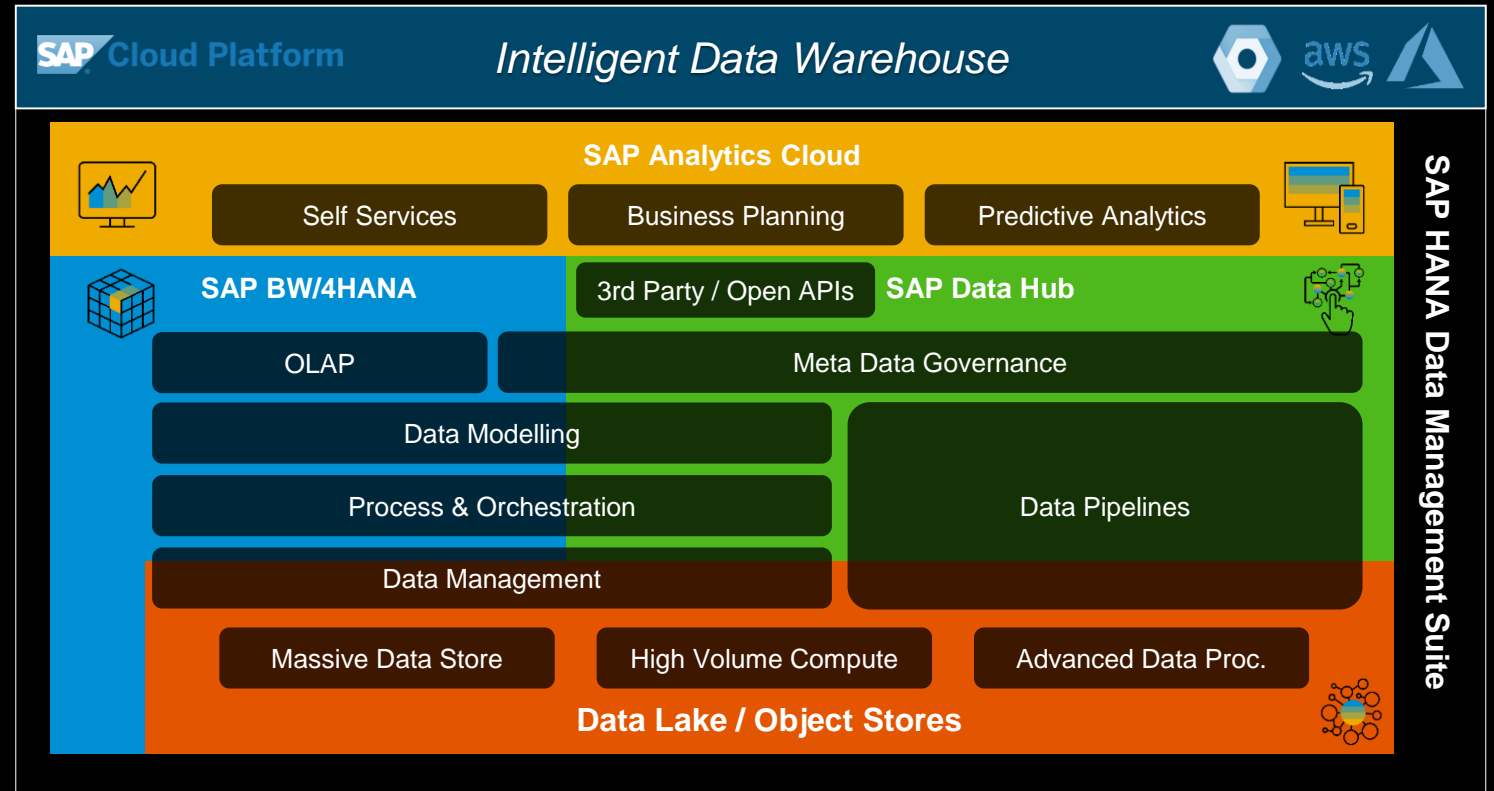
The combination of SAP Data Hub & SAP BW/4HANA is a perfect fit to these challenges!



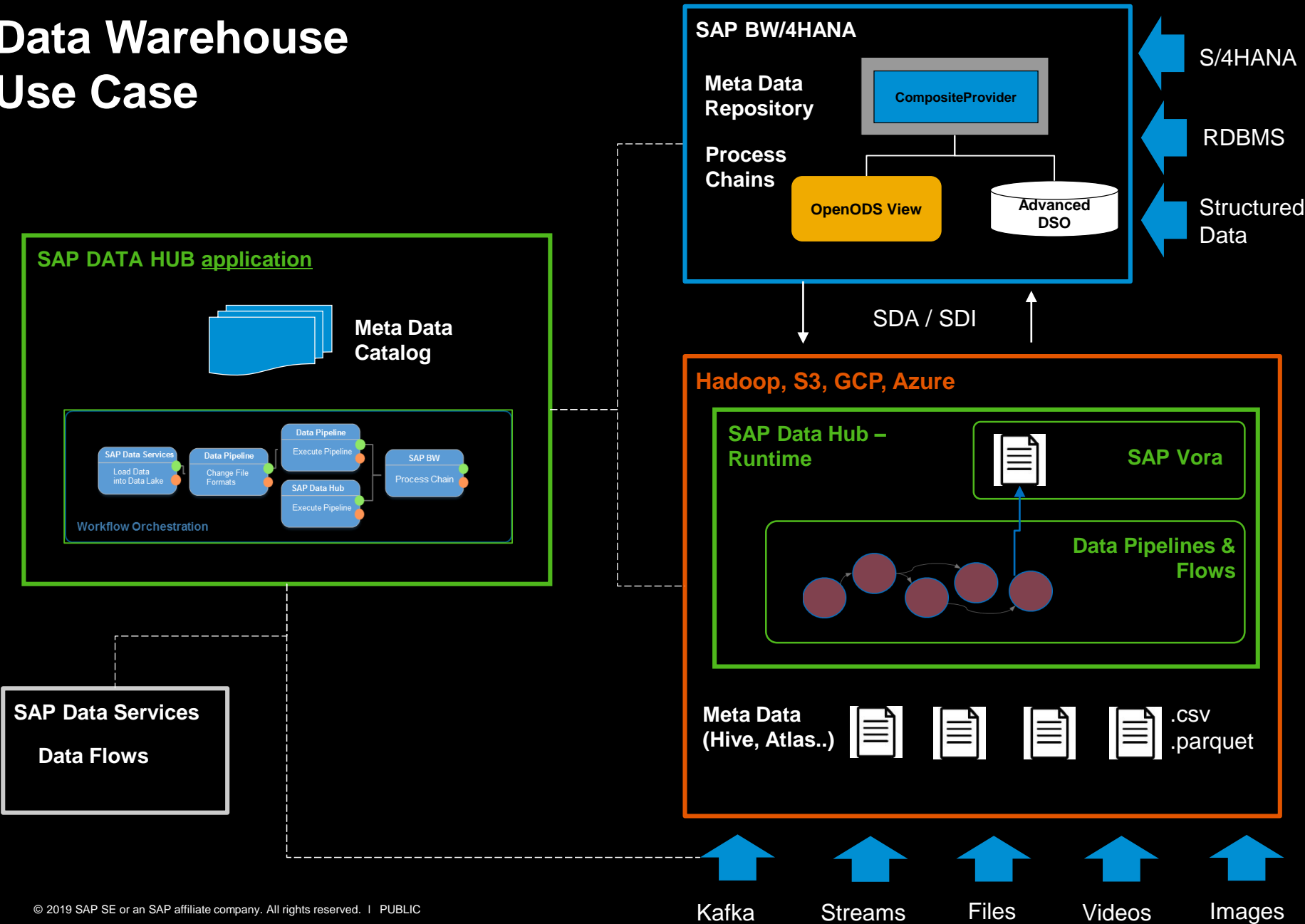
The Intelligent Data Warehouse

THE Data Management platform for the Intelligent Enterprise based on SAP HANA

- Close Integration with S/4HANA & Cloud Applications (SFSF, Ariba,...)
- Schema-less DWH with high automation and minimal Modeling
- Data Lake as primary high volume and computation persistency
- Scalable Storage and Data Processing capabilities in Cloud / On-Premise
- Data Processing beyond OLAP with ML / Predictive Analytics etc.



Data Warehouse Use Case



Analytics Model

Integration

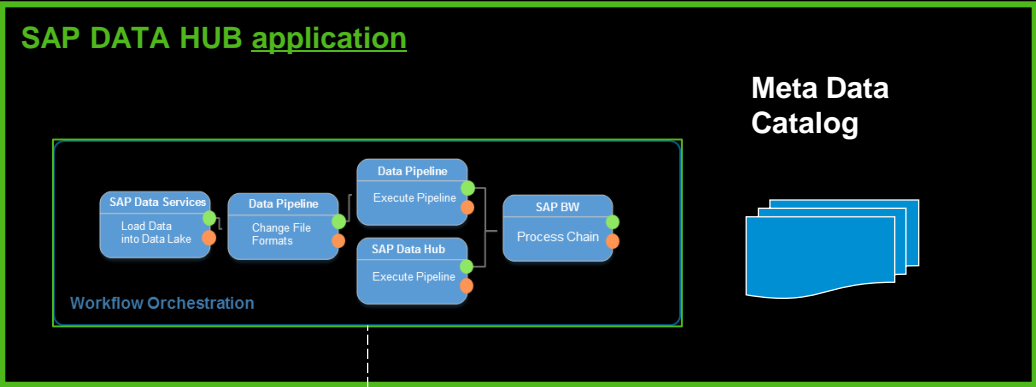
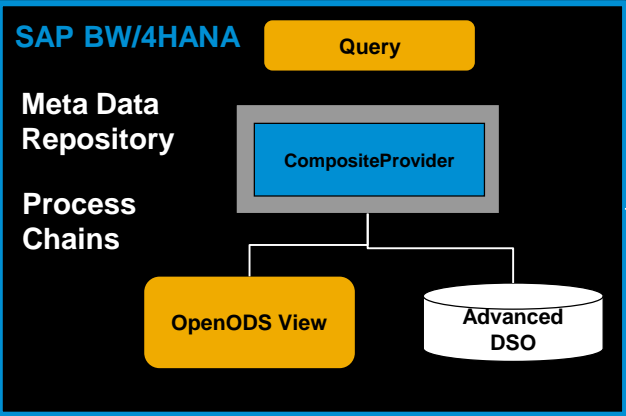
Refined Data

Big Data Processing

Ingestion

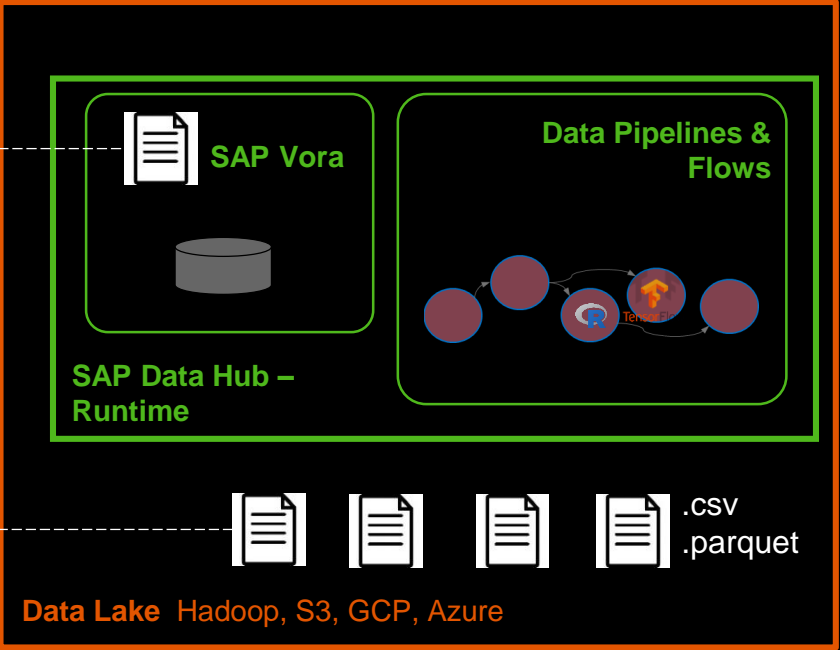
Demo

ML & Data Science Use Case

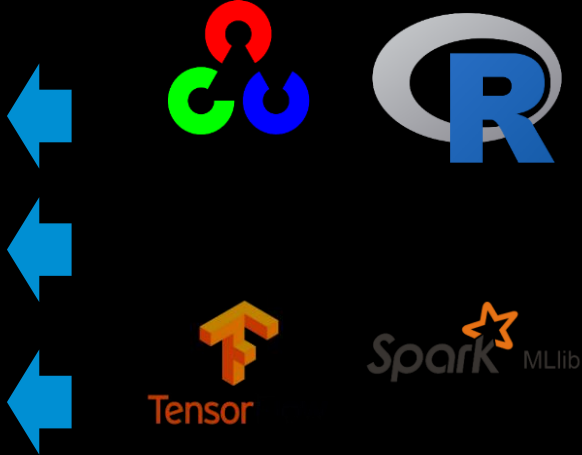


Snapshots

Tiering



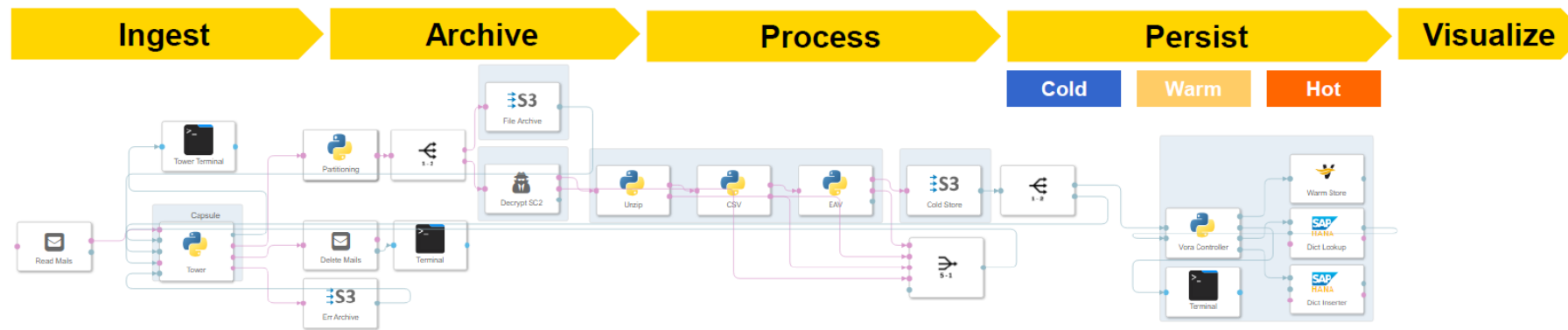
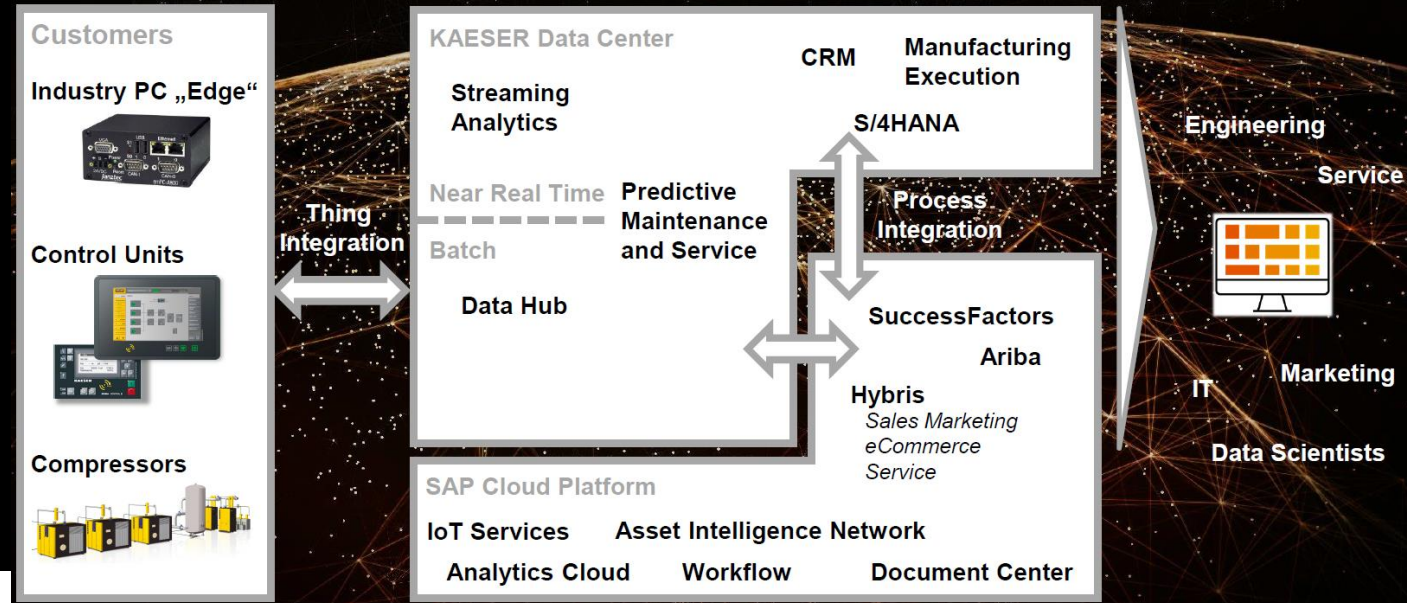
Data Science Community



How to provide SAP Data?
How to productize?



Customer Example: Kaeser (IoT)



- Built an Operator to **retrieve Mails** from Corporate Inbox
- Built an Operator to crawl content from **SAP Document Center / SAP IoT Services File Upload**

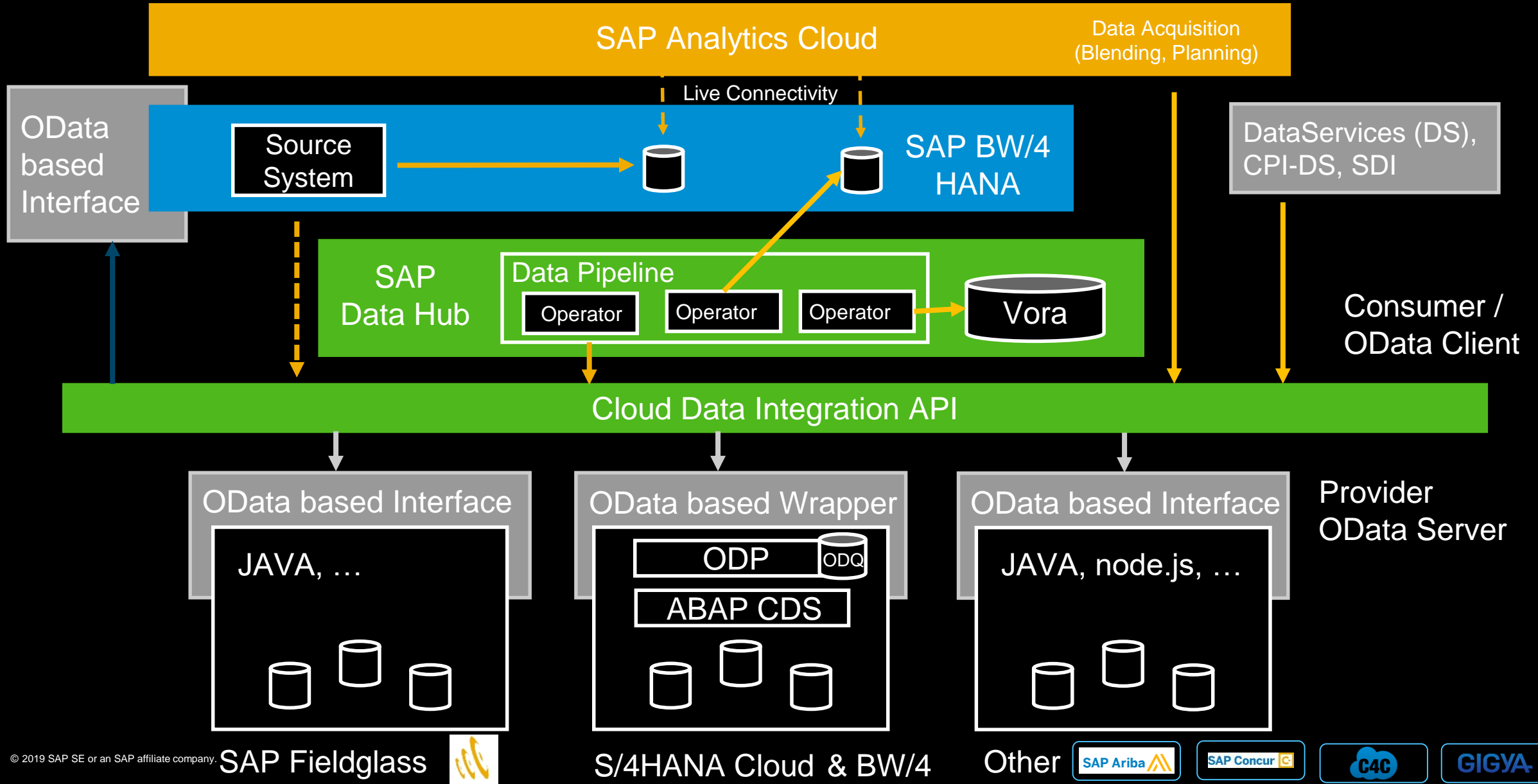
- Store unmodified ingest into S3 / Object Store for **archiving**
- **Error Handling** in receiving & archiving

- Built an Operator to **decrypt and verify** received data
- **Unpack** received data
- **Convert** received binary data into readable format

- Cold – Persist **readable data** in S3 (CSV files)
- Warm – Persist data in **Vora Disc Engine**
- Hot – **Calculation of Aggregates**

- Custom App
- **SAP Analytics Cloud**

Deep Dive “Cloud Data Integration” (CDI)



Agenda

Modern Data Warehouses

BW/4HANA + Data Hub

Conclusions

Key Take-Aways

- An **intelligent data warehouse (IDW)** is an evolution of a traditional DW.
- **BW/4HANA** + **Data Hub** is SAP's highly integrated toolset to build an IDW.
- This is live.

Thank you.

Contact information:

Thomas Zurek

Vice President BW/4HANA

@TFXZ