



SAP DATA HUB ON RED HAT

Sherry Yu
Principal SAP Solution Architect
Red Hat, Inc.

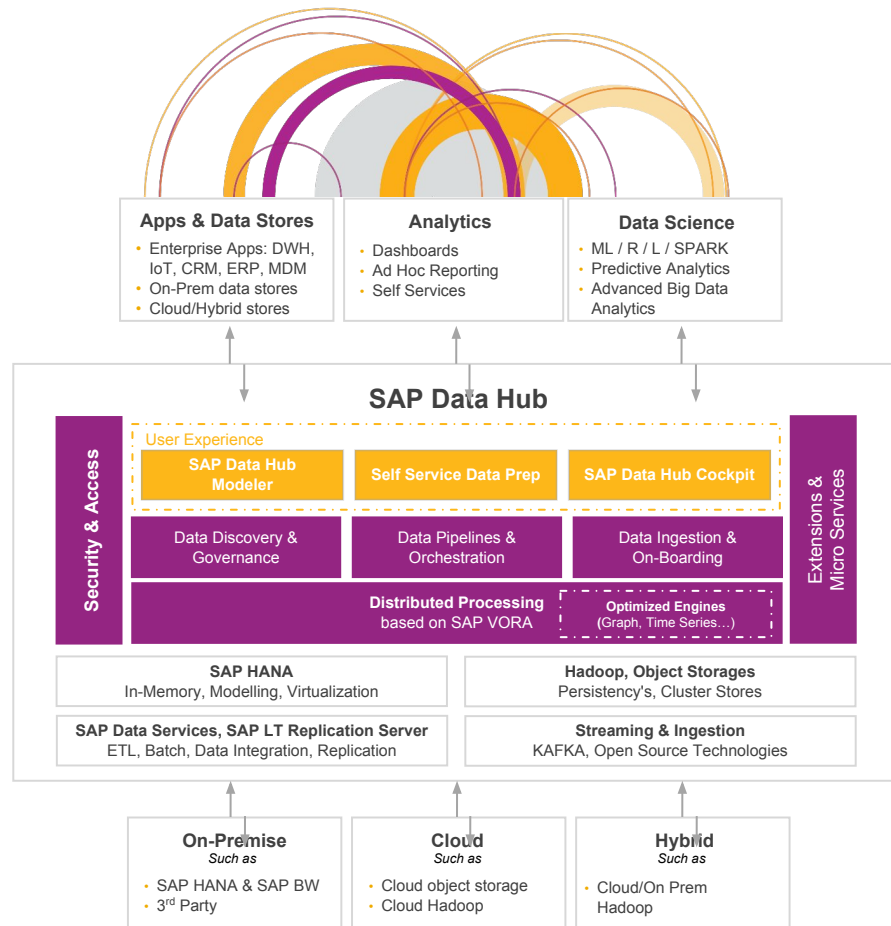
AGENDA

- SAP Data Hub Overview
- SAP Vora Overview
- Red Hat Support and Certifications for Data Hub
- Use Cases
- Roadmap
- Why Red Hat
- Resources

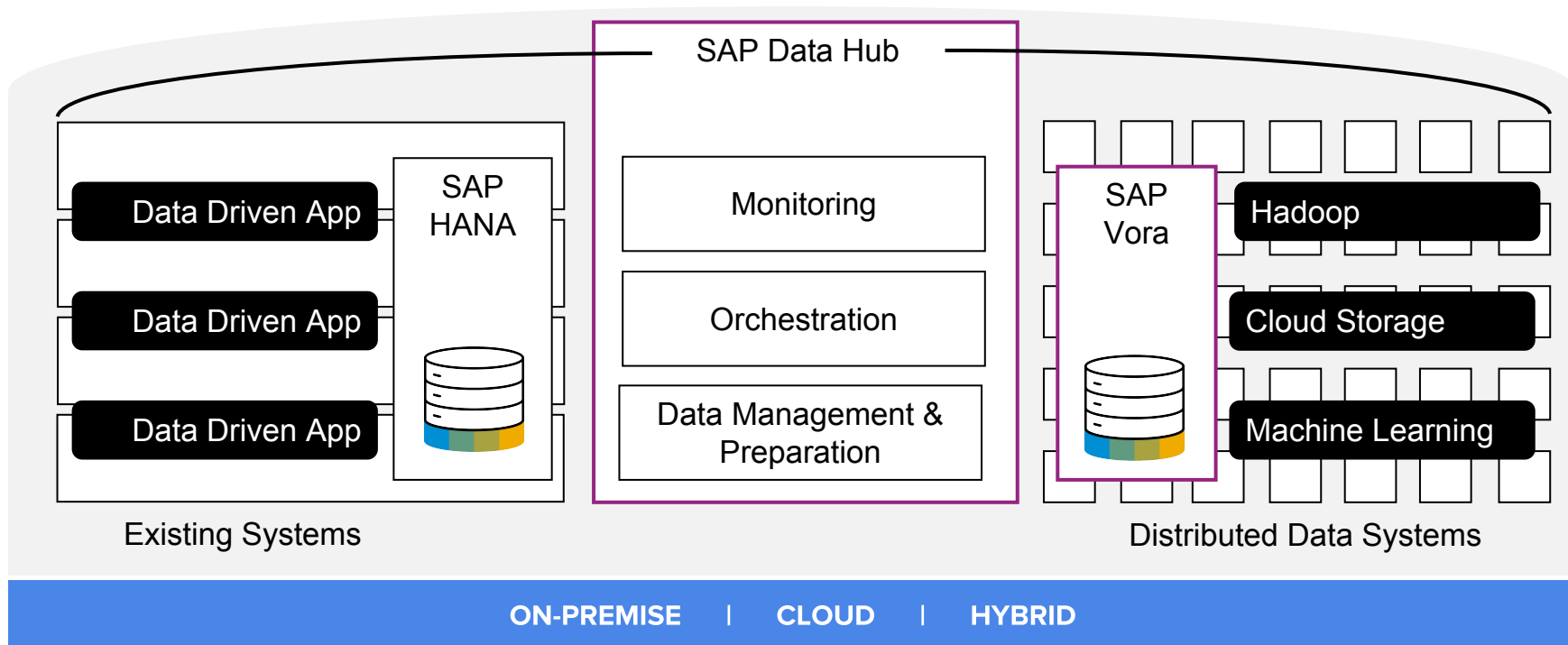
SAP Data Hub Overview

Define data driven processes across complex enterprise landscapes

- Access on-premises, cloud, or hybrid data sources – SAP or non-SAP (Amazon, Hadoop)
- Leverage robust enterprise integration capabilities
- Connect easily to SAP data management and application solutions as data sources
- Connect to SAP and non-SAP applications and analytic solutions as endpoints

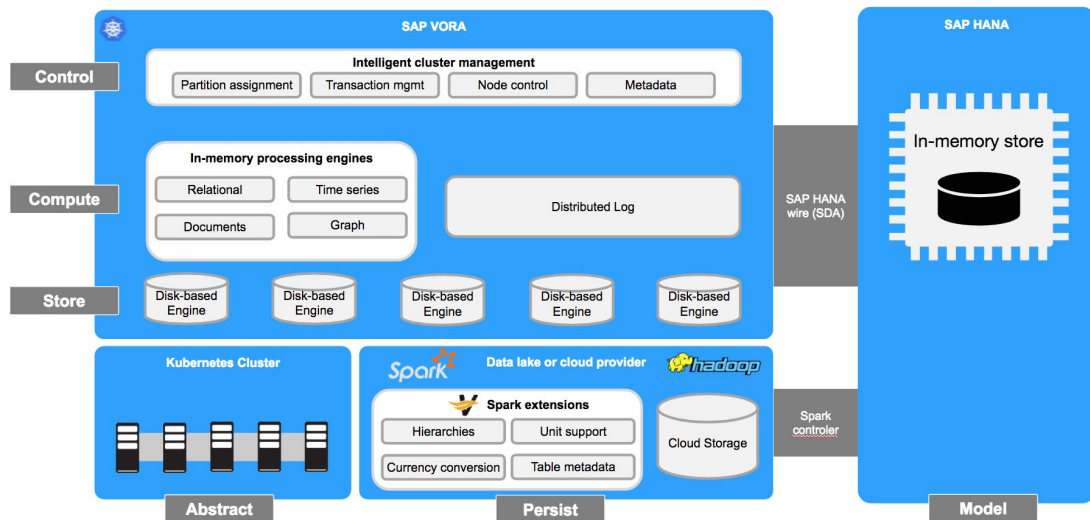


Integrate, Orchestrate and Manage Big Data Infrastructures

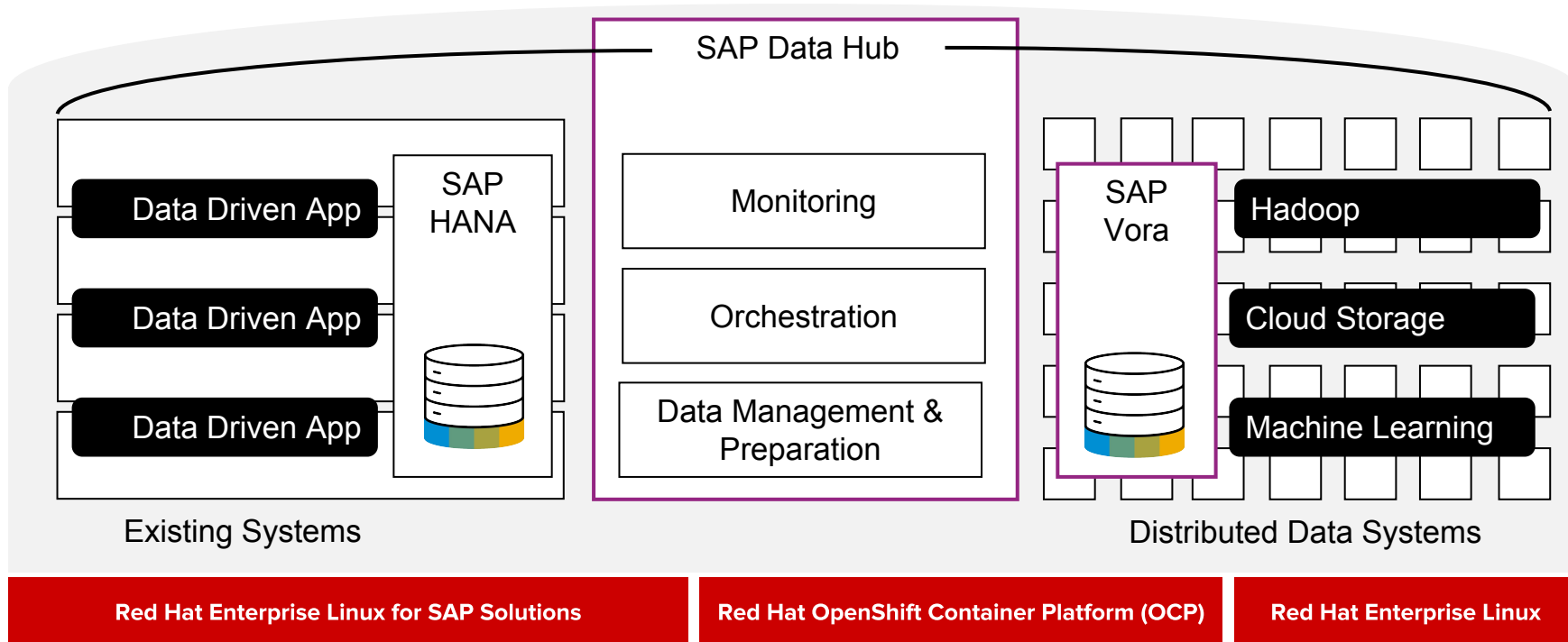


SAP Vora Overview

- In-memory, distributed computing solution for Big Data
- Part of SAP Data Hub
- There is no more public marketing for SAP Vora, but only SAP Data Hub
- SAP Vora 2.x is containerized and developed on Kubernetes
- Red Hat conducts SAP Vora 2.x validation on OpenShift Container Platform (OCP)



Red Hat Support and Certifications



SAP Vora Product Availability Matrix (PAM)

- SAP Vora PAM will only mention supported Kubernetes version
- It's up to Red Hat to validate OCP

Supported Infrastructure, Container, and Hadoop Platforms

Version of SAP Vora	Infrastructure Platform	Container Application Platform for Distributed Runtime ⁽⁶⁾	Required Hadoop Services for Vora Spark Integration
2.1	On premise (x86-64 architecture)	Kubernetes 1.7.x ⁽¹⁶⁾	see table below
2.x ⁽¹⁴⁾	Google Cloud Platform	Google Kubernetes Engine	
2.x ⁽¹⁴⁾	Microsoft Azure	Microsoft Container Service	
2.y ⁽¹³⁾	Amazon Web Services	Amazon Elastic Container Service for Kubernetes	

Required Hadoop Services for Vora Spark Integration

Version of SAP Vora	WebHDFS, YARN, HDFS ⁽²⁾⁽³⁾	Spark ⁽²⁾⁽³⁾	Java ⁽²⁾
2.1	as of Hadoop 2.7.0 or higher	2.1	8

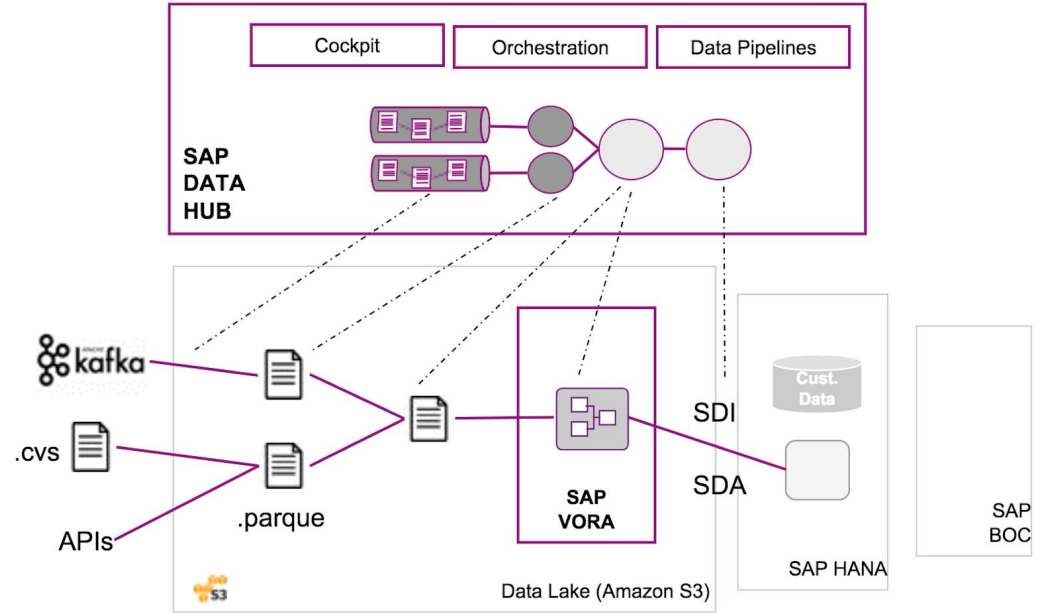
SAP Vora / Red Hat OCP Validation Matrix

- The following matrix has been validated by Red Hat

SAP Vora	Red Hat OCP	Notes
SAP Vora 2.0	OCP 3.6	Installation Guide
SAP Vora 2.1, 2.2	OCP 3.7	Installation Guide

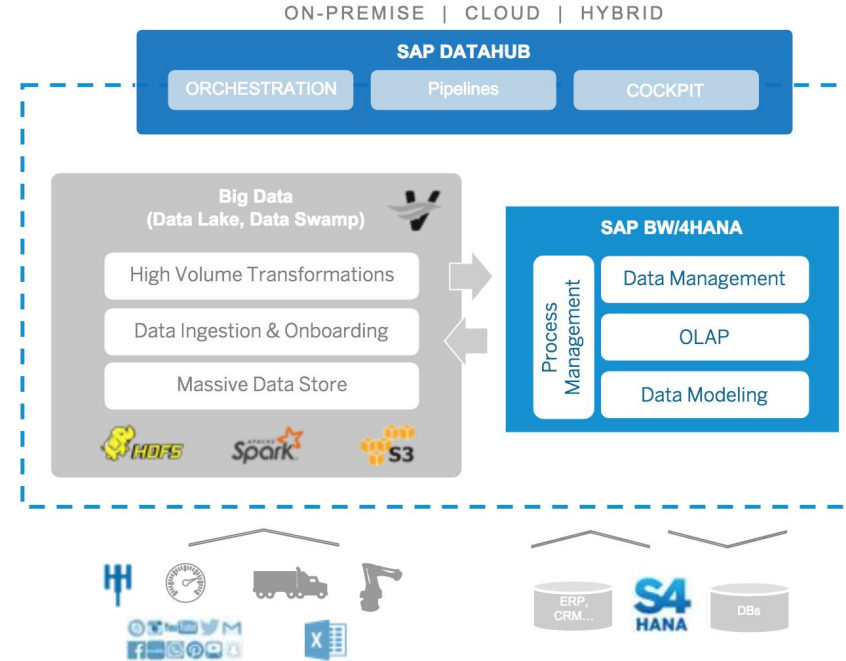
Use Case: Streaming and Processing IoT Data

- How to understand customer behavior and to drive insightful decisions?
- Solution with SAP Data Hub
 - Refine business value from data ingestion to enterprise applications
 - Visual modeling environment
 - Governance and Data Management
 - Orchestration and scheduling to define automated data driven processes



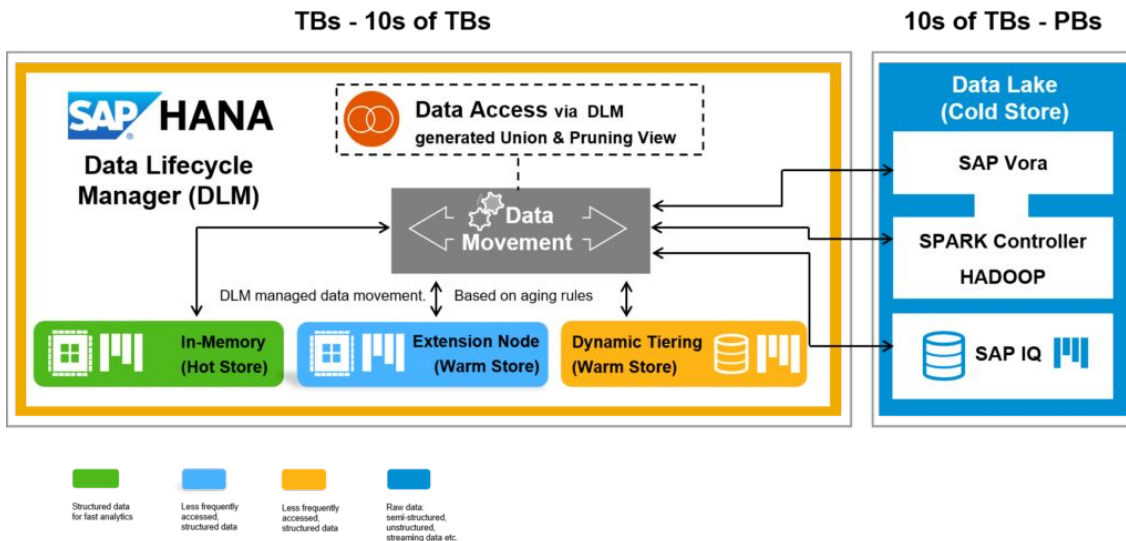
Use Case - Big Data Warehouse

- Build a modern, open, and hybrid DWH offering any data
- BW/4HANA as modern and simplified core data warehouse solution
- Implement and execute high volume transformations on Big Data Clusters Data Lake
- Leverage Big Data landscapes for data onboarding and ingestion
- Data Hub as orchestration and refinery application to address end to end processes



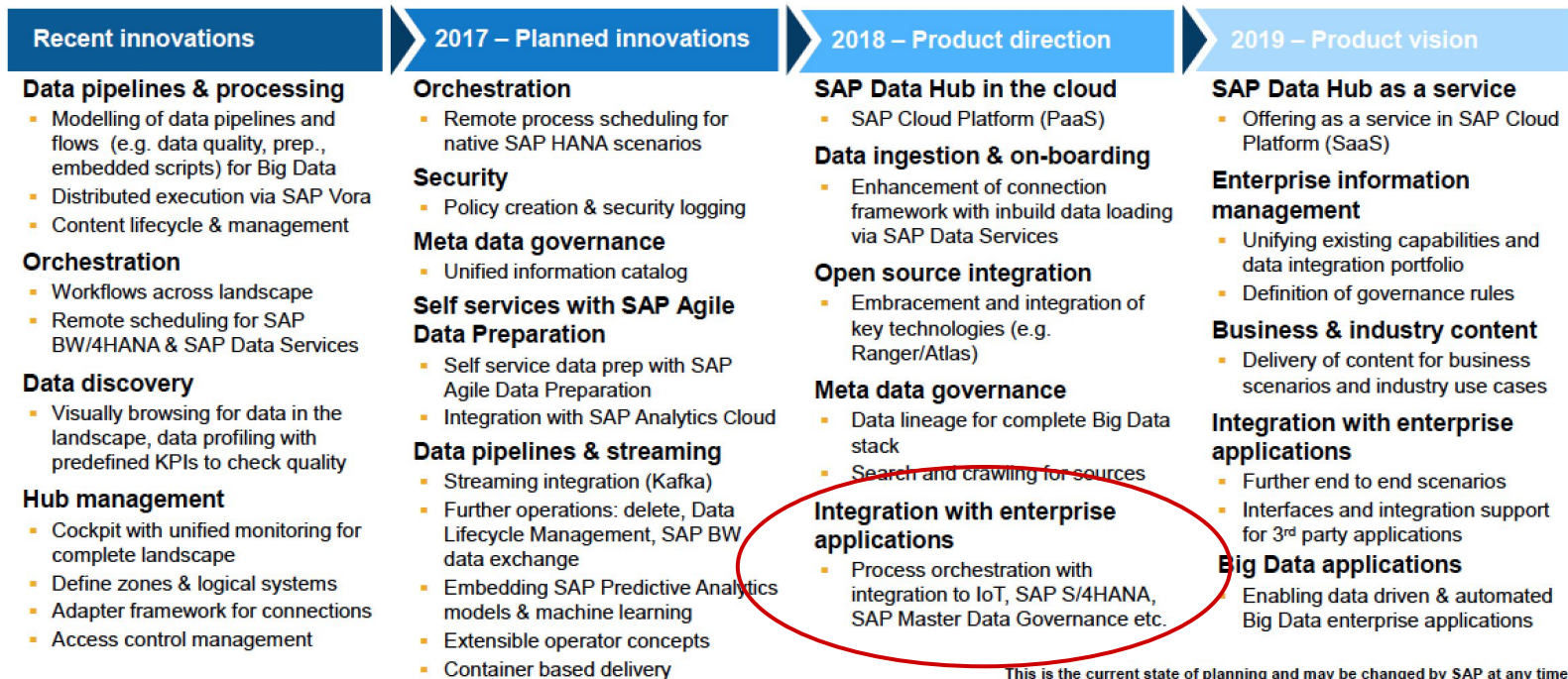
Use Case - Tiered HANA Management

- SAP HANA Dynamic Tiering
 - Utilize lower cost disk storage for historical data within a single SAP HANA database, or even a single multistore table.



SAP Data Hub Roadmap

Product road map overview - key themes and capabilities



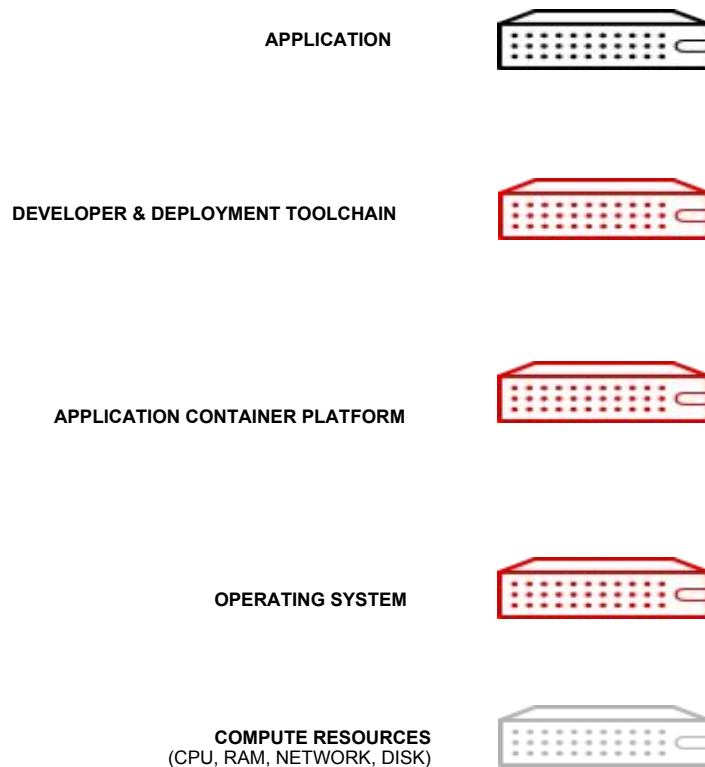
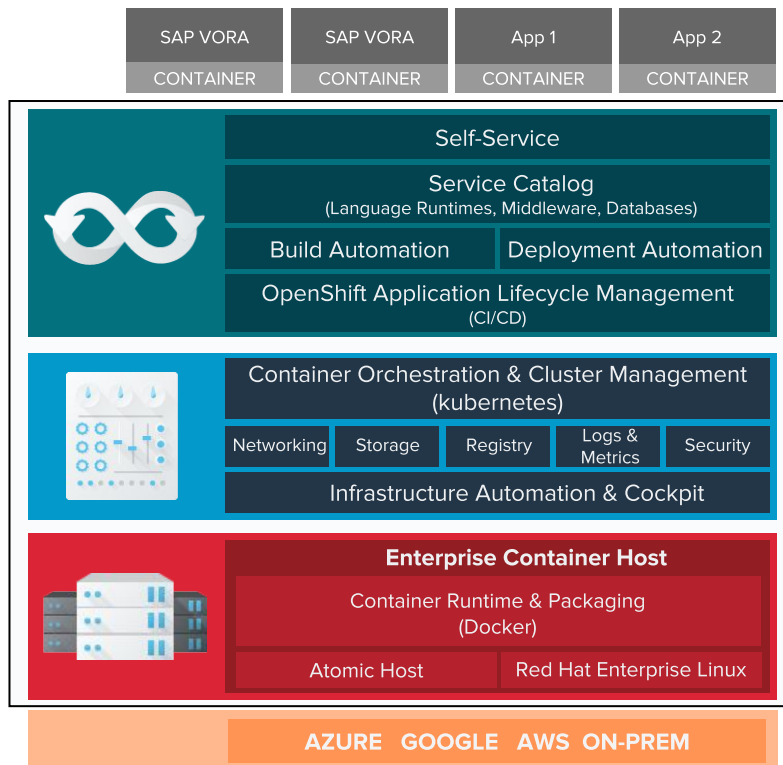
This is the current state of planning and may be changed by SAP at any time.

Why Red Hat for SAP?



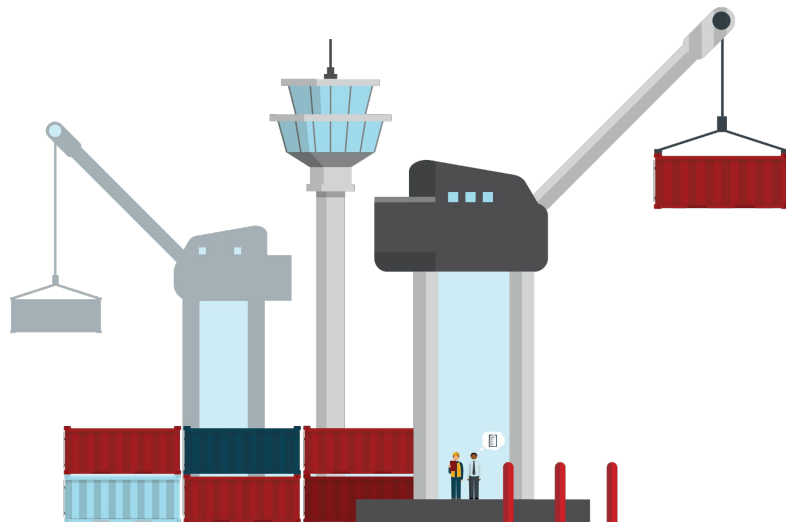
- **Market Leadership**
 - RHEL, 70% commercial market
 - OCP, enterprise grade PaaS
 - And more ... (management tools, automation, etc.)
- **Standardization**
 - As seen on the roadmap, with the integration between SAP S/4HANA (BW/4HANA) and SAP Data Hub, it's more important than ever to choose the right next-generation platform for your digital transformation journey
- **Support**
 - Award winning support organization
 - Integrated support process with SAP
 - SAP TAM

What Else Does A “Containerised” SAP Data Hub System Need To Run?



Build, Deploy, And Manage At Scale

WITH RED HAT OPENSIFT CONTAINER PLATFORM

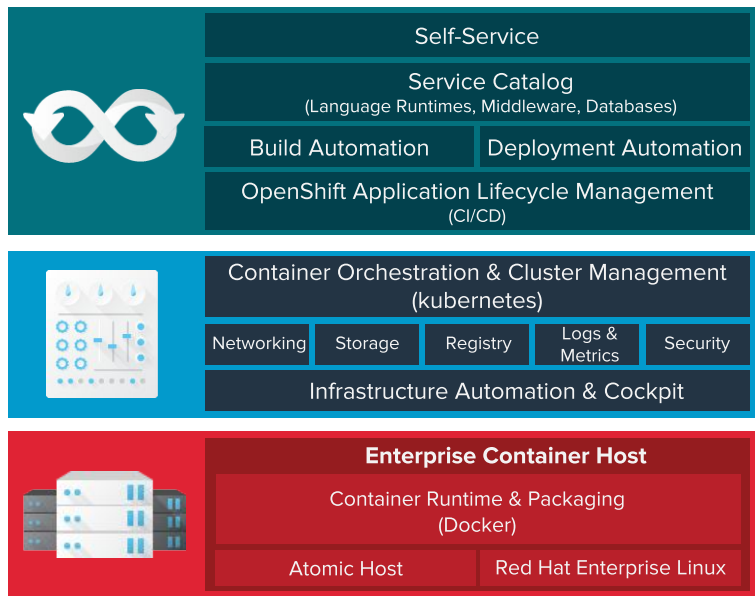


The industry's most secure and comprehensive enterprise-grade container platform based on industry standards, Docker, Kubernetes and Red Hat Enterprise Linux.



“Next Generation Hybrid Cloud Operating System”

Address InfoSec & DevOpsSec



1. Container Host & Multi-tenancy
2. Container Content
3. Container Registries
4. Building Containers
5. Deploying Containers
6. Container Platform
7. Network Isolation
8. Storage
9. API Management
10. Federated Clusters

[Ten Layers of Red Hat Container Security Whitepaper](#)

Resources

- SAP Vora Product Availability Matrix
 - https://support.sap.com/content/dam/launchpad/en_us/pam/pam-essentials/SAP_Vora_PAM.pdf
- SAP Note [2213226](#)
 - Prerequisites for installing SAP Vora: Supported Infrastructure, Container, and Hadoop Platforms
 - <https://launchpad.support.sap.com/#/notes/2213226>
- SAP Vora Technical Documents - Generic
 - https://help.sap.com/viewer/p/SAP_VORA
- SAP Vora Installation Guide on OCP
 - Install SAP Vora 2.0 on OCP 3.6
 - <https://access.redhat.com/articles/3320071>
 - Install SAP Vora 2.1 on OCP 3.7
 - <https://access.redhat.com/articles/3299301>



THANK YOU

Data Hub Components

Cockpit

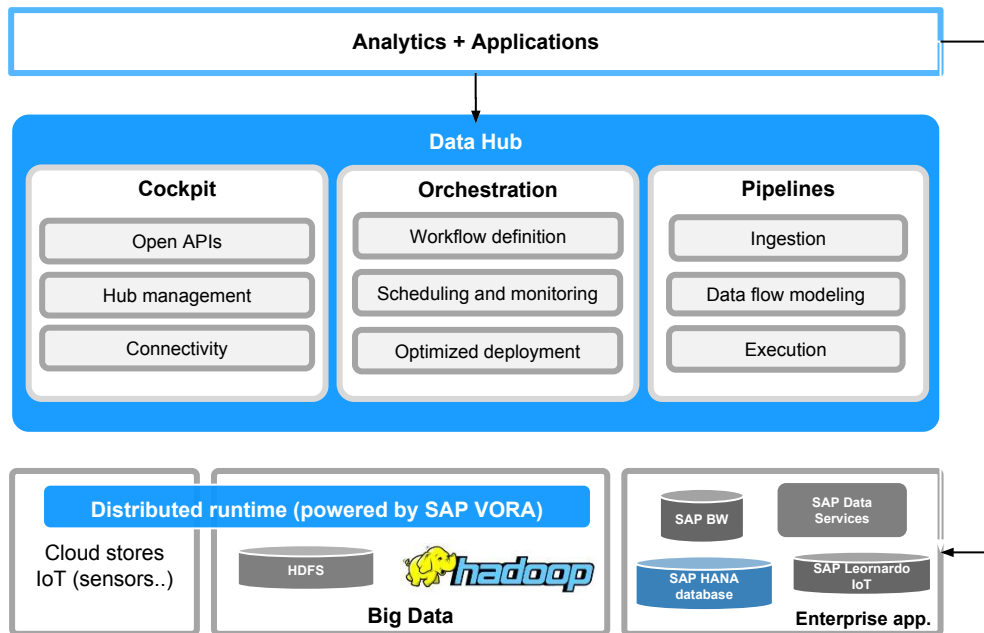
- Hybrid System Management (role-based hub)
- Zone supervision & security
- Data discovery

Orchestration

- Workflow definition
- Scheduling and monitoring
- Pipeline deployment (zone-dependent)
- File-system operations (HDFS, S3, local, ...)

Data Pipelines

- Data ingestion and integration
- Pre- and post-processing
- Complex Algorithms and service endpoints



SAP Data Hub

Architecture View

SAP Data Hub

Application

SAP HANA, XS Advanced Model

Platform Services

UAA Jobs Git ...



Metadata
Catalog



Data Discovery
& Profiling



Scheduling
& Monitoring



Data
Pipelines



Access
Policies



Remote
Orchestration

Distributed Runtime

Kubernetes Cluster

SAP Vora

Containerized

Relational Time-Series
Graph Document
DB Engines

SAP Data Hub Pipelines

Serverless infrastructure

Scripting (JS, Python) Built-in Connectors
Templates Custom Operators
Flow-based applications

Distributed Runtime

Hadoop Cluster

SAP Data Hub Adapter

VORA Spark Extensions

Connected Systems

SAP Integration & Open Connectivity

SAP BW

Process Chains

Data Warehousing Processes

SAP HANA

SDI Flowgraphs

Data Integration into SAP HANA

SAP Data Services

Data Services Job

Heterogeneous Landscapes

3rd party and Open Source

Direct Connectivity

Storage, Messaging, APIs