Hewlett Packard Enterprise | Red Hat

Solution brief

Check if the document is available in the language of your choice.

GOVERN DATA ACROSS ENTERPRISE LANDSCAPES WITH SAP DATA HUB

Red Hat OpenShift Container Platform and Red Hat Virtualization running on HPE Synergy and HPE Nimble Storage

Hewlett Packard Enterprise (HPE) and Red Hat—creating a consistent IT foundation for growth.
Red Hat® OpenShift® Container Platform and Red Hat Virtualization running on HPE Synergy and HPE Nimble Storage provide an SAP® Data Hub solution that delivers end-to-end data governance across complex diverse data landscapes.

This reference architecture provides an SAP Data Hub solution that:

- Delivers end-to-end data governance to deploy, scale, and manage environments.
- Unifies data management, and centralizes governance and pipelining capabilities of complex data landscapes.
- Helps govern and manage metadata assets spread across diverse systems.
- Provides data protection, individual privacy, and security using the SAP Data Hub ecosystem.
- Efficiently lays out an OpenShift configuration using a mix of virtual machines (VMs) and bare-metal hosts.

WHAT IS DATA GOVERNANCE?

Data governance is a system of decision rights and accountabilities for information-related processes. This model describes who can take what actions, when, with what information, under what circumstances, and using what methods.

The goals of implementing a data governance program include enabling better decision making, protecting the needs of data stakeholders, building repeatable processes, and ensuring transparent processes.

WHY SAP DATA HUB?

It is getting harder and costlier for organizations not only to understand the data they have, but also to work across the different systems that need to use it, and then apply end-to-end governance to capture the maximum value.

SAP Data Hub delivers a simpler, more scalable approach to data landscape management. With enterprise-spanning data integration, processing, and governance, SAP Data Hub provides unprecedented insight and access across the complex network of data in the modern enterprise.

SAP DATA HUB FEATURES

- Built to deploy on-premises, on cloud, or in hybrid landscapes
- Create orchestration scenarios and monitor operations and processes end-to-end
- Modern UI for hub management
- Control of an end-to-end data landscape that connects enterprise data with big data
- Duplicate data sets for data scientists to facilitate training machine learning models
- Fine grain, end-to-end security provided by policy management
- Manage access control of all data, enterprise, and big data

99.9999%

Guaranteed availability
HPE Nimble Storage all-flash arrays combine a flash-efficient architecture with HPE InfoSight predictive analytics to achieve fast, reliable access to data.[1]

Red Hat OpenShift Container Platform on HPE Synergy provides an end-to-end fully integrated container solution that, once assembled, can be configured within hours. This eliminates the complexities associated with implementing a container platform across an enterprise data center and provides the automation of hardware and software configuration to quickly provision and deploy a containerized environment at scale.

Red Hat OpenShift Container Platform provides organizations with a reliable platform for deploying and scaling container-based applications. HPE Synergy provides the flexible infrastructure you need to run the container platform to dynamically provision and scale applications.

LEARN MORE AT

hpe.com/info/synergy
hpe.com/us/en/storage/nimble
redhat.com/en

[1] HPE.com, "HPE Get 6-Nines Guarantee," September 2017.

Make the right purchase decision.
Contact our presales specialists.

Chat Email Call

Get updates

Hewlett Packard Enterprise

- Data discovery capabilities including profiling of data natively without leaving the source
- Rich data transformation through SAP Enterprise Information Management tools
- Data pipelining to connect data in different formats and make them accessible to SAP Vora® for analytics
- Easy-to-use advanced data pipeline to move data from various sources, including Amazon S3 to Hadoop HDFS, HDFS to SAP Vora, and SAP Vora to HDFS

SEE NEW OPPORTUNITIES IN YOUR DATA WITH SAP DATA HUB

SAP Data Hub provides unprecedented visibility into and access across the complex network of data in the modern enterprise. By providing a broad, detailed, and easily understood view of the entire data landscape—from sources like Hadoop and Amazon S3 to SAP HANA® and ERP—it helps organizations deeply understand data sources, uses, interconnections, quality, and impacts. This allows enterprises to see new opportunities from data, resolve emerging data issues, and ensure that data is flowing to where it needs to go.

A HIGHLY AVAILABLE AND SECURE SAP DATA HUB SOLUTION

The SAP Data Hub solution focuses on data governance using a Red Hat OpenShift Container Platform deployment on HPE Synergy Composable Infrastructure. It includes key features of the SAP Data Hub application, specifically end-to-end data governance and details on the design and configuration of the environment.

The Red Hat OpenShift Container Platform can be set up to take advantage of HPE Synergy Composable Infrastructure and HPE Nimble Storage, which provides the persistent storage the SAP Data Hub needs for containers and registry, VM storage, and data management.

MAKE BETTER DECISIONS AND MANAGE DATA SETS ACROSS SYSTEMS

One of the key design points for SAP Data Hub is to ensure data reliability, traceability, and compliance in accordance with your business. SAP Data Hub helps you manage your data across different systems by using the Metadata

Explorer, which gathers information about the location, attributes, quality, and sensitivity of data. This allows you to make informed decisions about which data sets to publish and determine who has access to use or view information about the data sets.

SAP Data Hub Metadata Explorer lets you govern and manage metadata assets spread across diverse systems and disparate sources so you can:

- Preview data in the data sets
- Create indexes about the data set contents to aid in searching
- Profile data to view information about the contents of different data sets
- Publish data sets to allow others to view and search the data
- Label the data set with keywords, which also helps in searching
- Prepare the data sets by applying data quality enhancements
- Conduct lineage analysis to learn where the data set is used and how it is transformed
- Create validation rules to ensure that your data passes data quality standards
- Monitor the status of tasks

Effective data governance can be achieved through metadata management, data lineage and impact analysis, and data access and security. Metadata is managed through indexing, publishing, and profiling. Indexing extracts the metadata, so the data can be searched. The extracted information is available in the search and browse pages.

a50001472ENW, April 2020