Whitepaper

Enterprise Linux

SAP® Data Hub on SUSE Container as a Service Platform

Introduction

Enterprise data is exploding, making it both a challenge and an opportunity. Companies are discovering ways to transform their data into services that help to differentiate the business and create new lines of revenue. Unfortunately, managing and fully utilizing the information stored in data silos (e.g., cloud databases, Hadoop clusters, social media feeds) has become incredibly complex due to requirements for security, governance, and specialized training.

SAP Data Hub provides a GUI-based, business-wide view of a broad array of data systems, databases and assets, enabling your analytics and business intelligence teams to manage your entire data landscape through an intuitive "single pane of glass."

In what is likely to become a trend, SAP uses open source to focus on their core competencies, making them leaders in mission-critical enterprise applications. This is the promise of the SAP Intelligent Enterprise. Supporting SAP with innovation, SUSE has been the market leader for SAP applications for over 20 years. SUSE is the trusted and preferred open source platform for SAP customers who want to unlock data intelligence, drive innovation and run with the best.

As an example of this open source mandate, SAP Data Hub is deployed on a Kubernetes-compatible container platform. **SUSE CaaS Platform** enables you to extend your SUSE Enterprise Linux for SAP environment to container-based application delivery.

Target Audience

This paper is directed to professionals involved in both IT Operations and Analytics/Business Intelligence. The recommended framework supplies the requirements to implement SAP's Data Hub, which is certified on SUSE CaaS Platform.

Business Problem

Today's business leaders are under increasing pressure to drive their business with data-driven decisions. This presents a particular challenge for those executives who strive to bring together the right combination of disparate data sources to unlock new value for their business.

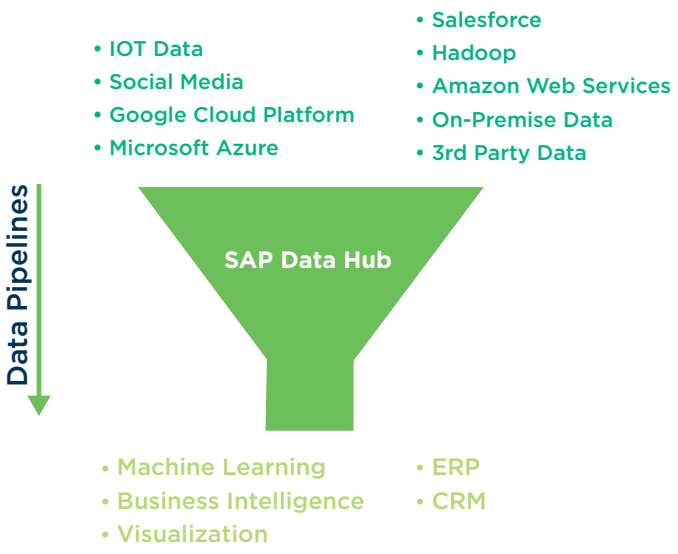


Figure 1. *SAP Data Hub Data Pipeline Funnel*

This difficulty is compounded by the very nature of how data is collected and stored, which results in independent data silos

that have no easy way to make critical associations across them. These data silos might be stored geographically close together or far apart. Some might be built with on-premises resources and some housed in one or more public clouds. Valuable data is often found in structured and unstructured databases, Hadoop data lakes, data warehouses and even in text files. Gaining new insights into potential customers and business opportunities could involve nearly all of the data silos a company has available to it.

What's more, finding a business application to perform the task isn't even the hardest part. Businesses need to ensure that their data-analysis software investment can meet the scale of their current and future application/data landscape, as well as enforce data governance. Equally important is the resiliency and scalability of the underlying infrastructure. Experienced leaders know that enterprise-grade software is a poor investment, unless it is built on enterprise-grade infrastructure.

Business Value

SAP Data Hub is a containerized solution designed to be deployed on enterprise-grade Kubernetes clusters such as SUSE CaaS Platform. For more than 17 years, SAP has developed its software on SUSE Linux Enterprise Server (SLES) and SUSE solutions such as our CaaS Platform. (Note: See SAP Note 2693555 for certified systems.)

SAP Data Hub is built on a next-generation data-aggregation model that does away with the need for expensive data warehouses. Instead, SAP Data Hub allows for data extraction and formatting to be done on the platform where the data resides. This is in contrast to the current practice of using cumbersome, single-use Extract, Load, Transform (ELT) operations that are used to populate data warehouses. SAP Data Hub provides formatted, refined and cleansed data from multiple sources directly to the data consumers.

SAP Data Hub leverages data pipelines, which are built from reusable application components. Data pipelines are computational models that are executed natively on the data source. They define what data should be gathered from which sources and how that data should be formatted at the source. Pipelines also specify the refinements and cleansing that each stream of data should go through to make it compatible with the other data streams in the pipeline. Finally, the data pipelines identify which consumer(s) the collated data should be sent to. Because SAP Data Hub does not need to persist data, it eliminates the need for expensive, scale-limiting data warehouses.

Data pipelines can be created through a graphical user interface to leverage existing data sources such as SAP HANA, SAP Vora, Apache Spark and Apache Hadoop, as well as all major open and closed source OLTP, OLAP and NoSQL databases.

Before implementing a data-analytics solution, consider the specific problem you are working to solve. Below are some use cases for SAP Data Hub that can help you zero in on the type of solution you are pursuing.

Example Use Cases

This section outlines potential use cases for SAP Data Hub built on SUSE Containers as a Service Platform. In general, SAP Data Hub excels in pulling information from multiple types of internal and external data resources to enable insight into very complex analytical problems. The use of machine learning and Big Data analytics platforms (SAP, Hadoop, MapR, Cloudera, etc.) require access to large pools of unstructured data in a highly automated, systematic and secure way.

Fraud Detection

Credit card fraud has become an epidemic, with losses in the billions of dollars. Financial institutions need the ability to create profiles that alert them to probable fraud on large volumes of transactions. The more information they can cross-reference, the more accurate their models will become. SAP Data Hub can pull in transactional data from ERP systems, credit reporting bureaus, email from a Hadoop cluster, social media data and "Dark Web" databases, enabling data scientist to build very precise detection methodologies.

Manufacturing Equipment Maintenance

Global manufacturers rely on the uptime of their equipment to meet product delivery targets. Unscheduled maintenance or equipment failure can result in lost profits, poor quality and unmet commitments. Conversely, over-scheduling maintenance activities also impacts cost and output. Manufacturers were early adopters of Internet of Things (IoT) technology for the real-time monitoring of equipment sensors (temperature, vibration, humidity, motor loading, etc.) to gain a better understanding of the state of their environment. What if predictive models and machine learning (ML) could be used to optimize maintenance scheduling? SAP Data Hub can be used to orchestrate the end-to-end data flow needed to feed an ML platform to predict impending outages and then schedule corrective maintenance before any disruptions occur.

Customer Affinity Recommendations

E-Commerce sites routinely use various data sources to recommend additional purchases or fine-tune searches to more relevant items. Early attempts were based solely on purchasing behavior at an individual retailer. However, state-of-the-art e-commerce now requires data input from email, social media, browser search data, clickstream data, and credit card reporting sites. SAP Data Hub enables you to easily build this pipeline to feed real-time recommendations into an active session on a purchasing website. This information can greatly increase the revenue per transaction metric that is critical to success.

Requirements

As your IT organization evaluates solutions to manage data growth and migration challenges, here are some key requirements to consider:

With a solid understanding of your requirements, you can begin to design the solution. The following section outlines the key concepts in the software architecture of the SAP Data Hub reference configuration.

Software Architecture

SAP Data Hub

SAP Data Hub offers data management capabilities to help customers manage their growing volume of data. This solution combines data governance, management of data pipelines and data integration, using a single visual interface and without the need for moving data into a central data warehouse. Figure 2 shows a high-level view of the architectural components designed to handle a wide range of enterprise applications scenarios. The optional Hadoop cluster can be used as the main software platform for handling the composition of application data.

Requirement Type	Details
Existing Data Stores	Access data from a variety of data sources, including Hadoop data lakes, object stores, databases and data warehouses, both in the cloud and on-premises. • Perform data transformations, data quality and data preparation processes. • Define data pipelines and streams. • Embed and productize scripts, programs and algorithms of the Data Scientist. • Productize open libraries or ML algorithms in one framework.
Distributed Data Processing	Distribute computational tasks to the native environments where the data reside Remote process scheduling: • SAP Business Warehouse process chains • SAP Data Services dataflows • SAP HANA Smart Data Integration FlowGraph
Governance	Establish and manage zones in a landscape with attached policies and services levels. Security and Access Control capabilities
Orchestration	Workflow creation of operations and processes across the landscape, with monitoring and analysis capabilities Execution of end-to-end data processes, starting with the ingestion of data into the landscape (e.g., the data lake), including data processing, and leading up to the delivery or integration of the resulting data into enterprise processes and applications.
Data Ingestion and Processing	Data integration, cleansing, enrichment, masking and anonymization.
Data Discovery	Data profiles for big data sets, showing quality and comprehensive structure information. Ability to crawl, discover and tag data elements. Expose discovered data for further usage.
Scalability	Scalable architecture, from small to big, test to production deployment.

Deployment	Easy deployment, using a proven-to-work combination of the components.
Fault Tolerance	Single-component errors will not lead to the whole system being unavailable.
Ease of Management/Ops	Reduced complexity for solution management.
Physical Footprint	Compact solution that works within your existing infrastructure models.
Flexibility	Flexible building block approach that allows sizing according to customer needs.
Security	Provides the means to secure customer infrastructure.
High performance	Best practices are designed into the solution to ensure the best performance results.

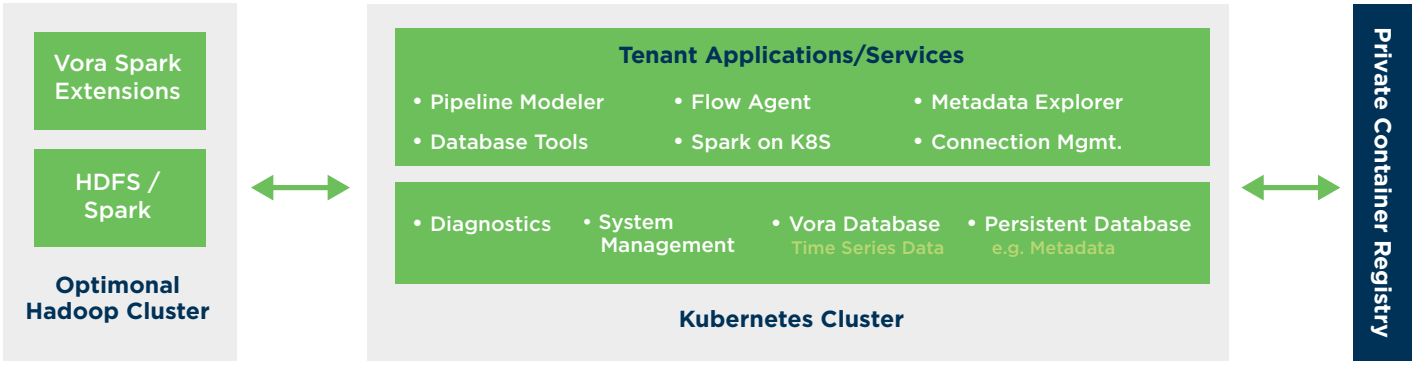
Tenant Applications and Services are the core of SAP Data Hub. SAP Data Hub provides various tools for development and administration, as well as applications that are accessible through the SAP Data Hub application launchpad.

- *SAP Data Hub Pipelines are the connectors between the various SAP Data Hub data sources. They provide reusable, configurable operations to process data from the various sources (including CSV files, web services APIs, and SAP's data stores) and can be flexibly designed.*
- *The SAP Data Hub Modeler allows for the creation and configuration of such pipelines through a graphical user interface.*
- *The Metadata Explorer provides information about the location, attributes, quality and sensitivity of data. With this information, you can make informed decisions about which datasets to publish and determine who has access to use or view information about the datasets.*
- *The Connection Management block enables connections to managed systems or external storage. Services such as Amazon S3, Google Cloud Services, Microsoft Azure (ADL, WASB), data services or Hadoop HDFS can be connected, as well as databases (Oracle, SAP HANA, SAP VORA) or business warehouses (SAP BW).*

SAP Vora Distributed Database

SAP Vora is a horizontally scalable, distributed database that can store and process structured data, time-series data (i.e., IoT streams), graph data and semi-structured documents in-memory and/or on disk. SAP Vora is only available with SAP Data Hub, running in Kubernetes as a fully containerized application.

Figure 2. *Data Hub Architecture*



It can store analytics data in Kubernetes pods, as well as provide a bi-directional Spark2 interface between SAP Data Hub and an optionally co-located Hadoop cluster. Like SAP Data Hub, Vora requires a Kubernetes cluster of at least three Worker Nodes, but runs alongside Data Hub on the same cluster.

Persistent Database

This database holds all of the required persistent data required by SAP Data Hub (e.g., metadata). This instance is automatically installed, sized, and maintained as part of the overall Data Hub installation process. No special consideration is required.

Container Registry

SAP Data Hub requires a container registry for container images. This can be a publicly accessible site or a private collection of workload images. Other public or private registry sites can be used to provide files such as Helm charts to deploy complete services. Although the private private container registry is not part of the SUSE CaaS Platform, you can either:

- *Build an on-premises instance using the Containers Module Add-on included with SUSE Linux Enterprise Server for SAP, along with the SUSE Portus (port.us.org) package. or:*

- *Deploy this as a container directly on SUSE CaaS Platform. Portus is an open source on-premises authorization service that enables users to administrate and secure their private container registries with fine-grained control.*

Optional Hadoop Cluster

An optional Hadoop cluster can be built on dedicated nodes and co-located with SAP Data Hub. This associated Hadoop Data Lake can be used as a local computational/storage medium for SAP Data Hub original and uploaded content. The SAP Data Hub Spark Extensions are used to interface with the Spark2 environment on the Hadoop cluster for processing and storing data. When utilizing this cluster, Data Hub users can leverage the analytical strengths of SAP Vora to analyze and store data in HDFS through the SAP Data Hub Vora Spark Extension. SUSE has extensive experience deploying bare-metal and virtualized Hadoop clusters on SUSE Linux Enterprise Server. While this Hadoop cluster uses dedicated nodes, its HDFS storage is built on block storage from the **SUSE Enterprise Storage** cluster that also serves SAP Data Hub

SUSE CaaS Platform

SUSE CaaS Platform is an integrated software platform that automates the tasks of building, managing and upgrading Kubernetes clusters. It combines the benefits of an enterprise-ready operating system with the agility of an orchestration platform for containerized applications such as SAP Data Hub.

While there are several top-tier Kubernetes offerings in the market, SUSE CaaS Platform stands out for its ease of installation and configuration, DevOps integration **(via SUSE Cloud Application Platform)** and enterprise-level operability and scalability.

One of the biggest challenges for Kubernetes operators is matching the scalability of the node-level infrastructure with that of the overlaying container infrastructure. Inconsistently applied software changes, as well as node configuration drift, create ticking time bombs in production Kubernetes clusters.

SUSE CaaS Platform (figure 3) resolves these problems with a combination of SUSE MicroOS as the container host operating system and Salt for configuration management. SUSE MicroOS is a mission-specific derivative of SUSE Linux Enterprise Server (SLES). While MicroOS leverages the same codebase and packages, its implementation ensures that software changes are applied atomically and within a snapshot-protected environment. The combination of MicroOS and Salt guarantees that all nodes in a cluster are always in a known and consistent state. The troubleshooting nightmares of discovering a single node with a partially failed configuration or software change are a thing of the past.

A SUSE CaaS Platform (figure 4) consists of the following node types:

Administration Node
The Administration Node of the SUSE CaaS Platform manages the deployment of the cluster and runs central services such as:

- *Velum: Web-UI dashboard used to administer the cluster.*
- *Salt Master: Manages the configuration of the cluster nodes.*
- *MariaDB Database: Stores Velum data and Salt master daemon events.*
- *Dex Identity Service: Provides user authentication and a robust, role-based access control (RBAC) system.*

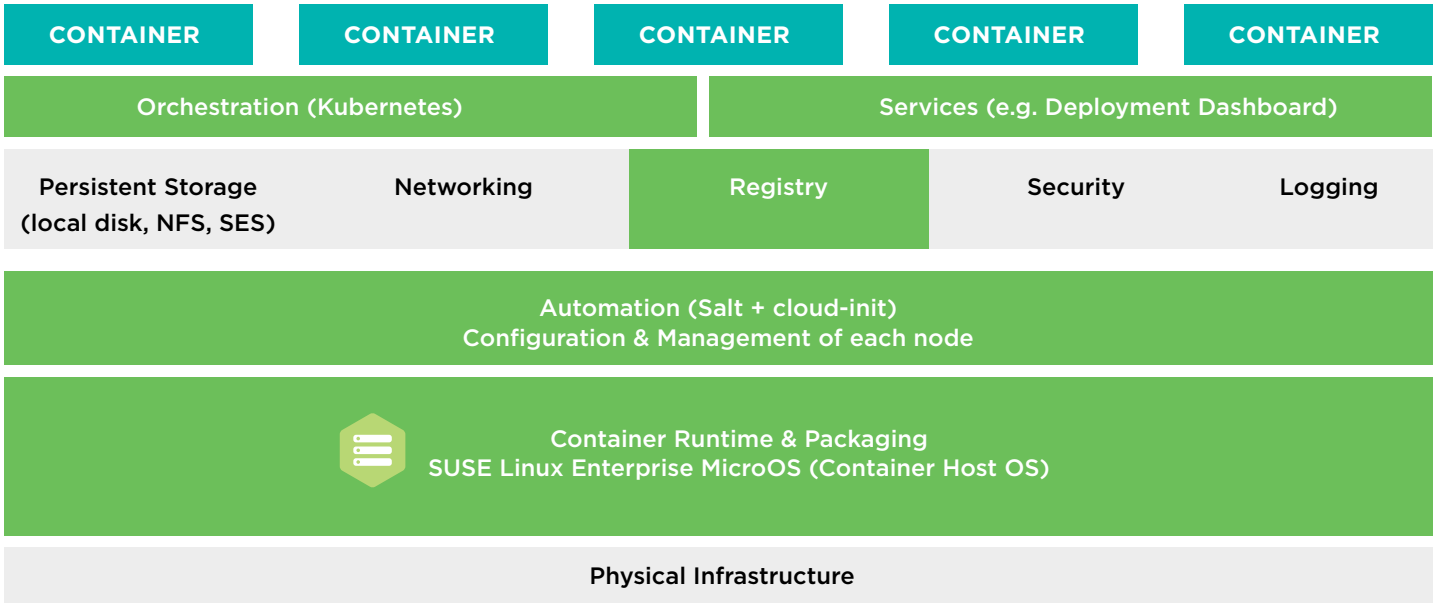


Figure 3. *SUSE CaaS Platform Architecture*

Kubernetes Master Nodes

The CaaS Platform Master Nodes maintain the Kubernetes control plane services. These services run as containers on the Master Nodes. While three or more Master Nodes (always an odd number) are required for high availability of the Kubernetes control plane, a single Master Node is acceptable for demonstration purposes.

Kubernetes Worker Nodes

The CaaS Platform Kubernetes Worker Nodes run the SAP Data Hub application containers. SAP Data Hub requires a minimum of three Kubernetes Worker Nodes (four worker nodes for production). SUSE currently supports CaaS Platform clusters of up to 150 nodes. Additional Worker Nodes can be added to a Production CaaS Platform cluster non-disruptively.

(Note: SAP specifies that each worker node must have a least 8 cores and 64GB of main memory.)

Optional SUSE Cloud Application Platform

SUSE Cloud Application Platform is a modern application delivery environment used to bring an advanced cloud-native DevOps experience to container-based infrastructure. SUSE's implementation is based on the open source **Project Eirini,** which uses Kubernetes to orchestrate application containers while maintaining the Cloud Foundry user experience. This Platform as a Service (PaaS) environment is used by developers to streamline lifecycle management of traditional and cloud-native applications. Together, these technologies accelerate innovation, improve IT responsiveness, and maximize return on investment.

Storage Architecture

The storage layer of this solution leverages the Software Defined Storage capabilities of SUSE Enterprise Storage (SES). SES is a commercially supported distribution of the Ceph enterprise-grade, scale-out storage solution. SAP requires a certified solution for storage that supports Rados Block Devices as well as Dynamically Provisioned Volumes. (See SAP Note 2686169 for certified storage options.)

(Note: SAP Data Hub 2.x no longer supports the NFS protocol. See SAP Note 2712050.)

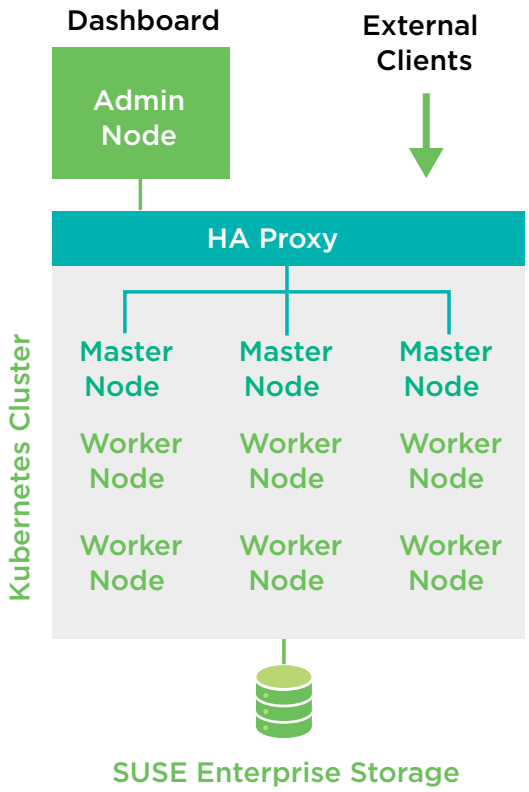


Figure 4. *SUSE CaaS Platform Node Configuration*

Ceph is a scale-out, distributed object store that provides excellent performance, scalability and reliability. In most use cases, clients use Linux kernel libraries to read and write object and block data directly to/from a storage node in the SES cluster. SES also provides gateway options to support data access via iSCSI, NFS, S3 and Swift protocols.

The storage capacity of the SES solution can be expanded easily by integrating additional storage nodes into the cluster. Existing storage nodes will take care of redistributing the data to the newly added nodes without interrupting the availability of storage services to the clients.

SES provides a reliable, scalable storage layer for the complete solution, which supports:

- *Dynamically provisioned block storage volumes to the pods running on SUSE CaaS Platform*
- *(Optionally) Block storage volumes for the co-located Hadoop cluster nodes, if configured*
- *Object storage through an S3-API-compatible interface, for additional data storage and backups*

Dynamically Provisioned Storage Volumes

In addition to providing block storage to the optional Hadoop cluster, a pod running on CaaS Platform can gain access to dynamically provisioned Kubernetes persistent volumes (PV) through Kubernetes persistent volume claims (PVC). Persistent volumes are created as block devices in the supporting SES cluster. CaaS Platform uses persistent volume claims (PVCs) to obtain dynamically provisioned persistent volumes through the Software Defined Storage mechanisms in SES. When a PVC is removed, the persistent volume and its associated block storage device in SES are automatically removed.

Software and Systems Management

While SAP Data Hub doesn't require an external SAP HANA instance in order to function, most users of this solution will be attaching to an existing HANA database to build their data pipelines. After assembling this combined data pipeline and writing to your HANA database, you can take advantage of **SAP Advanced Analytics Processing** capabilities, including machine learning/predictive analytics, spatial intelligence (location awareness) and streaming data processing. The scale-out capabilities of SAP HANA support rapid data growth, but it is important to have a dependable method of updating your SAP HANA servers. SUSE Manager can mirror CaaS Platform installations and update packages to help enforce consistency across your organization. SUSE Manager can also analyze the container images in your private container registry as well as containers running on your SUSE CaaS Platform for known vulnerabilities, outstanding patches, or pending package updates. **SUSE Manager** enables you to efficiently manage a set of Linux systems and keep them up to date. An SAP HANA scale-out setup offers these benefits:

Reduced Complexity of Managing SAP HANA Environments

- *Ensure consistent management of SAP HANA and all other cluster systems.*
- *Manage your data environment across physical, virtual and cloud environments.*
- *Manage your channels effectively.*

Create/Manage Development, QA and Production Channels

- *Add and manage third-party channels.*
- *Simplify compliance.*

Audit the Patch Status for SAP HANA and Subsystems

- *Track the configuration changes and make sure all administrators have the right authority for changes.*
- *Slash costs of ownership.*

Automate System Management Tasks for SAP HANA and All Other Subsystems

- *Leverage a single, web-based interface to see the status of all your servers.*
- *Use your resources effectively.*

Hardware Architecture

This reference defines a private proof-of-concept cluster. Other guides will address cloud-based and production environments. The proof-of-concept cluster is a starting point for prototyping real-world applications that are meant to go into production. As such, a proof-of-concept cluster can easily be grown into a production environment. Further, a private deployment is useful for secure, in-house prototyping, but is not restricted from using cloud-based applications and data. SAP Data Hub can manage data from locations both behind the firewall and in the cloud.

The backbone of the application environment is the Kubernetes cluster, as implemented in the SUSE CaaS Platform software. As described in the architecture section above, this cluster can be implemented on low-cost, commodity hardware, typically 1U racked boxes with 2-socket processors. The application environment reflects a key aspect of growth, from prototyping to production. In a PoC, there are typically only a few application containers, because developers are gaining experience building, deploying and managing code as containers on the new software environment. As such, there are fewer demands on the hardware for resources, particularly main memory. Memory can easily be expanded as the cluster needs to grow. Processor speed, on the other hand, is something to consider carefully. It is not as easy to swap processors, so choosing a medium to high-end Xeon processor with many cores is a good idea.

It is recommended to start with a minimum 2-socket Xeon processor with 8 to 12 cores per CPU. Memory can start at 64GB. Local storage on the cluster is used primarily for the operating system and any temporary application data. A common choice is a RAID1 configuration of operating system disks, each with a minimum of 256GB. SSD is the recommended choice, because it speeds operations and doesn't add considerable cost to the cluster in this case. Networking is another key performance area, so 10GbE NICs are a minimum recommendation. You can start with 1 NIC, but that is a bare minimum; 2 or more NICs will be needed for specific network requirements, based on your application needs. In particular, high-bandwidth data pipelines will move a lot of data to the cluster storage and can easily saturate a single NIC. Consider mapping 1 high-speed NIC to those storage requirements.

As described in the SUSE CaaS Platform section above, a minimum of 5 nodes will be required: 1 as the administration node, 1 master node and 3 worker nodes. Any SUSE YES certified IHV platform could be used for the physical nodes of this deployment, as long as the certification refers to the major version of the underlying SUSE operating system required by the SUSE CaaS Platform release. One key benefit of this data analytics implementation is that IHV industry-standard servers can fulfill each of the resource node's computational and additional storage needs. To reduce the time spent on hardware specification for an initial proof-of-concept implementation, the hardware should be general purpose and allow for a wide range of configuration options. The following sections detail the attributes of the IHV overall portfolio and some specific recommendations of platform models.

Given the overlay of all the software-defined infrastructure components, many choices of industry standard servers can form the basis for this solution. You can select any of your favorite hardware providers, obtain the desired number of nodes and build your solution. The appendices contain some respective component and resource sizing guidelines for each of the node roles.

Compute

The following considerations for the system platforms should be emphasized:

- *Ensure that all similar system devices are consistent and up to date with regard to BIOS/UEFI/device firmware versions, to reduce potential troubleshooting issues later.*
- *Reset the BIOS setup configuration to the default setting, in order to have a known baseline configuration for consistency.*
- *If possible, set up RAID1 mirroring on the storage controller across a pair of drives for the operating system installation.*

Storage

As discussed in the Storage Architecture section, the storage layer of this solution leverages the software-defined-storage capabilities provided by SUSE Enterprise Storage.

SAP Data Hub and SUSE CaaS Platform are the base framework for a data analytics environment. The data analytics you execute will be defined by a set of application containers that run on the SUSE CaaS Platform. These containers will access data across your company's infrastructure and may store derived results in SUSE Enterprise Storage. As you define the workflow of your data analytics applications, you will need to access data across many different storage systems in your enterprise. This access is beyond the scope of the architecture described herein, but it is important to understand that data access will be required for a wide range of disparate storage systems.

The SUSE Enterprise Storage cluster can be utilized to store both intermediary and final results from the data analytics pipeline. In other words, as your data analytics applications derive new data results, those are typically stored in SUSE Enterprise Storage. This enables the data analytics environment (SAP Data Hub, SUSE CaaS Platform, and your data analytics applications) to be logically organized in one physical location.

Network

The following considerations for networking should be emphasized:

- *Configure 802.3ad for system port bonding in order to get the maximum performance of bonded network interfaces.*
- *Ensure that all similar switching devices are consistent and up to date with regard to firmware versions, to reduce potential troubleshooting issues later.*

Summary

SUSE CaaS Platform is an excellent environment for creating an SAP Data Hub implementation. This composable infrastructure enables you to define appropriate hardware from software descriptions. This means you can easily scale, adjust and customize your environment to fit your needs as you move from a proof-of-concept toward a production environment.

SUSE CaaS Platform is an enterprise Kubernetes container platform that provides software infrastructure for not only the SAP Data Hub software described in this reference, but also the data analytics applications you will build to ingest and manage your data.

All of the software environments in this reference architecture are supported products and have been tested to work together on industry-standard x86-64 gear.

Join the best. Run your SAP solutions on SUSE

Appendices/References

SAP Data Hub

- **https://www.sap.com/products/data-hub.html.**
- *Release Note* – **https://launchpad.support.sap.com/#/notes/2721708**
- *Prerequisites* – **https://launchpad.support.sap.com/#/notes/2686169**
- *Install Guide* – **https://help.sap.com/viewer/e66c399612e84a83a8abe97c0eeb443a/2.4.latest/en-US**

SUSE CaaS Platform

- **https://www.suse.com/products/caas-platform**
- *Documentation* – **https://www.suse.com/documentation/suse-caasp-3**

SUSE Enterprise Storage

- **https://www.suse.com/products/suse-enterprise-storage**
- *Documentation* – **https://www.suse.com/documentation/suse-enterprise-storage-5**

SAP Software Bill of Materials

BILL OF MATERIALS – SAP DATA HUB SOFTWARE

Role	Quantity	Description	Notes	Installation
Software	1	SAP Data Hub 2 – Foundation – Installation Product	Select the desired version when downloading from launchpad.support.sap.com. Please consult the installation guide @ help.sap.com for SAP Data Hub prerequisites.	Required
Software	1	SAP Data Hub 2 – Spark Extension – Installation Product	Select the above corresponding version when downloading from launchpad.support.sap.com. Please consult the installation guide @ help.sap.com for SAP Data Hub prerequisites.	Optional

SAP Software Bill of Materials

BILL OF MATERIALS – SAP DATA HUB SOFTWARE

Role	Quantity	Product #	Description	Notes
Software	1	874-007633	SUSE CaaS Platform, x86-64, 1-2 Sockets with Unlimited Virtual Machines, L3-Priority Subscription, 3 Year	For each node of the deployment

BILL OF MATERIALS – SUSE ENTERPRISE STORAGE SOFTWARE

Role	Quantity	Product #	Description	Notes
Software	1	874-007044	SUSE Enterprise Storage Base Configuration, x86-64, 4 OSD Nodes with 1-2 Sockets, L3-Priority Subscription, 3 Year	Includes 4 OSD Nodes plus 6 infrastructure nodes (e.g., Admin, Mon, gateway).
Software	1	874-007046	SUSE Enterprise Storage Expansion Node, x86-64, 1 OSD Node with 1-2 Sockets, L3-Priority Subscription, 3 Year	Includes 1 additional OSD Node plus 1 infrastructure node.l

IHV Hardware Bill of Materials

(Tip: Any SUSE YES certified IHV platform could be used for the physical nodes of this deployment, as long as the certification refers to the major version of the underlying SUSE operating system required by the SUSE CaaS Platform release

BILL OF MATERIALS – INDUSTRY STANDARD SERVERS (KUBERNETES CLUSTER)

Role	Quantity	Component	Description
Admin Node	1	CPU	>= 2 GHz, >= 4 cores, >= 1 socket
		Memory	>= 16 GB RAM , Divided evenly across CPU sockets
		Storage (OS)	>= 64 GB Disk, >= 1 drive
		Network	>= 1+ GbE, >= 1 NIC
Master Node	3	CPU	>= 2 GHz, >= 8 cores, >= 1 socket
		Memory	>= 32 GB RAM, Divided evenly across CPU sockets
		Storage (OS)	>= 128 GB Disk, 2 drives (RAID1)
		Network	>= 1+ GbE, >= 1 NIC
Worker Node	4	CPU	>= 2 GHz, >= 8 cores, >= 1 socket
		Memory	>= 64 GB RAM, Divided evenly across CPU sockets
		Storage (OS)	>= 128 GB Disk, 2 drives (RAID1)

BILL OF MATERIALS - INDUSTRY STANDARD SERVERS (STORAGE CLUSTER)

Role	Quantity	Component	Description
Admin Node	1	CPU	>= 2 GHz, >= 4 cores, >= 1 socket
		Memory	>= 16 GB RAM, Divided evenly across CPU sockets
		Storage (OS)	>= 64 GB Disk, >= 1 drive
		Network	>= 10+ GbE, >= 1 NIC
Monitor Node	3	CPU	>= 2 GHz, >= 8 cores, >= 1 socket
		Memory	>= 32 GB RAM, Divided evenly across CPU sockets
		Storage (OS)	>= 128 GB Disk, 2 SSD drives (RAID1)
		Network	>= 10+ GbE, >= 2 NIC
Storage/OSD Node	4	CPU	2 GHz of logical CPU core per OSD drive, >=1 socket
		Memory	2 GB RAM per TB of OSD capacity, Divided evenly across CPU sockets
		Storage (OS)	>= 128 GB Disk, 2 drives (RAID1)
		Storage (OSD)	>= 128 GB Disk, >= 8 drives
		Storage (OSD journal)	>= 128 GB Disk, Ratio SSD:spinner-OSD = 1:5-7, NVMe:spinner-OSD =1:10
		Network	>= 10+ GbE, >= 2 NIC

Additional contact information and office locations:
www.suse.com

www.suse.com