Running SAP Data Hub with SUSE Container as a Service Platform
Overview, Installation, Outlook

Markus Gürtler
Sr. Technical Alliance Manager & Team Lead
mguertler@suse.com
Agenda

- About the SAP Linux Lab and the SUSE/SAP Alliance
- SAP Data Hub and where it runs
- SUSE Container as a Service Platform & SUSE Enterprise Storage
- Installation of SAP Data Hub 2.4 on CaaS Platform 3.0
- Outlook and other cool stuff ;-)}
About SUSE at SAP
About the SAP Linux Lab

- Founded in 1999 (20th anniversary this year!)
- Interdisciplinary team with employees from SAP, Hard- and Software vendors
- Innovation projects
- Development support
- 3rd level customer support
- Our mission: Make Linux #1 platform for SAP workloads
Unrivaled Relationship Making SUSE the Smart Choice for SAP Workloads

- 95%+ of SAP HANA customers rely on SUSE Linux
- 70% of SAP NetWeaver-on-Linux customers rely on SUSE Linux
- 20 years of joint testing and development at the SAP LinuxLab
- Seamless support from SAP and SUSE
- SLES for SAP is the #1 choice for SAP and SAP customers
- SUSE OpenStack Cloud & SUSE Enterprise Storage power SAP Cloud Platform
- Joint collaboration on Cloud Foundry
- Data Hub 2 validated with SUSE Container as a Service Platform
SUSE Solutions for SAP

SAP HANA, S/4 HANA, SAP NetWeaver, ...

SAP Data Hub

SAP Cloud Platform

SUSE Linux Enterprise Server for SAP Applications

SUSE Container as a Service Platform

SUSE Open Stack Cloud SUSE Enterprise Storage

High Availability

on-premise and in the cloud
SAP Data Hub and where it runs
SAP Data Hub
Enterprise Data

- Data Sharing
- Data Pipelining
- Data Governance

SAP Data Hub

- Big Data Lakes
- BI
- Cloud Apps
- Enterprise Apps
- 3rd Party Apps
- MDM, GDM
- EDW, Data Marts
SAP Data Hub 2 - Overview

Connected Apps
- SAP S/4HANA
- SAP BW/4 HANA
- SAP HANA
- Other Databases
- SAP Cloud Applications
- 3rd Party Apps

Data Storage
- Hadoop
  - Data Lake, Machine Learning, Predictive Analytics
- Cloud Storages
  - AWS S3, GCP GCS
  - Azure ADL, WASB

SAP Data Hub
- Data Discovery & Pipelines
- Orchestration & Monitoring
- Ingestion & Integration

Container as a Service Platform
SUSE Enterprise Storage
SUSE Linux Enterprise Server for SAP Applications
SAP Data Hub 2 - Overview

Connected Apps
- SAP S/4HANA
- SAP BW/4 HANA
- SAP HANA
- Other Databases
- SAP Cloud Applications
- 3rd Party Apps

SAP Data Hub
- Data Discovery & Pipelines
- Ingestion & Integration

containerized

Data Storage
- Hadoop
  - Data Lake, Machine Learning, Predictive Analytics
- Cloud Storages
  - AWS S3, GCP GCS
  - Azure ADL, WASB

Container as a Service Platform
SUSE Enterprise Storage
SUSE Linux Enterprise Server for SAP Applications
Architectural Overview: SAP Data Hub on SUSE CaaS Platform

**SAP Data Hub Foundation**
- Tenant Applications / Services (Managed by System Management)
  - Pipeline Modeler
  - Flow Agent
  - Metadata Explorer
  - Database Tools
  - Spark on K8S
  - Connection Management

**SUSE CaaS Platform**
- Kubernetes Cluster
- Optional Hadoop Cluster
  - Vora Spark Extension
  - HDFS / Spark
- Docker Registry

**Containerized**
- SUSE CaaS Platform

**Not Containerized**
- Tenant Applications / Services
- Pipeline Modeler
- Internal SAP HANA

**Components**
- Diagnostics
- Vora Database
- System Management
- Internal SAP HANA
- Optional Hadoop Cluster
- Vora Spark Extension
- HDFS / Spark
- Docker Registry
SAP Data Hub Validation

SAP Data Hub 2.3 validated for
SAP Data Hub 2.4

SUSE Container as a Service Platform 3.0
SUSE Enterprise Storage

see SAP note #2686169 - Prerequisites for installing SAP Data Hub 2
SUSE Container as a Service Platform
SUSE CaaS Platform

Speed application delivery to improve business agility

SUSE CaaS Platform is a Kubernetes-based container management solution used by application development and DevOps teams to deploy, manage, and scale container-based applications and services.
SUSE CaaS Platform

What are the main components of the SUSE CaaS Platform?

- Kubernetes
- Salt
- SUSE MicroOS

Easy to deploy, manage, update ...
SUSE CaaS Platform

- Orchestration (Kubernetes)
- Services (e.g. Deployment Dashboard)
- Persistent Storage (local disk, NFS, SES)
- Networking
- Registry
- Security
- Logging
- Automation (Salt + cloud-init)
  Configuration & Management of each node
- Container Runtime & Packaging
  SUSE Linux Enterprise MicroOS (Container Host OS)

(Physical) Infrastructure
SUSE Enterprise Storage
SUSE Enterprise Storage

Unlimited Scalability with Self Managing Technology

Management Node

Object Storage

Block Storage

File System

Monitor Nodes

Storage Nodes
Deploying SAP Data Hub on CaaSP with SES
Prerequisites
Minimum Requirements from SAP

Test
• 3 x worker nodes with $\geq$ 32 GB RAM and $\geq$ 4 cores per node
• 1 x admin node, 1 x master node, 1 x docker registry

Production
• 4 x worker nodes with $\geq$ 64 GB RAM and $\geq$ 8 cores per node
• 1 x admin node, 1 x master node, 1 x docker registry
Lenovo  Reference Architecture

Created by Lenovo, SUSE and SAP

• SAP Data Hub 2.4 on SUSE CaaSP 3.0
• Hadoop cluster on the same infrastructure
• On-premise S3-Storage for data
  (a benefit of using SUSE Enterprise Storage)

Minimum Requirements

• CaaS Platform: 3 x worker nodes with 192 GB RAM and 1 Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz with 8 cores
• 1 x admin node, 1 x master node, 1 x docker registry
• SUSE Enterprise Storage: 4 x nodes with 2 x Intel(R) Xeon(R) CPUs, 192 GB RAM and 4 x 960 GB SSD for data
Lenovo Reference Architecture

Compute Nodes
- Lenovo ThinkSystem SR650
- 1 or 2 CPUs, 192 or 384 GB RAM per Node
- Drives only for OS, Data on Storage
- SLES 12 SP3 / SUSE CaaSP 3

Redundant Fabric for Networking & Storage
- 10GbE for Compute, 25GbE for Storage, 1GbE for Management

Storage Nodes
- Lenovo ThinkSystem SR650
- 4+ 960GB SSD Drives per Node
- SUSE Enterprise Storage 5

Hadoop Data Lake

SAP Data Hub
- 3+ Kubernetes Worker
- 1+ Kubernetes Master
- Docker Image Repository for SAP Data Hub
- CaaSP Cluster Management

Internal Fabric and Connectivity to Managed Systems
- or external DataSources / Sinks, Up to 100GbE

Storage for everything
- Dynamically provisioned storage for SAP Data Hub
- Storage for Docker Registry
- Storage for Hadoop nodes
- S3 Object Storage
Bringing all pieces together

- SAP Data Hub 2.4
- SUSE Container as a Service Platform 3.0 cluster
- SUSE Enterprise Storage 5 cluster
- Linux Workstation with Kubernetes client, Docker, Helm, Data Hub 2.4 Installer
- Private Docker Registry SLES 12/15 with Docker Registry and Portus
Installation
Installation order

1. Linux Workstation
   with Kubernetes client,
   Docker, Helm,
   Data Hub 2.4 Installer

2. SUSE Enterprise Storage 5 cluster

3. SUSE Container as a Service Platform 3.0 cluster

4. Private Docker Registry
   SLES 12/15 with
   Docker Registry and Portus

5. SAP Data Hub 2.4
1 Prepare Management Workstation

- Install OpenSUSE Leap or Tumbleweed (can also be MacOS, Ubuntu, etc.)
- Install Helm (make sure version fits to Tiller, 2.8.2)
- Install Kubernetes client (incl. kubectl) 1.10, 1.11, 1.12
- Install Docker
- Download kubeconfig to workstation from CaaS Platform dashboard
- Download and unpack SAP Data Hub 2.4 Foundation
2 Install SUSE Storage

• Min. 4 x nodes

• According to SES 5 Deployment Guide
3 Install CaaS Platform 3.0

Install admin node
3 Install CaaS Platform 3.0

Initially configure CaaS Platform

![CaaS Platform Configuration](image-url)
3 Install CaaS Platform 3.0

- Install 4 additional nodes (i.e. via PXE and AutoYast)
- Assign roles in CaaS Platform dashboard
3 Install CaaS Platform 3.0

Completed cluster setup with

1 x admin node
1 x master node
3 x worker nodes
Install Docker Registry

- SLES 12 or SLES 15 with container module

- Install Apache, Docker Registry and Portus

- Generate SSL certificates for container-registry host
  (Signed by a valid CA or self-signed)

- **Configure Docker Registry and Portus** (according to SLES installation guide)

- Distribute the Root CA certificate to all CaaS Platform hosts and to the management workstation

- Test registry with: “docker login https://my-registry:5000 -u <user>"
5 Install SAP Data Hub

- Install SAP Data Hub
  - use the provided **install.sh** script or
  - Maintenance Planner and Software Lifecycle Plugin 1.0

- See https://help.sap.com/viewer/p/SAP_DATA_HUB
Install SAP Data Hub

```bash
./install.sh
--cert-domain=master-node.testlab.intern
--sap-registry-login-type 2
--sap-registry-login-username=<S-USER>
--cert-domain=master-node.testlab.intern
--sap-registry-login-password=<password>
--vora-admin-username=admin
--vora-admin-password=<password>
--vora-system-password=<password>
--extra-arg=vora-vsystem.vSystem.nodePort=32123
--extra-arg=vora-dqp.components.txCoordinator.nodePort=30343
--extra-arg=vora-cluster.components.txCoordinator.hanawire.portNumber=32215
--namespace=datahub24
--registry=container-registry.testlab.intern:5000
--interactive-security-configuration=no
--enable-checkpoint-store=no
--image-pull-secret=regcred
--pv-storage-class=nfs-client
--disk-storage-class=nfs-client
--dlog-storage-class=nfs-client
--consul-storage-class=nfs-client
--hana-storage-class=nfs-client
--accept-license
--confirm-settings
--install
```
Install SAP Data Hub

1. Pulls images from Private Docker Registry
2. Deploys with HELM charts to SUSE Enterprise Storage 5 cluster
3. Validates Data Hub installation

SAP Data Hub 2.4 runs in containers / starts containers
SUSE Container as a Service Platform 3.0 cluster
SUSE Enterprise Storage 5 cluster

Data Hub Installer
Running at the Management Workstation
SAP Container Registry
Install SAP Data Hub

```
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

# Examples of commands
1. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

2. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

3. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

4. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

5. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
```

```
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

# Examples of commands
1. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

2. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

3. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

4. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

5. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
```

```
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

# Examples of commands
1. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

2. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

3. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

4. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

5. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
```

```
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

# Examples of commands
1. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

2. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

3. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

4. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

5. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
```

```
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation
us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

# Examples of commands
1. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

2. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

3. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

4. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD

5. `us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
   us@is3507:~$ datehub/SAPDataHub-2.3.152-Foundation$ echo $INSTALLCMD
```
Take a “Brotzeit”
SAP Data Hub Installation successfully completed ✔️
Demo Video
Outlook
Outlook and other cool stuff!

**SAP Data Hub**
- SAP Data Hub 2.5 (not yet released) for CaaS Platform 3.0 currently in validation
- CaaS Platform 4.0 will be validated with SAP Data Hub as well

**Other Container Projects**
- **SAP Gardner**
  - SUSE helps with OS based on SLE 15 JeOS + Docker + Tooling
  - SUSE helps with development resources
- **SAP Hana as a Service**
  - Cloud service of SAP running in Public Clouds (starting with GCP)
  - SAP HANA runs containerized on SLES based container
  - Container runtime is based on SLES 15 JeOS
Resources

**SUSE CaaS Platform & SUSE Enterprise Storage**

- Product landing-pages
  - https://www.suse.com/de-de/products/caas-platform/
  - https://www.suse.com/de-de/products/suse-enterprise-storage/

- Deployments Guides, Administration Guides & more
  - https://www.suse.com/documentation/suse-enterprise-storage-5/

**SAP Data Hub**

- Product landing-page

- Installation Guide, Administration Guide & more
Unpublished Work of SUSE LLC. All Rights Reserved.
This work is an unpublished work and contains confidential, proprietary and trade secret information of SUSE LLC. Access to this work is restricted to SUSE employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of SUSE. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer
This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. SUSE makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for SUSE products remains at the sole discretion of SUSE. Further, SUSE reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All SUSE marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.
Additional Slides
Operations & Maintenance
Management Tools
command-line and web-based

- SUSE CaaS Platform Management web interface
- Kubernetes Dashboard or kubectl (on your management workstation)
- Portus (Docker Registry) web interface
- OpenATTIC / SUSE Enterprise Storage Manager
- SAP Data Hub web interface
Monitoring & Diagnostics

CaaS Platform and SUSE Enterprise Storage
• Log analysis with Fluentd, Elasticsearch & Kibana
• Monitoring with Promethus and Grafana

SAP Data Hub
• Log analysis with Kibana
• System and application metric monitoring with Grafana
Maintenance

CaaS Platform Update
- Minor version updates via the CaaS Platform web interface
- Major version upgrades via docker & salt
- Refer to the CaaS Platform Administration Guide for more details

SAP Data Hub Platform Update
- via SL Plugin and Maintenance Planner
- alternatively via install-Script “install.sh –update”
- Refer to the SAP Installation guide for more details
SAP Data Hub
SAP Data Hub Configuration

Login to Data Hub
SAP Data Hub Configuration

Do some post-configuration stuff following the SAP Installation Guide
Ready to use...

Start building your data pipelines!