

Lenovo Intelligent Insights with SAP Data Hub

Lenovo

A scalable landscape to orchestrate, manage and drive improved decision making from data

Manage data-driven processes for end-to-end Big Data scenarios

The Data Challenge

Data sets grow rapidly. More and more data sources are being tapped, beyond the traditional enterprise data. Sensor networks, software logging, mobile devices, cameras, microphones, and other new sources of information create huge amounts of data, which holds a business opportunity. At the same time, corporate data landscapes are growing increasingly complex. This makes it hard and costly to capture the maximum value from the available data, by understanding it, working with it across different systems, and also to apply governance in an end-to-end fashion.

Usually data is kept in silos across the enterprise, e.g. in enterprise applications, databases, plain files, Hadoop data lakes, data warehouses, or various forms of cloud storage. Combining data across those silos is needed to unlock its value, but this process is complex, time-consuming and therefore costly. Many of today's data integration tools are point-to-point, complex to use, and highly manual making it challenging to rapidly connect and implement desired data outcomes. The increased complexity of enterprise landscapes also makes it challenging to enabling appropriate and effective governance. In order to trust and rely on data accuracy, an end-to-end governance across all data sources is required otherwise acting on the data, either through analytical or operational applications using the data, is at risk.

Enterprise readiness of Big Data technologies is not a given. When trying to solve the complexity of the data landscape by simply storing all data in a Hadoop data lake, businesses often encounter limited governance, little automation of scheduling data processing, lack of common security and access management, and limited monitoring and tracing capabilities. Also, implementing Big Data initiatives and creating value from them requires very specialized skillsets. These specialized resources can be difficult to find and retain. Managing and processing data across silos is also a challenge, but it is also an opportunity to unlock the value of the data by combining data from different sources to create new insights.

Highlights

- Improves decision making whether data is in the cloud, on premise, in data lakes, or business systems
- Scales your data center landscape to enable new use cases and meet demanding data growth
- Prepares you for future information management needs with an extensible and open system

Built on server infrastructure rated #1 in reliability¹

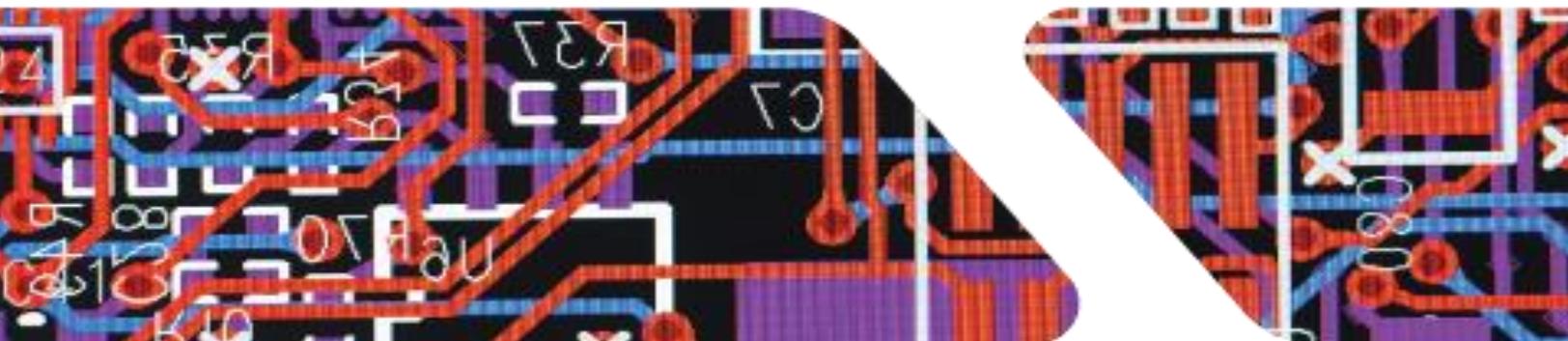
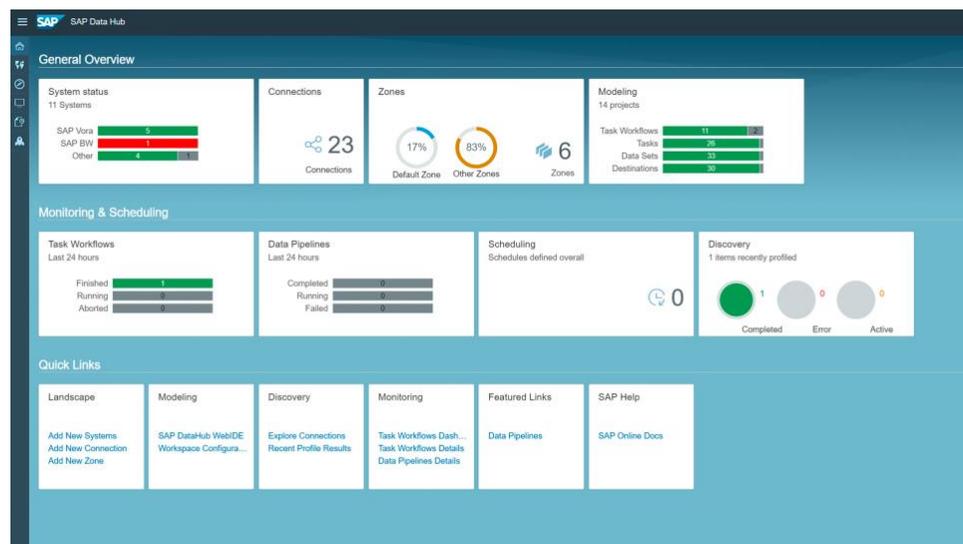




Figure 1: SAP Data Hub
Experience a simpler, more scalable approach to data operations and landscape management



MANAGE DATA-DRIVEN PROCESSES FOR END-TO-END BIG DATA SCENARIOS

Lenovo Intelligent Insights with SAP Data Hub delivers a simpler, more scalable approach to managing complex data landscapes. It provides data integration, processing and governance, and provides visibility into and access across the complex network of data in the modern enterprise. Lenovo Intelligent Insights helps organizations to better understand data sources, interconnections, quality and impact by providing a broad and detailed view of the entire data landscape. Lenovo Intelligent Insights with SAP Data Hub provides a single data management pane for data from various data sources, like Hadoop, cloud storage, SAP HANA, business applications, and more. This enables enterprises to discover new business opportunities, resolve emerging data issues, and ensure data flowing to where it needs to be.

Lenovo Intelligent Insights with SAP Data Hub allows you to create powerful data pipelines that access, harmonize, transform, process, and

move information from a variety of sources to a variety of destinations. Data pipelines can be easily and quickly created in a single, visual design environment.

Most of the data that Lenovo Intelligent Insights with SAP Data Hub will process typically comes from non-SAP systems. It is designed to leave the source data in place (even if the aggregated data goes into SAP HANA).

Data Hub uses "push down" processing techniques to enable the source systems to process the data in place as much as possible before pushing results to SAP Data Hub.

Leaving data where it is, is important. Data movement is time-consuming and expensive. It creates superfluous copies of the data, which can be a compliance nightmare.

Pipeline processing is done on SAP Vora, the data processing platform based on Apache Spark. SAP Vora has the option to store the processed data in SAP HANA if available.

Lenovo Intelligent Insights with SAP Data Hub provides an easier way to understand, manage, and get greater value from a complex data landscape, including data held on premise and in the cloud, in data lakes, data warehouses, and data marts. It allows to quickly create data-driven applications and analytics that leverage data from across the organization, and makes it easy to combine and integrate an enterprise landscapes with big data.

DataOps management

Experience a simpler, more scalable approach to data operations and landscape management with enterprise-spanning data governance and easier data integration.

Accelerate data projects

Expand data projects by easily creating powerful data pipelines in a single, visual design environment.

Data-driven applications

Accelerate your business with fast pipeline results enabled by "push-down" distributed native processing.

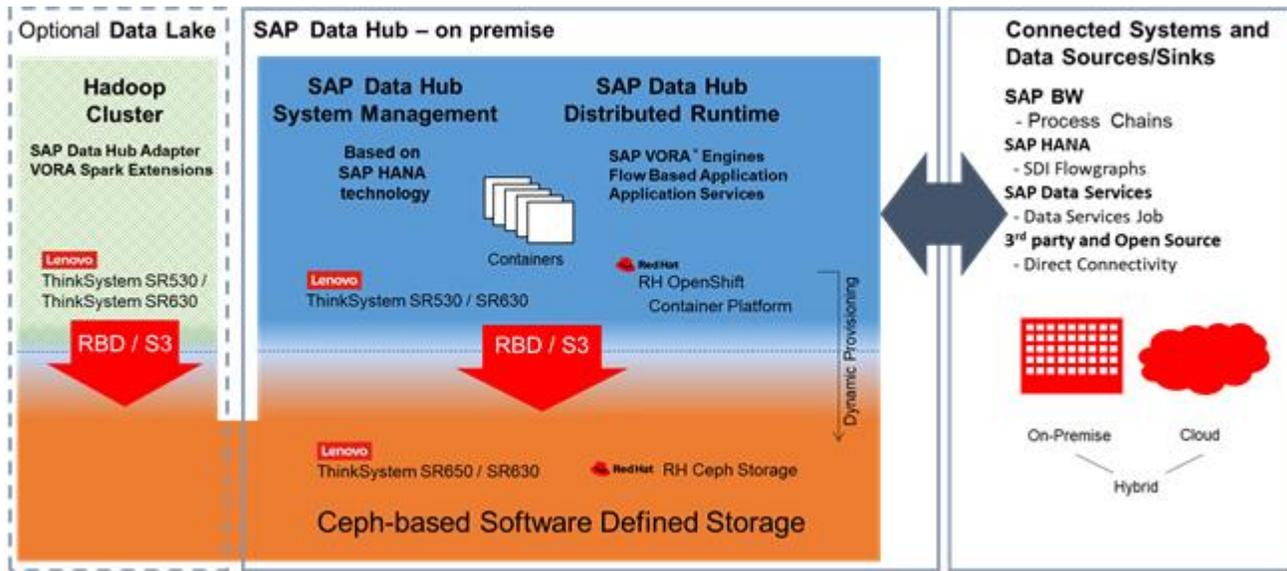


Figure 2: Lenovo Intelligent Insights with SAP Data Hub High Level Architecture

SCALE YOUR DATA CENTER LANDSCAPE

Lenovo Intelligent Insights with SAP Data Hub is a critical starting point to enable a successful Digital Transformation. The solution enables unified analytics of multi-layer, multi-vendor, multi-location environments (on-premises, cloud, hybrid cloud). In addition, it provides executive-level oversight of devices and data.

A complete Lenovo Intelligent Insights with SAP Data Hub implementation is composed of several components including the SAP Data Hub Cockpit, SAP HANA, SAP Data Hub distributed runtime connected systems and solution storage.

The Lenovo Intelligent Insights with SAP Data Hub is an integrated end-to-end

solution built on selected, proven-to-work components. A flexible building block approach guarantees a high scalability, from proof of concept to large production implementations. SAP, Red Hat and Lenovo Best Practices are built into the solution for best availability, performance results and predictable outcomes.

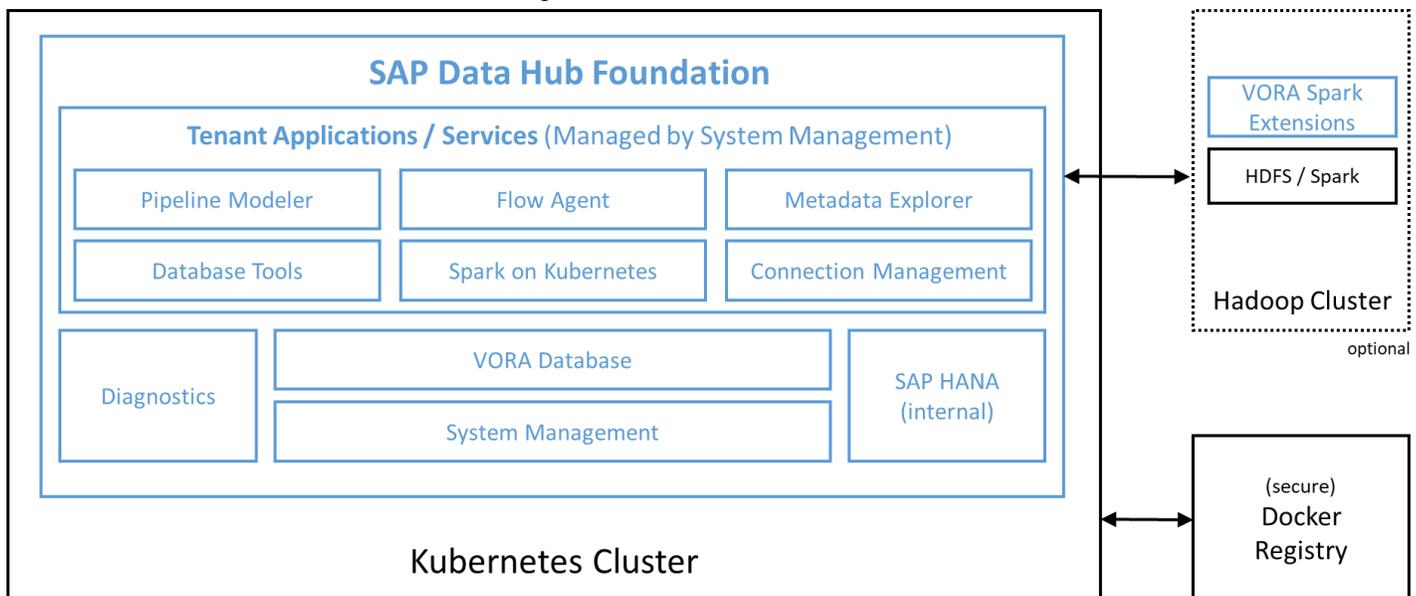


Figure 3: Lenovo Intelligent Insights with SAP Data Hub Component Overview



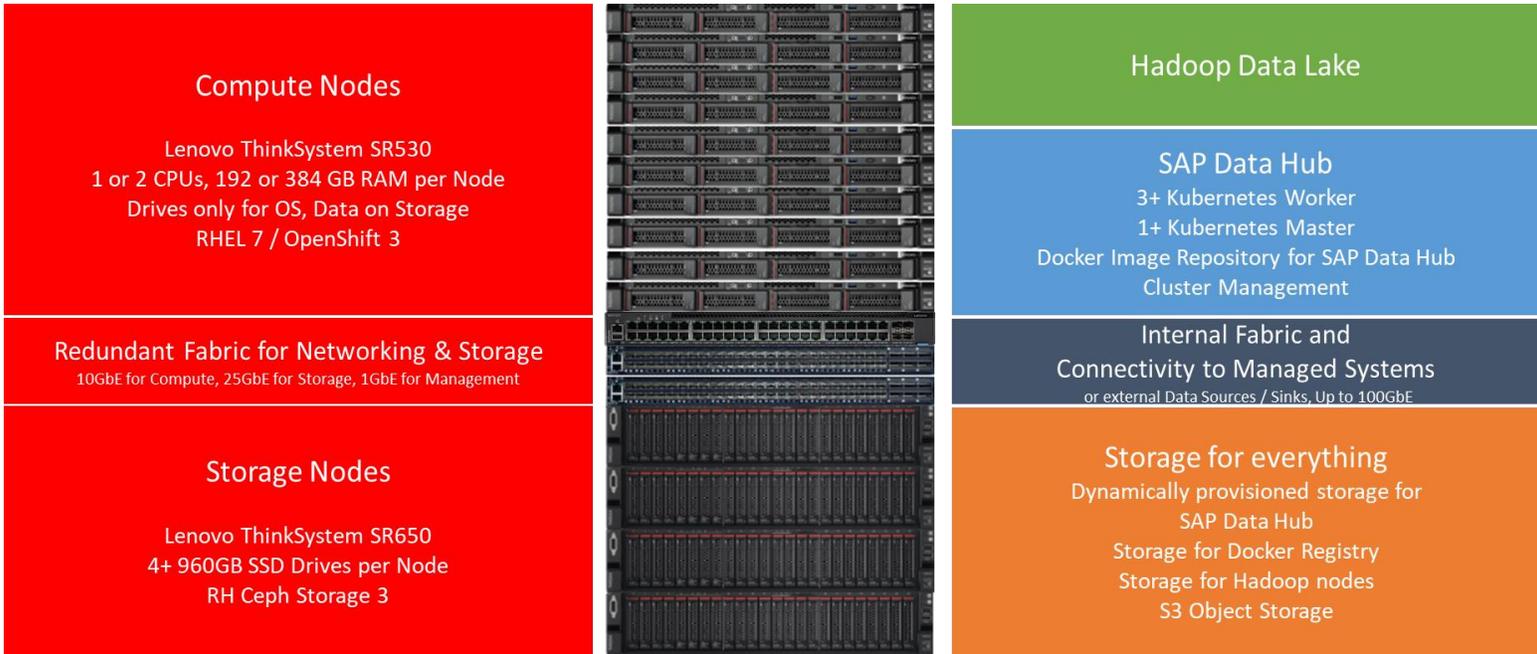


Figure 4: Operational Model for SAP Data Hub with Red Hat OpenShift and Red Hat Ceph Storage.

AVOID THE IMPACT AND RISK OF DOING NOTHING

Lenovo Intelligent Insights with SAP Data Hub solution components include:

- **SAP HANA** The internal SAP HANA database is used to store metadata, such as information about the data pipelines and other configuration data. It does not store any business data, so the memory footprint is very small.
- **SAP Data Hub Systems Management and SAP Data Hub Distributed Runtime** run containerized inside the Red Hat OpenShift Container Platform, provisioning and managing the containers and the application components as needed by SAP Data Hub.

SAP Vora is a distributed database, which stores and processes structured data, graph, data and semi-structured documents in-memory and on disk which provides enriched interactive analytics on data stored in Hadoop or other sources. It runs containerized in the Kubernetes infrastructure and provides a bi-directional interface between data from SAP Data Hub data sources to Hadoop.

- **SAP Data Hub Pipelines** are the connectors between the various SAP Data Hub data sources and provide reusable, configurable operations to process data from the various sources. SAP Data Hub Pipelines run containerized in the Kubernetes infrastructure.

- **SAP Data Hub Vora Spark Extension** provides access to SAP Vora data sources from a Spark environment. It is installed on each Spark worker node in a Hadoop cluster.
- **Red Hat OpenShift Container Platform** is built around a core of application containers powered by Docker, with orchestration and management provided by Kubernetes, on a foundation of Red Hat Enterprise Linux. It provides many enterprise-ready features like enhanced security, multitenancy, simplified application deployment, and continuous integration/continuous deployment tools. Red Hat OpenShift gives administrators a single place to implement and enforce policies across multiple teams, with a unified console across all Red Hat OpenShift clusters.”

A **Red Hat OpenShift Container Platform** consists of several types of nodes:

- Two or more nodes running RHEL 7+
- Series of Ansible playbooks
- Several Kubernetes nodes
- Private Docker Registry

Red Hat Ceph Storage is a commercially supported distribution of Ceph, an enterprise grade, scale-out storage solution. Ceph is a distributed object store which provides excellent performance, scalability and reliability.

Lenovo Intelligent Insights with SAP Data Hub is deployed on **Lenovo ThinkSystem** platforms:

- **SAP Data Hub application and distributed runtime:**
 - Lenovo ThinkSystem SR530 or
 - Lenovo ThinkSystem SR630
- **Storage nodes:**
 - Lenovo ThinkSystem SR630 or
 - Lenovo ThinkSystem SR650
- **Lenovo ThinkSystem NE2572 Rackswitch** – Dedicated SAP Data Hub network
- **Lenovo RackSwitch G7052** – Management network

Lenovo Professional Services provide end to end lifecycle management as well as deployment, maintenance and managed services for SAP Data Hub, SAP HANA and big data applications.

Ensure Business Transformation Success, the Lenovo Solution for SAP Data Hub:

- Improves decision making whether data is in the cloud, on premise, in data lakes, or business systems
- Enables new business use cases
- Prepares you for future information management needs with an extensible and open system
- Scales to meet ever growing enterprise and big data requirements based a server infrastructure rated #1 in reliability¹

#1 in x86
Reliability
In the latest ITIC reliability survey Mid-year 2018, **Lenovo** again was the **#1 x86 vendor** for six years running*

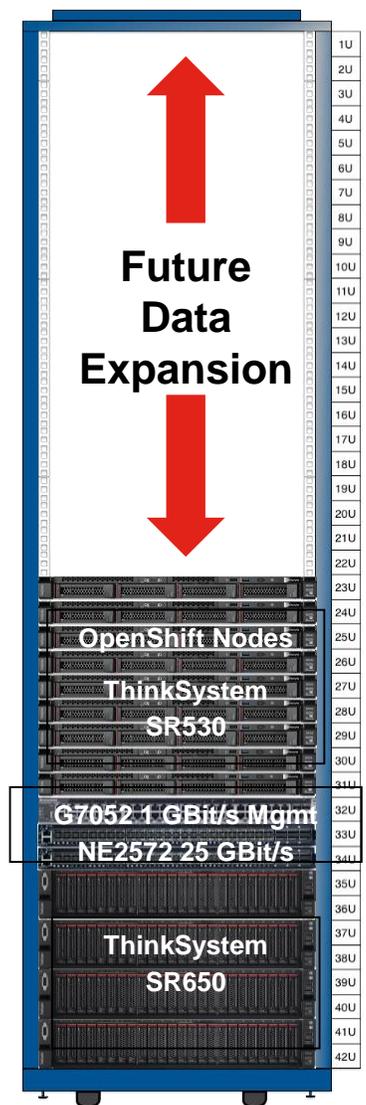


Figure 5: Example configuration for Lenovo Intelligent Insights with SAP Data Hub Rack Implementation

Drive Successful Digital Transformation

Lenovo

Why Lenovo

Lenovo is a leading provider of x86 servers for the data center. Featuring rack, tower, blade, dense and converged systems, the Lenovo server portfolio provides excellent performance, reliability and security for SAP applications. Lenovo has shipped more than 12,000 SAP HANA systems and also offers a full range of networking, storage, software, solutions, and comprehensive services supporting business needs throughout the IT lifecycle. With options for planning, deployment, and support, Lenovo offers expertise and services needed to deliver better service-level agreements and generate greater end-user satisfaction.

For More Information

To learn more about Lenovo Intelligent Insights for SAP Data Hub contact your Lenovo sales representative or Business Partner and visit: www.lenovo.com/sap



© 2019 Lenovo. All rights reserved.

Availability: Offers, prices, specifications and availability may change without notice. Lenovo is not responsible for photographic or typographical errors. **Warranty:** For a copy of applicable warranties, write to: Lenovo Warranty Information, 1009 Think Place, Morrisville, NC, 27560, Lenovo makes no representation or warranty regarding third party products or services. **Trademarks:** Lenovo, the Lenovo logo, System x, ThinkServer are trademarks or registered trademarks of Lenovo. Microsoft and Windows are registered trademarks of Microsoft Corporation. Intel, the Intel logo, Xeon and Xeon Inside are registered trademarks of Intel Corporation in the U.S. and other countries. Other company, product, and service names may be trademarks or service marks of others. SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries.

1 ITIC Mid-year 2018 Global Hardware, Server OS Reliability Report: <https://en.resources.lenovo.com/analyst-reports/itic-2018-global-reliability-survey-mid-year-update>

CRN: XXXXXXXXXXXX

04/2019