

Analyze High-Velocity, High-Volume Data from the Internet of Things

Solution Brief
September 2016

With Cisco Solutions for End-to-End IoT Analytics for SAP



Highlights

Build a Foundation

- Deploy Cisco UCS® Integrated Infrastructure for Big Data and Analytics with SAP solutions.

Tap Into High-Velocity Data

- Use data stored in repositories with data gathered in real time to enhance your data sets and extend analysis.

Make More Informed Decisions

- Quickly collect, process, and analyze massive amounts of data to get better insight in less time.

Deploy a Validated Solution

- Use a Cisco and SAP solution that consists of clusters of reference architectures that are each prevalidated to simplify deployment and accelerate data analysis.

Accelerate Response Times

- Rapidly provision high-performance solutions for faster analytics.

Massive amounts of data are coming your way in real time through the Internet of Things (IoT). We help you capture that data and turn it into business insight.

The Internet of Things, or IoT, is considered one of the most profound transitions in technology today, creating new opportunities for businesses. With the capability to harness the variety and volume of information stored in repositories with high-velocity data gathered in real time, the IoT is poised to disrupt the way your business and users interact with and analyze digital information. With Cisco® solutions, you can create an IoT foundation for SAP software that lets you quickly collect, process, and analyze massive amounts of data to get better insight in less time.

Energizing Analytics with the Internet of Things

The IoT links smart objects to the Internet, allowing new types of data exchange. With the capability to plug in devices and access information located anywhere—from handheld devices and sensors to cars and smart city infrastructure—the IoT promises to deliver a new level of data-driven intelligence and services. To be successful, your organization needs IT solutions and applications that can locate, gather, analyze, and use this information whether it is in storage or in motion.

What's Needed in an IoT Foundation

The right IoT foundation lets you connect, manage, and control previously unconnected devices and use the data collected from them in your decision-making processes. These devices can be almost anything that interacts with systems and users to gather and store information. Data injection mechanisms send the collected information over the network to a message broker that authenticates client devices and interacts with systems that subscribe to or publish information.

Continuous, real-time data processing is performed by a stream-processing framework. Unlike systems that assume the use of static data, the stream-processing component receives data from the message broker in real time. Pattern discovery and recognition software identify data anomalies and separate them from verified or estimated data. When digital data sources are used, such as speech, video, and audio, signal processing solutions filter the data and make it available to applications. Sampling techniques reduce the volume of data generated to accelerate performance and reduce the size of data sets.

Typically, data is pipelined, allowing the system to transform or convert the captured data to another format for processing. Data may be aggregated to reduce the likelihood that the system will process spurious information, and data may be enriched with additional data sets that can improve the quality of data analysis. Finally, the resulting data is stored either as structured or unstructured data that can be loaded into in-memory databases and processed by applications.

Cisco and SAP: Delivering a Better IoT Foundation

Co-innovation between Cisco and SAP resulted in a Cisco solution for end-to-end IoT analytics. Using Cisco Validated Designs for SAP HANA, SAP HANA Vora, and streaming solutions as building blocks, the solution provides a foundation for big data analytics from the data center to the edge of the network. The solution can be deployed on your premises or in a cloud

environment, depending on the needs of your users. Clusters run in secure tenants, with each cluster running a different software technology to help ensure isolation. Clusters can scale from one to thousands of nodes, and multiple clusters can be deployed in a tenant.

Built for Big Data

The solution supports the industry-standard Lambda Architecture, providing a robust, fault-tolerant platform for big data. It allows applications to use IoT data from clusters running SAP HANA and SAP HANA Vora through the use of queries, as well as data from Kafka alerts. Within the solution:

- All data that enters the system is rapidly analyzed to search for known patterns
- Data that matches a known pattern is sent to a speed layer for immediate use by IoT applications
- All data is sent to a batch layer so that data scientists can discover new, previously unknown patterns to use in future search operations
- A serving layer enables ad-hoc queries of data in both the speed layer and the batch layer

Solution Architecture

The solution consists of Cisco and SAP technologies that work together to help you integrate your IoT data with business, analysis, and operational processes for better business outcomes. Cisco UCS Integrated Infrastructure for Big Data and Analytics, SAP and third-party solutions, and software-defined, policy-driven networking come together in a highly scalable IoT foundation.

Cisco UCS Integrated Infrastructure for Big Data and Analytics

The foundation for the Cisco and SAP approach to end-to-end IoT analytics is Cisco UCS Integrated Infrastructure for Big Data and Analytics. This complete and easy-to-order package includes computing, storage, connectivity, and unified management capabilities. The solution can scale from 80 nodes in a single Cisco UCS management domain up to thousands of nodes by interconnecting multiple domains.

Hadoop Distributed File System

The Hadoop Distributed File System (HDFS) is a distributed, scalable, and portable file system that stores data on the nodes in a cluster. It uses compression and encryption techniques to help enhance data security and system performance.

Apache Spark

Apache Spark is an optimized, general-purpose engine for processing large amounts of data. The platform brings together many capabilities—including batch and real-time processing, analytics, and interactive exploration—so that you can enhance your Hadoop deployments and other big data repositories to explore data faster.

SAP HANA Vora

SAP HANA Vora is an in-memory query engine that lets you use the abundance of information stored in your Hadoop platform. By combining your business information with data from other sources—including streaming, interactive queries, and machine learning—you can accelerate and add context to your decision-making processes.

SAP HANA

SAP HANA is a highly optimized in-memory database platform that allows you to analyze structured data in real time. In addition to storing and retrieving data, SAP HANA can integrate data from multiple data sources with real-time transaction information so that you can get better insight in less time. Built-in application services make it easy for you to perform text, streaming, and predictive analytics as well as text searches and spatial data processing.

SAP HANA Smart Data Streaming

SAP HANA Smart Data Streaming captures data from large numbers of IoT devices so that SAP HANA can receive and act on information as soon as it is available. A lightweight version, SAP HANA Streaming Lite, runs on remote devices and IoT gateways to preprocess events prior to sending them to the data-streaming software.

SAP SQL Anywhere

In the solution, SAP SQL Anywhere solutions support the exchange, synchronization, and management of data captured from IoT devices and systems in real time. This enterprise-class database replicates transactions among remote devices and your data center and uses encryption to keep data secure as it is transmitted.

Apache Kafka

Using a publish-and-subscribe messaging approach, Apache Kafka allows massive amounts of data sent from hundreds or thousands of clients to be coordinated for processing. It divides data streams into chunks and

sends the data fragments to cluster nodes for processing in parallel.

Unified Management

Cisco Unified Computing System™ (Cisco UCS) management lets you define policies for creating server definitions, with assignment of specific roles that can use those policies to establish server instances. You can easily deploy and maintain all of your SAP environments by eliminating the complex, time-consuming, manual processes that can otherwise slow the process of deploying new environments and scaling existing ones.

Cisco UCS service profiles support rapid and consistent server configuration, and automation simplifies ongoing system maintenance activities such as firmware updates across the infrastructure as a single operation. Advanced monitoring capabilities raise

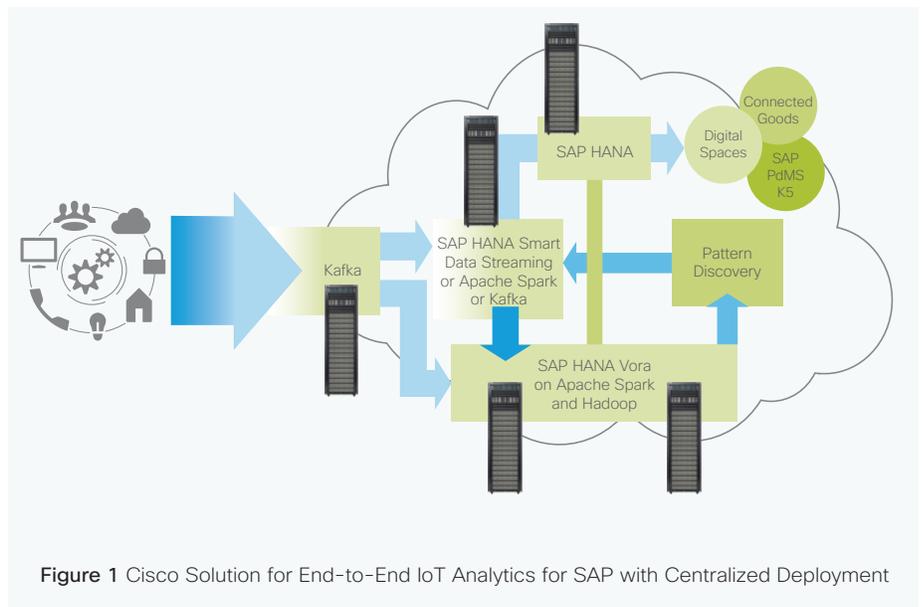
alarms and send notifications about the health of the entire infrastructure so that you can proactively address issues before they affect data analysis.

Deployment Options

You can deploy the solution in different ways, based on whether your data is stored in your data center or at the edge of your network.

Centralized Deployment

In a centralized deployment, the SAP software is deployed in your data center, and other third-party components are deployed at the edge of your network (Figure 1). Data generated by sensors and IoT data sources is taken into the system using Apache Kafka, routed to the streaming processing software, and sent in real time to the in-memory SAP HANA database, where it can be used by



Analyze High-Velocity, High-Volume Data from the Internet of Things With Cisco Solutions for End-to-End IoT Analytics for SAP

analytical applications. At the same time, the data stream is processed and sent to SAP HANA Vora and Hadoop for persistent storage. Data storage and processing can take an alternate route, starting with delivery to SAP HANA Vora and Apache Spark, followed by transmission to Hadoop and SAP HANA for access by analytical applications.

Distributed Deployment

In a distributed deployment, the SAP software is deployed in your data center and at the edge of your network (Figure 2). IoT data is sent to a local message broker or remote streaming facility for pattern recognition, machine learning, and deep-learning methodologies and tools. Resulting data is stored in a local SAP SQL Anywhere

database, and raw data is stored in the Hadoop platform at the edge of the network. A subset of data is sent to SAP HANA Vora, where it is processed. Aggregated data is made available through the in-memory SAP HANA database, and it is simultaneously stored in the Hadoop platform in the data center for persistence.

Conclusion

If you need to rapidly provision high-performance solutions for faster analytics, consider an IoT foundation based on Cisco and SAP solutions. With this innovative solution, you can capture the massive amounts of data that are coming your way in real time through the IoT and turn it into

Cisco Firsts for SAP Deployments

Cisco was the first company to deliver the following solutions certified by SAP:

- Policy-based Tailored Data Center Integration (TDI) for management, security, and provisioning
- SAP HANA Vora and big data reference architecture
- One-click deployment of SAP HANA applications
- Software-defined storage for SAP HANA
- 40-Gigabit Ethernet for SAP HANA
- Cisco Solution for end-to-end IoT analytics for SAP software with distributed deployment

business insight. Many organizations have achieved 528 percent return on investment (ROI) by running their SAP solutions on Cisco infrastructure.¹ Let us help you do the same.

For More Information

For more information about Cisco and SAP solutions for the IoT, visit <http://www.cisco.com/go/sap>.

For more information about Cisco Validated Designs for SAP solutions, visit http://www.cisco.com/c/en/us/solutions/enterprise/data-center-designs-application-networking/landing_sap.html.

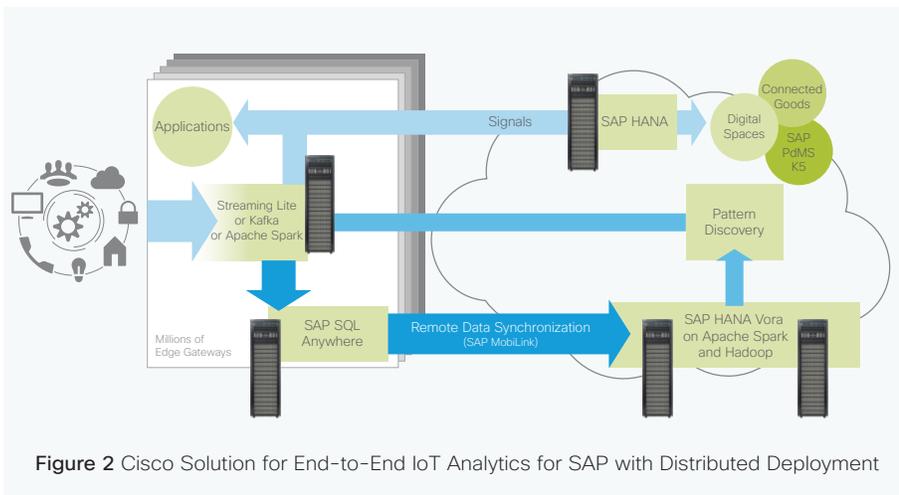


Figure 2 Cisco Solution for End-to-End IoT Analytics for SAP with Distributed Deployment



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

¹ IDC: The Business Value of Cisco UCS Integrated Infrastructure Solutions for Running SAP Workloads, March 2016, US41084916.