Cisco and SAS Edge-to-Enterprise IoT Analytics Platform

Comprehensive analytics platform for the Internet of Things with Cisco Unified Computing System and SAS analytics.

The Internet of Things (IoT) is leading one of the most profound global market transitions in history. An enormous number of things are being connected to the Internet, including devices, people, and entire business processes. But they are not just connected to the Internet. They are also connected to everything else. This "connection of the unconnected" transforms the Internet of Things into the Internet of Everything (IoE), and it creates a vast range of new opportunities for businesses around the world.

The IoE will transform nearly every industry, locally and globally. It brings networking technology to places where it was previously unavailable or impractical. The true power of smart, connected devices and the data and insights they generate will create the next business era.

According to Cisco[1], 50 billion devices will be connected to the Internet by 2020, and 500 billion devices are expected to be connected to the Internet by 2030. Each device includes sensors that collect data, interact with the environment, and communicate over a network. The IoT is the network of these connected devices. As these connections multiply, the result is exponential change and a digital disruption that creates new revenue streams from better customer and social experiences.

These smart, connected devices generate data that analytic applications need to collect, aggregate, and analyze to deliver informed, actionable insights. The challenge is to build the right digital infrastructure and enable the right set of applications to harness this data.

[1] https://www.cisco.com/c/dam/en/us/products/collateral/se/internet-of-things/at-a-glance-c45-731471.pdf

In fact, Gartner[2] estimates that by 2020, 75 percent of businesses will be digital or will have digital business transformations underway. However, only 30 percent of those efforts will be successful, in part because of a lack of personnel with specialized talent and technical expertise. All these changes require a fundamental rethinking of business and operational strategies as organizations adopt new models for agile information technology, edge analytics, and platform-based security.

In the connected world of the IoT, petabyte-scale data is generated in real time from cars to factories to farms and other entities, and the capability to capture, monitor, and rapidly process information is essential for the modern enterprise. Cisco and SAS have partnered to create the Cisco® and SAS Edge-to-Enterprise IoT Analytics Platform, which allows enterprises to quickly collect, process, and analyze massive amounts of data in real time, both at the network edge and in the enterprise data center.

The goal of the IoT is valuable outcomes. The only way to achieve this value is through the timely application of analytics. This document describes the Edge-to-Enterprise IoT Analytics Platform from Cisco and SAS using use cases in the renewable-energy sector. Although the use cases are specific, the architecture has broad applicability across many industries and streaming analytics scenarios.

Use Cases

The Cisco and SAS Edge-to-Enterprise IoT Analytics Platform enables a cohesive guidance layer to align the expertise of partners—providers of sensors, applications, data analysis, security, services oversight, and so on—and to get these partners and their assets all operating effectively together to gather, analyze, and deliver data and a shared experience. This experience promotes optimal, timely decisions that support an abundance of new opportunities for revenue generation and cost control.

This architecture supports multiple use cases such as the following:

- Energy and utilities
- Connected automobiles
- Smart manufacturing

Energy and utilities

The pace of change in the energy industry is accelerating faster than ever before. The entire landscape is shifting: where energy comes from, who creates it, and how it gets where it needs to go. Two areas receiving a lot of attention today are predictive asset management and the smart grid.

Predictive asset management

Predictive asset management improves the overall availability and operational performance of assets while reducing their total cost of ownership (TCO). Edge-based analytics enable organizations to move from scheduled maintenance models to predictive ones, giving them dynamic control over their equipment and maintenance resources.

These capabilities help organizations avoid catastrophic failures by identifying critical issues that occur inside the normal scheduled maintenance window. They also lower costs by eliminating unnecessary maintenance tasks and make more effective use of scarce and expensive resources (both human and capital). This strategy is essential for large industrial organizations.

Smart grid and smart meter management

The smart grid, regarded as the next-generation power grid, uses a two-way flow of electricity and information to create a widely distributed and automated energy delivery network.

The smart grid provides advanced information about electricity generation, communications, and infrastructure by capturing sensor information. This information helps organizations understand delivery patterns and consumption patterns by analyzing the data collected by the system.

The smart metering system provides advanced information metering, monitoring, and management. Smart meters connected to sensors provide data about consumption patterns to help manage electricity in an optimal way.

[2] https://www.cisco.com/c/dam/en_us/about/business-insights/docs/digital-business-transformation.pdf

Connected automobiles

The automotive industry is entering a period of profound change. Autonomous vehicles are close to reality, using advanced machine learning powered by IoT data. Vehicle-to-vehicle and vehicle-to-infrastructure communications technologies are in development, taking advantage of unified, secure network connectivity. These technologies will enable transmission of real-time vehicle telematics, GPS tracking, and geo-fencing. This data can be used to support improved safety, mobility, and efficiency and to pave the way for new business models such as pay-as-you-go auto insurance. Proactive maintenance decreases costs and reduces vehicle downtime.

Smart manufacturing

Smart manufacturing solutions provide intelligent, timely information and collaboration in context to transform manufacturing businesses through continuous innovation, a differentiated customer experience, supply-chain agility, and operational excellence. Smart manufacturing enables organizations to reduce downtime, increase productivity, and maintain industry compliance, while optimizing processes and improving overall equipment effectiveness.

The Role of Big Data

Every enterprise is becoming a data-driven business. Data is the lifeline that guides intelligent decision making. Any discussion of IoT is incomplete without considering big data because IoT devices generate significant amounts of data. Connected devices, sensors, and algorithms all operate in ways that involve massive amounts of data. This data must be transferred, stored, analyzed, and managed.

This rise of big data has led to the modernization of data infrastructure, with Apache Hadoop as a critical foundational element for data storage and processing. An edge-to-enterprise analytics system uses big data technology as a primary component. Beyond its obvious role in storing collected data for historical analysis and model development, big data technology is used to manage the transfer of data in a highly available way. It also enables analytics processes for streaming data and extract, transform, and load (ETL) processing for traditional business intelligence and data warehouse systems.

Emergence of Analytics at the Network Edge

In the IoT, objects or sensors with embedded computing devices connect to the Internet to send and receive data. This behavior represents a significant architectural challenge because these devices generate enormous amounts of data, currently on the order of exabytes per day. In addition, these new IoT-enabled devices produce many different types of data. This data is often very noisy, is produced continuously, and uses a variety of protocols, and it all needs real-time analysis and response.

Traditional computing models send the data to the enterprise data center for analysis. However, this approach is impractical in many scenarios because of the volume of data being produced and the need for real-time analysis and response times, often measured in milliseconds.

Simply collecting data from connected sensors, systems, or products is not enough. To benefit from the promise of IoT data, businesses need to be able to expand the way that analytics processes are run: from traditional enterprise data centers out to devices on the edge. The challenges arise from the complexity—and risks—inherent in capturing and analyzing the huge volumes and varieties of data flowing from the ever-increasing number of things.

As a result, a new model for analyzing IoT data at the edge of the network has emerged. This model moves the analysis and response close to the devices that generate the data, reducing latency and also reducing the load on the network and the enterprise data center.

This document describes an architecture for analyzing IoT data, including real-time analysis and response at the edge of the network as well as historical analysis, operational control, and model development in the enterprise data center. This architecture takes into account the often harsh environments that exist outside the computing center.

Cisco and SAS Edge-to-Enterprise IoT Analytics Platform

Together, the use of big data analytics in the enterprise data center and the application of targeted analytics at the network edge enhance the value of IoT analytics. The IoT analytics architecture provides an opportunity to think differently about where and when analytics processes are run. It allows organizations to challenge the barriers of latency, data volume, and connectivity and the costs associated with them.

The Edge-to-Enterprise IoT Analytics Platform supports analytics across the entire architecture, including the edge, the network, and the enterprise data center.

Cisco and SAS Support Analytics From the Edge to the Enterprise

The edge includes physical devices and sensors that interact with the world and collect data. They are the data producers. They can be very small in terms of their computing power, perhaps only able to collect data and transmit it over an industrial protocol. They can also be very powerful, with built-in computing and the capability to route data directly over the network.

At the edge of the network, Cisco 829 Industrial Integrated Services Routers act as a bridge between the IoT devices and the enterprise data center. The edge network acts as a collection and routing point for the data produced by large numbers of IoT devices. SAS Event Stream Processing expands the edge network from a routing point to a fully fledged analytics processing center.

Event Stream Processing applies sophisticated analytics to streaming data in real time. By performing analysis at the edge, the system can respond very quickly: within milliseconds. By applying filters, the system dramatically reduces the volume of data traversing the network, reducing the network load and the data ingest and analysis in the enterprise data center.

Note that the conditions at the edge of the network are harsh relative to the controlled environment of the enterprise data center. Equipment at the edge needs to be specifically designed to survive these conditions. This requirement has an impact on the choices made regarding what services to put where. For example, the cost of using hardened industrial routers capable of running independent applications may be far less than the cost of maintaining a remote facility for less rugged equipment.

The transfer layer includes real-time data pipelines, such as Apache Kafka, that provide a reliable, fault-tolerant, linearly scalable message bus for data in transit from the edge to the enterprise. The data is retained until it is processed and stored in the enterprise data center.

The enterprise data center fulfills its traditional role of housing the data and providing the tools to analyze and use the data. The big data is stored using Hadoop, and the analysis needs are met using an integrated suite of tools from SAS: SAS Event Stream Processing Server, LASR Analytics Server, Visual Analytics, and Visual Statistics.

Figure 1 shows SAS Visual Statistics displaying a linear regression model that estimates the phase angle at a point on the power grid. An analyst can take data from substations that have sensors and use it to estimate values in other locations.
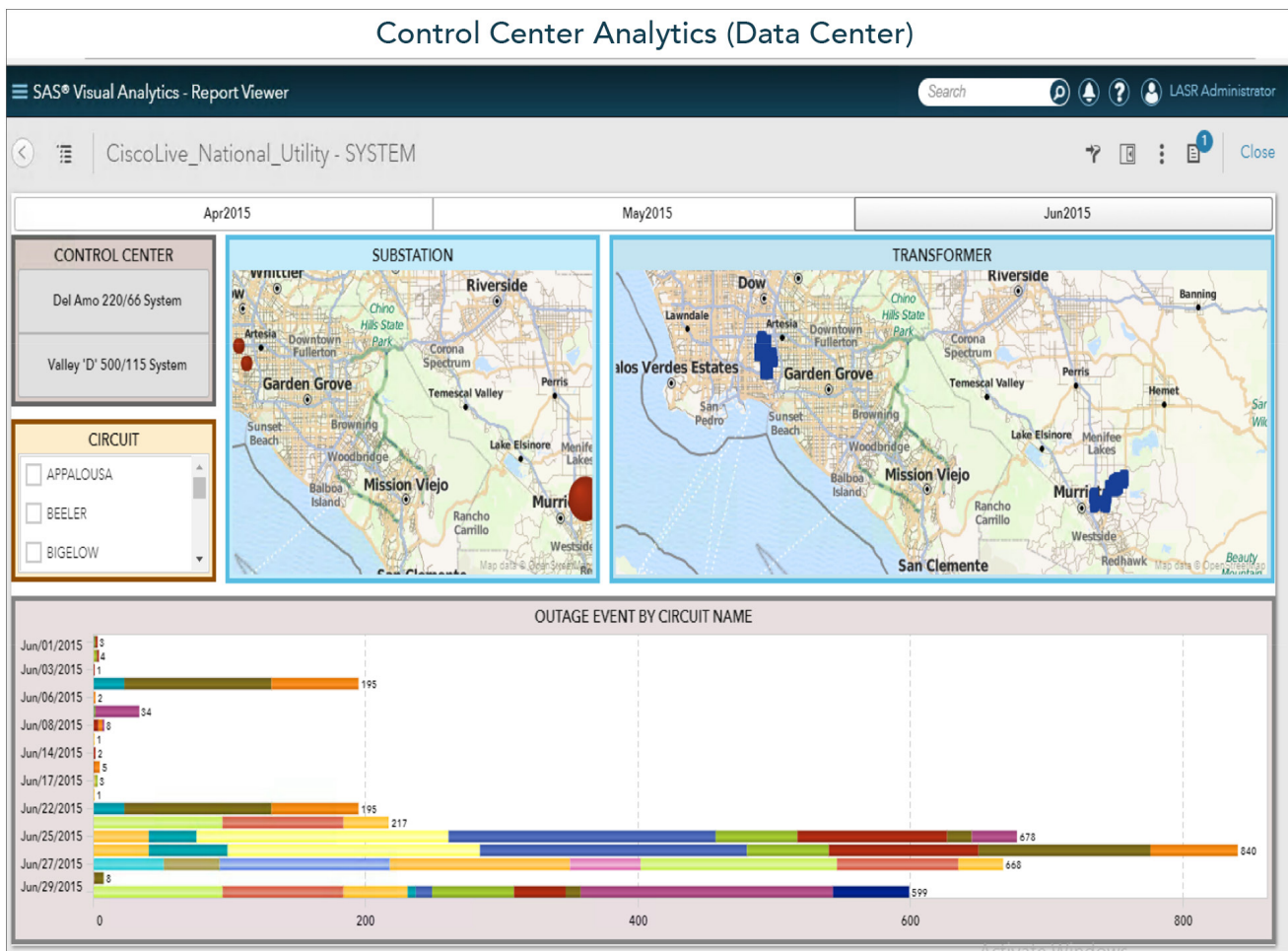
Figure 1.   SAS Visual Statistics Dashboard

An important function in the enterprise is the development of models to be used by the edge systems to analyze the data at the edge. The capability to learn from the deep analysis in the enterprise and then roll out improved analytical models at the edge provides a continuous feedback loop. Support for a transparent workflow between model development and model deployment is an important differentiator of edge-to-enterprise analytics systems.

SAS Event Stream Processing Server, installed in the enterprise, enables organizations to develop pattern-detection capabilities across multiple streams of data coming from the edge. Organizations also can develop real-time control-center dashboards based on the data streaming from various edge locations. For example, Figure 2 shows the number of signals emerging from the various substations in a utilities scenario.
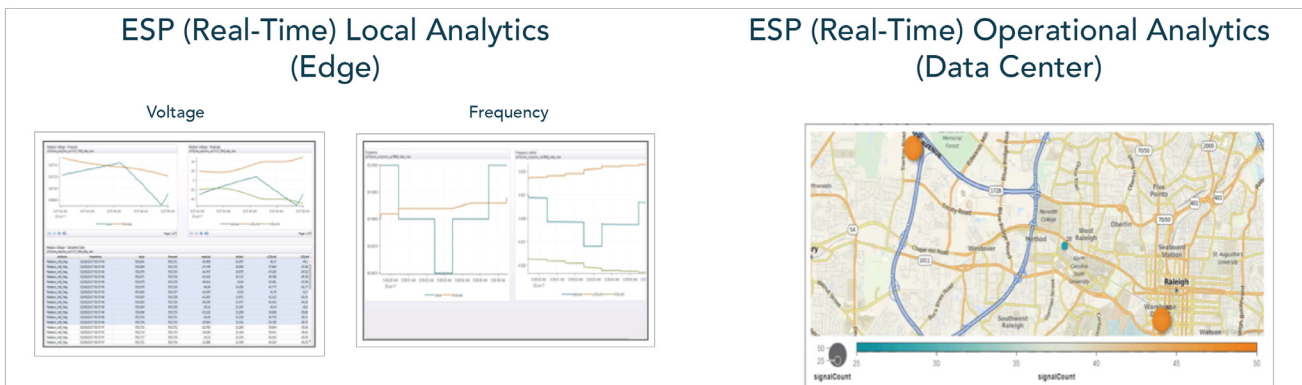
Figure 2. SAS Event Stream Processing Real-Time Control Center Dashboard

As shown in Figure 3, the Edge-to-Enterprise IoT Analytics Platform uses Cisco infrastructure and software technologies to connect to IoT devices at the edge, collect and route data traffic, and run SAS software at both the edge and in the enterprise. The SAS software at the edge filters data, applies models, and issues alerts. Filtered (or relevant) IoT data is routed to the enterprise data center by Apache Kafka, a secure, highly available, distributed, message broker capable of handling the large volumes of data common in IoT analytics systems.

In the enterprise, data is stored using Hadoop and analyzed both historically and in real time using SAS software. Historical analysis, real-time operational control, and model development are all performed using SAS's integrated suite of tools: Event Streaming Processing Server, LASR Analytics Server, Visual Analytics, and Visual Statistics.
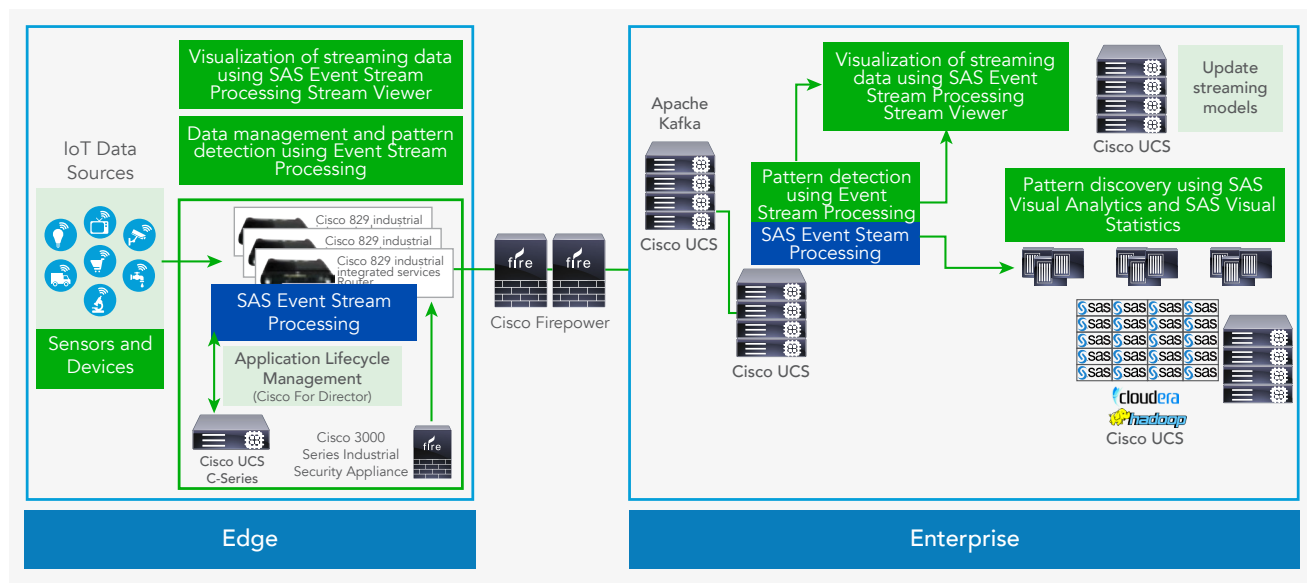

Figure 3. Cisco and SAS Edge-to-Enterprise IoT Analytics Platform

Edge Components

The reference architecture includes the edge components discussed in this section.

Cisco Unified Computing System

At the edge, Cisco Unified Computing System™ (Cisco UCS®) rack-mount servers or Cisco UCS Mini is recommended.

- Cisco UCS C-Series Rack Servers deliver unified computing in an industry-standard form factor to reduce TCO and increase agility. Each server addresses varying workload challenges through a balance of processing, memory, I/O, and internal storage resources.
- Cisco UCS Mini is optimized for branch and remote offices, point-of-sale locations, and smaller IT environments. It is an excellent solution for customers who need fewer servers but still want the comprehensive management capabilities provided by Cisco UCS Manager. Cisco UCS Mini delivers servers, storage, and 10 Gigabit Ethernet networking in an easy-to-deploy, compact form factor.

The reference architecture presented in this document uses the Cisco UCS C220 M5 Rack Server, a 1-rack-unit (1RU) server. This server runs Cisco Fog Director and also serves as an edge network server to address additional reporting and analysis requirements.

Cisco 829 Industrial Integrated Services Router

The Cisco 829 Industrial Integrated Services Router is a ruggedized integrated services router designed for deployment in harsh industrial environments. The router has a compact form factor, an integrated 9V to 32V DC power input, and multimode third-generation (3G) and fourth-generation (4G) Long-Term Evolution (LTE) wireless WAN and WLAN connections.

Cisco IOx software, an open, extensible environment for hosting applications at the network edge, and SAS Event Stream Processing software are installed on the Cisco 829 Industrial Integrated Services Router.

Cisco Fog Director

Cisco Fog Director provides the capability to manage large-scale production deployments of edge devices and applications on the Cisco 829 Industrial Integrated Services Routers.

Cisco Fog Director is used to deploy and manage the Event Stream Processing application on the Cisco 829 routers at the edge. It is also critical to managing the deployment of new models run by the SAS software.

Cisco 3000 Series Industrial Security Appliance

The Cisco 3000 Series Industrial Security Appliance provides the industry's widest range of operation technology-specific access control, threat detection, and application visibility for the most demanding environments. It works in harsh environments (with a temperature range of –40 to 60°C) and is hardened for vibration, shock, surge, and electrical noise immunity. It offers four high-performance Ethernet data links in a DIN rail or rack-mount form factor.

Cisco Firepower firewall

The Cisco Firepower® next-generation firewall is a fully integrated, threat-focused next-generation firewall with unified management. It uniquely provides advanced threat protection before, during, and after attacks while providing better security, faster speed, and a smaller footprint.

SAS Event Stream Processing Client

SAS Event Stream Processing Client analyzes and acts on events in real time as they occur. It is deployed on the Cisco 829 routers and provides real-time, low-latency, high-throughput event processing. It supports a variety of functions, including machine-learning algorithms (such as neutral networks, gradient boosting, and decision trees), text analysis to categorize text and extract entities, sentiment analysis, advanced pattern matching, and in-stream join operations.

Enterprise Components

The reference architecture includes the enterprise components described in this section running on Cisco UCS infrastructure.

Cisco Unified Computing System

Cisco UCS C-Series Rack Servers and S-Series Storage Servers are industry-leading 2-socket servers designed for both performance and expandability over a wide range of computing and storage-intensive infrastructure workloads.

Apache Kafka

Apache Kafka is an open-source, distributed messaging system that provides fast, highly scalable, and durable messaging through a publish-subscribe model. SAS Event Stream Processing uses Kafka to handle data pipelines for high-speed filtering and pattern matching. Kafka is co-located in the enterprise data center core and deployed on Cisco UCS C240 M5 Rack Servers.

SAS Event Stream Processing Server

SAS Event Stream Processing Server provides real-time event-stream processing in the enterprise data center, capturing data arriving continuously from edge devices, analyzing it, and acting on new information as it arrives. It issues alerts and notifications and streams live information to operations dashboards.

Event Stream Processing includes a visual model-development environment and a visualization component for building dashboards using streaming data.

SAS LASR Analytics Server

SAS LASR Analytic Server is an analytics platform, applying analytics to big data. The server provides speedy, secure, multiuser access to in-memory data in a distributed computing environment.

Apache Hadoop

Apache Hadoop is an open-source initiative focused on the big data challenge and a leading solution for enterprise big data implementations. Hadoop provides the software and system needed to store, process, and analyze huge amounts of both structured and unstructured data. Hadoop is designed for high availability and linear scaling.

SAS Visual Analytics

SAS Visual Analytics enables organizations to gain insight from all the data, no matter the amount of data, with no need to sample or create subsets of the data. It is implemented as an integrated suite of web applications that offer intuitive drag-and-drop interactions, rapid and highly visual responses, and role-based access to functions.

SAS Visual Statistics

SAS Visual Statistics enables organizations to derive predicted values from the predictive models. These new variables contain the prediction information for the models and can be used in other visualizations. Deployed on Cisco UCS C240 M5 servers, Visual Statistics is fully integrated into SAS Visual Analytics.

SAS Analytics for IoT Bundle

The SAS Event Stream Processing, Visual Analytics, and Visual Statistics components are encapsulated in a bundle, SAS Analytics for IoT, to provide an industry-independent foundational platform for IoT analytics.

Reference Architecture

The reference architecture includes the components shown in Figure 4 and listed in Table 1.
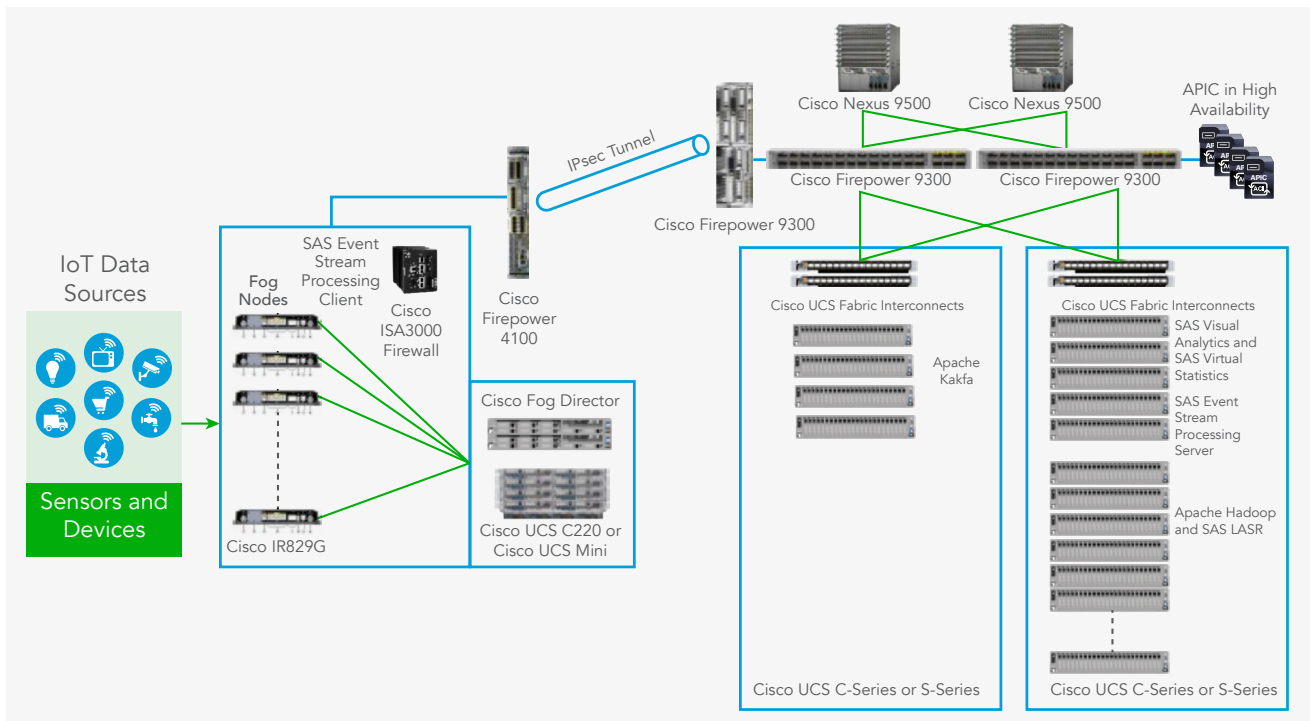
Figure 4. Reference Architecture for Cisco and SAS Edge-to-Enterprise IoT Analytics Platform

Location	Component	Configuration
Edge	• Cisco 829 Industrial Integrated Services Router	• 2 cores (1.25-GHz dual-core CPU, 2-GB DDR3, 8-GB embedded multimedia card [eMMC] storage, and 4 x 10/100/1000 Mbps) • Cisco IOx Release 1.3.0 • Guest OS: Ubuntu Release 14.04.01/03 • Cisco IOS® Software Release 15.6.3M2 • SAS Event Stream Processing Edge Release 4.3.0
	SAS Event Stream Processing Edge	Container application running on Cisco 829 Industrial Integrated Services Router
	Firewall	• Inbound (to Cisco 829 Industrial Integrated Services Router): Cisco 3000 Series Industrial Security Appliances • Outbound (from and Cisco Fog Director): Cisco Firepower 4100 Series

Location	Component	Configuration
	Cisco Fog Director	Cisco Fog Director can be run at the edge or enterprise data center on Cisco UCS Mini or Cisco UCS C220 M5 Rack Servers with: • 2 Intel Xeon Processor Scalable Family 6132 CPUs (28 cores) at 2.6 GHz • 12 x 32 GB 2666 MHz (384 GB) • Cisco 12-Gbps SAS Modular RAID Controller with 2-GB flash-based write cache (FBWC) • 8 x 1.8-TB 10,000-rpm hard-disk drives (HDDs) • Cisco UCS Virtual Interface Card (VIC) 1387 (with 2 x 40 Gigabit Ethernet) **Note:** Cisco Fog Director Release 1.3.0
Data pipeline (transfer)	Apache Kafka nodes	Four Cisco UCS C240 M5 Rack Servers each with: • 2 Intel Xeon Processor Scalable Family 6132 CPUs (28 cores) at 2.6 GHz • 12 x 16 GB 2666 MHz (192 GB) • Cisco 12-Gbps SAS Modular RAID Controller with 4-GB FBWC • 26 x 1.8-TB 10,000-rpm HDDs • 2 x 480-GB 6-Gbps 2.5-inch Enterprise Value SATA solid-state disks (SSDs) for boot • Cisco UCS VIC 1387 (with 2 x 40 Gigabit Ethernet Quad Small Form-Factor Pluggable [QSFP] ports) **Alternative reference architecture with Cisco UCS S-Series Storage Servers:** • Cisco UCS S3260 Storage Servers each with two nodes, and each node with: • 2 Intel Xeon Processor E5-2680 v4 CPUs (28 cores) at 2.4 GHz • 8 x 32 GB 2400 MHz (256 GB) • Cisco 12-Gbps SAS Modular RAID Controller with 4-GB FBWC • 24 x 6-TB 7,200-rpm HDDs • 2 x 480-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS system I/O controller (SIOC; with 2 x 40 Gigabit Ethernet QSFP ports)

Location	Component	Configuration
	Firewall	Inbound to Apache Kafka: Cisco Firepower 9300 Series
Enterprise data center core	Connectivity	Cisco Application Centric Infrastructure (Cisco ACI™) spine-and-leaf architecture consisting of: • 2 Cisco Nexus 9508 Switches (spine switches) • 2 Cisco Nexus 9332PQ Switches (leaf switches) • 1 Cisco Nexus 9372 platform leaf switch • 3 Cisco Application Policy Infrastructure Controller (APIC) M2 appliances • 4 Cisco UCS 6332 Fabric Interconnects
	Apache Hadoop	• Management nodes: 3 nodes • Data nodes: 16 nodes All nodes are Cisco UCS C240 M5 servers each with: • 2 Intel Xeon Processor Scalable Family 6132 CPUs (28 cores) at 2.6 GHz • 12 x 32 GB 2666 MHz (384 GB) • Cisco 12-Gbps Modular SAS host bus adapter (HBA) • 8 x 1.6-TB SSDs • 2 x 480-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS VIC 1387 (with 2 x 40 Gigabit Ethernet QSFP ports) **Alternative reference architecture with Cisco S-Series Storage Servers:** • Cisco UCS S3260 Storage Servers each with two nodes, and each node with: • 2 Intel Xeon Processor E5-2690 v4 CPUs (28 cores) at 2.6 GHz • 512 GB of memory • Cisco 12-Gbps SAS Modular RAID Controller with 4-GB FBWC • 8 x 1.6-TB SSDs • 12 x 6-TB 7,200-rpm HDDs • 2 x 480-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS SIOC (with 2 x 40 Gigabit Ethernet QSFP ports)

Location	Component	Configuration
	SAS LASR	Co-located with all data nodes **Note:** SAS LASR Release 9.4
	SAS Visual Analytics and SAS Visual Statistics	3 Cisco UCS C240 M5 Rack Servers each with: • 2 Intel Xeon Processor Scalable Family 6132 CPUs (28 cores) at 2.6 GHz • 12 x 32 GB 2666 MHz (384 GB) • Cisco 12-Gbps Modular SAS HBA • 8 x 1.6-TB SSDs • 2 x 480-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS VIC 1387 (with 2 x 40 Gigabit Ethernet QSFP ports) **Note:** SAS Visual Analytics and Visual Statistics Release 8.1
	SAS Event Stream Processing Server	2 Cisco UCS C240 M5 Rack Servers each with: • 2 Intel Xeon Processor Scalable Family 6132 CPUs (28 cores) at 2.6 GHz • 12 x 32 GB 2666 MHz (384 GB) • Cisco 12-Gbps Modular SAS HBA • 8 x 1.6-TB SSDs • 2 x 480-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS VIC 1387 (with 2 x 40 Gigabit Ethernet QSFP ports) **Note:** SAS Event Stream Processing Server Release 4.3

Table 1.   Reference architecture configuration details

Conclusion

The comprehensive Cisco and SAS Edge-to-Enterprise IoT Analytics Platform delivers industry-leading performance, scalability, and cost of ownership. This platform enables organizations to develop and deploy analytics easily with a simplified Cisco UCS and SAS deployment model to support powerful IoT and analytics processing.

For more information

- Cisco and SAS Edge-to-Enterprise IoT Analytics Platform Cisco Validated Design: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/Cisco_SAS_Edge_to_Enterprise_IoT_Analytics_Platform.html
- Cisco and SAS Validated Design for Visual Analytics Platform: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/Cisco_SAS_Visual_Analytics_Platform.html
- Cisco Big Data portal: https://www.cisco.com/go/bigdata
- Cisco UCS Integrated Infrastructure for Big Data: https://blogs.cisco.com/datacenter/cpav4/
- Cisco 829 Industrial Integrated Services Router: https://www.cisco.com/c/en/us/products/collateral/routers/829-industrial-router/datasheet-c78-734981.html
- Cisco 3000 Series Industrial Security Appliances: https://www.cisco.com/c/en/us/products/security/industrial-security-appliance-isa/index.html
- Cisco Firepower 4100 Series: https://www.cisco.com/c/en/us/products/security/firepower-4100-series/index.html
- Cisco Firepower 9300 Series: https://www.cisco.com/c/en/us/products/security/firepower-9000-series/index.html
- SAS and IoT: http://www.sas.com/en_us/insights/big-data/internet-of-things.html
- SAS Visual Statistics: http://support.sas.com/documentation/cdl/en/vaug/68648/PDF/default/vaug.pdf
- SAS Visual Analytics: http://support.sas.com/documentation/onlinedoc/vs/7.1/vsug.pdf
- SAS Event Stream Processing: http://www.sas.com/content/dam/SAS/en_us/doc/factsheet/sasevent-stream-processing-106151.pdf