SAS takes Event Stream Processing to Cisco's edge for Internet of Things stack

JASON STAMPER

17 MAR 2017

The business intelligence firm has signed a partnership agreement with Cisco, whereby SAS's Event Stream Processing technology runs at the edge of the network on Cisco gear.

451 Research®

Privately held software firm SAS Institute is able to deploy its Event Stream Processing (ESP) technology at the edge of the network on Cisco equipment, thanks to a new technology partnership arrangement.

THE 451 TAKE

No single vendor has what we might call a 'complete' Internet of Things (IoT) technology stack, and SAS and Cisco are no exception. However, by working together, they have shown how it is possible to do analytics from the edge of the network right through to the datacenter and various types of cloud, too. It may be a reference architecture and a coming together of what were existing products, but for companies looking to do IoT-type analytics, it still makes a lot of sense. It also highlights the importance of edge computing for its role in IoT, as well as the leading role that stream processing technologies are likely to play in the era of the Internet of Things.

CONTEXT

SAS is the largest privately held software vendor, specializing in business intelligence offerings and statistical software in particular. The Statistical Analysis System (SAS) was the first suite developed by SAS Institute for advanced analytics, originally between 1966 and 1976 at North Carolina State University, when SAS was then incorporated.

The two project leaders were Anthony Barr, who worked on variance and regression software for IBM mainframes, and student Jim Goodnight, who developed the software's statistical routines. Goodnight was later joined by John Sall.

Goodnight, 74, is CEO. In 2016, SAS posted revenue of $3.2bn. It employs over 14,000 staff and its software is installed at 83,000 government, business and university sites. It counts 94 of the top 100 companies on the 2016 Fortune Global 500 as customers.

TECHNOLOGY

One technology in SAS's vast portfolio of business intelligence and data management products is its ESP technology. Stream processing is gaining in popularity as more and more companies find themselves trying to analyze data in motion – stream processing is all about ingesting, filtering and analyzing data in flight.

For that reason, there is a requirement to be able to ingest tens of thousands to millions of events per second, and filter and analyze them in – or at the very least near to – real time. SAS's ESP is able to ingest and process millions of events per second. The stream processing engine runs in memory, which gives it the necessary speed.

Like others in the streaming space, SAS told us in a recent briefing that it is seeing tremendous growth in its stream processing business. It saw an 81% increase in ESP revenue between December 2016 and March of this year. It has 87 ESP customers all told, of which 30 were won in the last two years. However, the company did say that 80% of its ESP customers were already customers of another SAS product.

Among the use cases that SAS sees driving ESP business is IoT. Smart devices, sensors and other machine data is of most value as soon as it is created – there is often little value in analyzing it at all after weeks, days or even in some cases, hours.

To that end, in March SAS announced a technology partnership with Cisco, which has for many years evangelized its take on IoT, which it calls the Internet of Everything (IoE). The companies are calling their combined capability edge-to-enterprise analytics, and it is specifically targeted at IoT use cases. The reference architecture combines Cisco's networking, edge and datacenter infrastructure with SAS's capabilities in streaming and advanced analytics.

SAS said that depending on data volume, data velocity, latency and reporting requirements, multi-phase analytics enables it to detect unusual events as they occur at edge devices – Cisco gear – while maintaining aggregated views across a collection of devices and systems.

The companies are calling the reference architecture the SAS, Cisco IoT Analytics Platform. It combines Cisco's networking and datacenter infrastructure technology with SAS's streaming analytics and advanced analytics. The companies said there are three main elements in the SAS, Cisco IoT Analytics Platform. At the edge, Cisco's 829 Industrial Integrated Services Routers can now run SAS ESP. The 829 series of routers are designed for deployment in harsh

conditions. They are designed to withstand shock, vibration, dust, humidity and water, and they can tolerate a wide temperature range, too. Also at the edge, Cisco Fog Director software on servers is said to simplify the deployment of applications and models on edge routers.

Second, SAS and Cisco use the open source publish and subscribe mechanism, Apache Kafka, to transfer data from edge devices to the enterprise or the cloud. Today, ESP is certified to work in Amazon Web Services' cloud, but it can also be deployed in containerized/private cloud environments including Cloud Foundry BOSH and Chef. The company is considering certifying it to run in other clouds depending on customer demand.

Last – in the reference architecture – the companies envisage Apache Hadoop handling data storage with Cisco UCS Integrated Infrastructure for big data and analytics. It's possible to use SAS LASR Analytics Server, SAS Visual Analytics and SAS Visual Statistics in conjunction with Hadoop. Cisco UCS 6300 Fabric interconnects can be used for additional network management, connectivity, management and monitoring.

To validate their design, Cisco and SAS said they used sensor data from a smart grid containing millions of events. The data was ingested by SAS ESP deployed on Cisco 829 routers. It was then streamed to an ESP instance in a datacenter, and additionally then into Apache Hadoop for further analysis and for models to be built.

As well as potentially being used for smart grids (including smart factories and smart cities), the companies said use cases for the platform include smart electric grids or homes, connected cars, asset performance and predictive maintenance.

While models can be built to do the likes of machine learning in the datacenter, SAS's ESP does not currently support the statistical language R. The company said it might add this later if customers want it. Instead, models can be written in C++ or Python today. The company does, however, have a lightweight, HTML-based dashboard technology that enables streams to be viewed, customized, saved or shared: that product is called Stream Viewer.

Because SAS expects many companies that are attempting IoT-style projects to use ESP with its Visual Analytics and Visual Statistics technologies, it has also bundled the three into a suite called SAS Analytics for IoT.

COMPETITION

Competition would include the Predix platform from GE, although SAS has a connector that can link ESP to Predix because it feels it is in fact complementary. However, we have also seen some work on edge computing from Hadoop technology provider MapR, which also has a streaming engine called MapR Streams.

Another company that focused on edge analytics was Predixion Software, which was recently acquired by Greenwave Systems. What was Predixion's RIOT product has since been renamed AXON Predict, and made part of Greenwave's broader AXON IoT and mobile-to-mobile platform.

SAP has also done some work in the area of edge computing, as it is able to bring together its HANA In-memory database (which can run with a very small footprint if desired) and its SAP HANA Smart Data Streaming technology (which it acquired by buying Sybase, which itself had previously bought complex event processing company Aleri).

Software AG, similarly, has reference architectures for IoT use cases that typically involve its Apama Streaming Analytics technology, and sometimes also its Terracotta In-Memory data grid.

In the cloud – although not necessarily edge analytics technologies as such – customers could consider Amazon Web Services' Kinesis stream-processing service. If companies are focused on time series data in the cloud – such as that often generated by IoT use cases – they could look to technologies from startups Grove Streams or jKool.

Back on-premises, streaming technology competitors include Informatica, with Ultra Messaging Streaming Edition. There is also event-processing technology from TIBCO, which bought StreamBase Systems.

Smaller competitors include DataTorrent, which has a stream processing application that sits on a Hadoop cluster and can be used to analyze the data as it streams in, and SQL-based event-processing specialist SQLstream. Another young company, Striim (formerly known as WebAction) offers what it describes as streaming integration and intelligence, which could appeal for some use cases.

Cognitect says its Datomic distributed database has been used in edge-type situations, and the same goes for Edge Intelligence (formerly JustOne Database), which has just rebranded and redoubled its focus on edge use cases.

In the open source space there is also Spark Streaming or similar Apache Storm data streaming technologies.

SWOT ANALYSIS

STRENGTHS SAS has a well-proven data-streaming offering, which now through a partnership with Cisco enables data analytics from the edge of the network to the datacenter.	**WEAKNESSES** No single vendor has a complete IoT stack, and Cisco and SAS are no exception. The technology does not yet support the statistical language R yet, either.
OPPORTUNITIES More and more IoT projects are emerging, and those that have already been started tend to expand into more and more sophisticated projects.	**THREATS** While SAS and Cisco are working together on the SAS, Cisco IoT Analytics Platform, some companies may prefer a platform that has even broader, multi-vendor technology integration possibilities.