IBM Systems MEDIA

SAS Viya is Now Available on POWER

SAS Viya now runs on POWER9; learn what it means for you.

By Jaqui Lynch

12/02/2019

For years IBM has partnered with SAS to underpin SAS Grid Manager for platforms with Spectrum LSF, allowing clients to better manage their SAS environment across a grid. SAS itself has run for many years on POWER and is widely used for analytics and data mining on that platform. SAS Viya is the new computing architecture from SAS that is centered on the Cloud Analytics Server (CAS). With the addition of SAS Viya, IBM and SAS now provide an agile, full stack solution. SAS Viya is a parallelized architecture that can run across multiple nodes, holding as much data as possible in memory. This leads to reduced transaction times and can significantly reduce the time to solve complex problems and queries.

SAS and IBM have announced in the fourth quarter of 2019 that SAS Viya v3.5 will run on Red Hat Enterprise Linux 7.6 on IBM's POWER9 servers. This will allow Viya to take advantage of the CPU and memory bandwidth available on POWER9 as well as the incredible I/O performance. Additionally, on servers that have GPUs in them (i.e. the AC922), SAS Viya has the ability to utilize those GPUs and take advantage of the performance boost provided by the GPU to GPU and GPU to CPU performance. SAS Viya's in memory, parallel load approach requires extremely fast I/O, large bandwidth networking, and the ability to scale and deploy more threads per core–POWER9 addresses this need.

SAS Viya can be run in two execution modes–symmetric multi-processing (SMP) and massive parallel processing (MPP). Viya can run on a single server with one or more LPARs or can be spread across multiple servers. The simplest way to start is to run a single node to see how it works and then move to multiple nodes as needed. In SMP mode everything is done in the one LPAR, whereas in MPP mode CAS uses worker nodes to offload the analytics and to parallelize the processing of data. The worker nodes do not have to be identical–some may have GPUs, others may not. On POWER9, deployment can be either scale-out (several smaller nodes) or scale-up (one or more nodes on a larger system). Scale-up deployments are very popular when trying to co-locate current SAS 9.4 deployments with SAS Viya, although it is not a requirement to co-locate them. Mixed modes allow for SAS 9.4 and CAS to be co-located on a scale-up server such as the E980, while the worker nodes are on smaller servers such as the AC922 where they can take advantage of the GPUs in those nodes.

A sample configuration on an E980 might have an AIX LPAR running SAS 9.4 and multiple RHEL 7.6 LPARs providing CAS, microservices and worker nodes. Additional LPARs can be added to scale the environment. Although there are no GPUs on the E980, a configuration like this allows you to take advantage of fast communications across the hypervisor as well as the memory, I/O and CPU performance in the E980. It should be noted that the E980 can use other nodes such as the AC922 as worker nodes to provide GPU support for workloads requiring that.

Because of the parallelism, SAS Viya requires a shared filesystem that is very scalable. Spectrum Scale fits this

requirement but it is not yet confirmed whether Spectrum Scale will be part of SAS Viya or not. Spectrum Scale can grow to petabytes of data and provides great parallel performance. It can also be shared by multiple platforms, as AIX, Linux and Windows can all access the same data in the same filesystem. The use of IBM flash systems, such as the FS9150, has also been shown to significantly improve performance.

Architecture

The SAS Viya architecture consists of three components to consider. They can all be installed on the same host, or distributed over multiple hosts:

1. CAS Server

The CAS Server is required for all deployments, regardless of type (full or programming-only). Memory requirements are determined by the amount of data processed along with the level of user activity.

2. Programming Runtime

The Programming Runtime consists of multiple components that are required for all deployments, regardless of type (full or programming-only). It includes the SAS compute server, SAS Foundation, SAS Studio, SAS Workspace Server, SAS/CONNECT Server, and any SAS/ACCESS engines that you have licensed. CPU is typically licensed by cores and both CPU and memory requirements are workload dependent.

3. Service Layer

The service layer consists of components that are required for a full deployment, as well as services that support specific SAS products. This layer includes the Core Services host group and all the other services that support SAS Viya analytics processing. The host groups that compose the Service Layer can be deployed on multiple hosts, and with as many CPU cores as needed for optimal performance and availability.

Planning

As with any environment, SAS Viya requires some planning ahead of time in order to ensure success. Whether you deploy on POWER or x86, planning is critical. Some questions to ask yourself first are:

1. Architecture–SMP or MPP, Scale-up or Scale-out?
2. Am I co-locating some or all of SAS Viya with my current SAS 9.4 system?
3. Where will my CAS node go?

4. Where will my worker nodes go? Will I run all one type of node (i.e. AC922) or a mixture?
5. What is my shared filesystem?
6. How much CPU, Memory, I/O bandwidth, etc. will I need for each node?
7. How do I want to architect the network? Do I want to separate data traffic from management traffic, etc.?
8. What are the backup plans?
9. Do home directories need to be shared?
10. Are nodes needed to authenticate users against a shared infrastructure, such as LDAP or Active Directory?

These are just a few points to consider ahead of time. IBM and SAS are currently working on the sizing guides for SAS Viya on POWER9. Additionally, if RHEL is new to the environment, then some training or services may be required to get the new environment up and running.

Summary

SAS Viya is the third major evolution of the SAS distributed, in-memory technology. The ability to parallelize jobs in a properly architected environment will provide significant improvements in the time it takes to perform machine learning, analytics, deep learning and other artificial intelligence processes. POWER9 provides better capabilities with NVLink in comparison to Intel x86 due to the CPU to GPU performance in the AC922 servers. If you are considering SAS Viya today, it is highly recommended that you look at deploying it on POWER9 technologies so that you can take advantage of all that POWER9 and Viya offer together.

Many thanks to Harry Seifert from IBM's Washington Systems Center for the work he has done for many years in this area, especially in producing the SAS and POWER performance whitepapers.

References

1. http://www.circle4.com/movies/
2. SAS Viya Brochure

https://www.sas.com/content/dam/SAS/en_us/doc/overviewbrochure/sas-viya-108233.pdf (https://www.sas.com/content/dam/SAS/en_us/doc/overviewbrochure/sas-viya-108233.pdf)

3. SAS Viya for Linux (x86) Documentation

https://go.documentation.sas.com/?docsetId=dplyml0phy0lax&docsetTarget=p06vsqpjpj2motn1qhi5t40u8xf4.htm&docsetVersion=3.4&locale=en (https://go.documentation.sas.com/?docsetId=dplyml0phy0lax&docsetTarget=p06vsqpjpj2motn1qhi5t40u8xf4.htm&docsetVersion=3.4&locale=en)

4. POWER9 and FS9150 Performance Brief

https://www.ibm.com/downloads/cas/J6PBKQPG (https://www.ibm.com/downloads/cas/J6PBKQPG)

5. SAS Viya on Power Solutions Guide

https://www.ibm.com/downloads/cas/4EWGE0ZZ (https://www.ibm.com/downloads/cas/4EWGE0ZZ)

About the author
Jaqui Lynch has over 38 years of experience working with a projects and OSes across vendor platforms, including IBM Z, UNIX systems and more.