Paper 1834-2018

SAS Grid Infrastructure on AWS

Siddharth Prasad, CoreCompete LLC

ABSTRACT

SAS® Grid computing is a shared, centrally managed analytics computing environment that features workload balancing and management, high availability, and fast processing. A SAS® grid environment helps you incrementally scale your computing infrastructure over time as the number of users and the size of data grow. It also provides rolling maintenance and upgrades without any disruption to your users. This paper emphasizes on deploying SAS Grid on cloud (AWS infrastructure).

INTRODUCTION

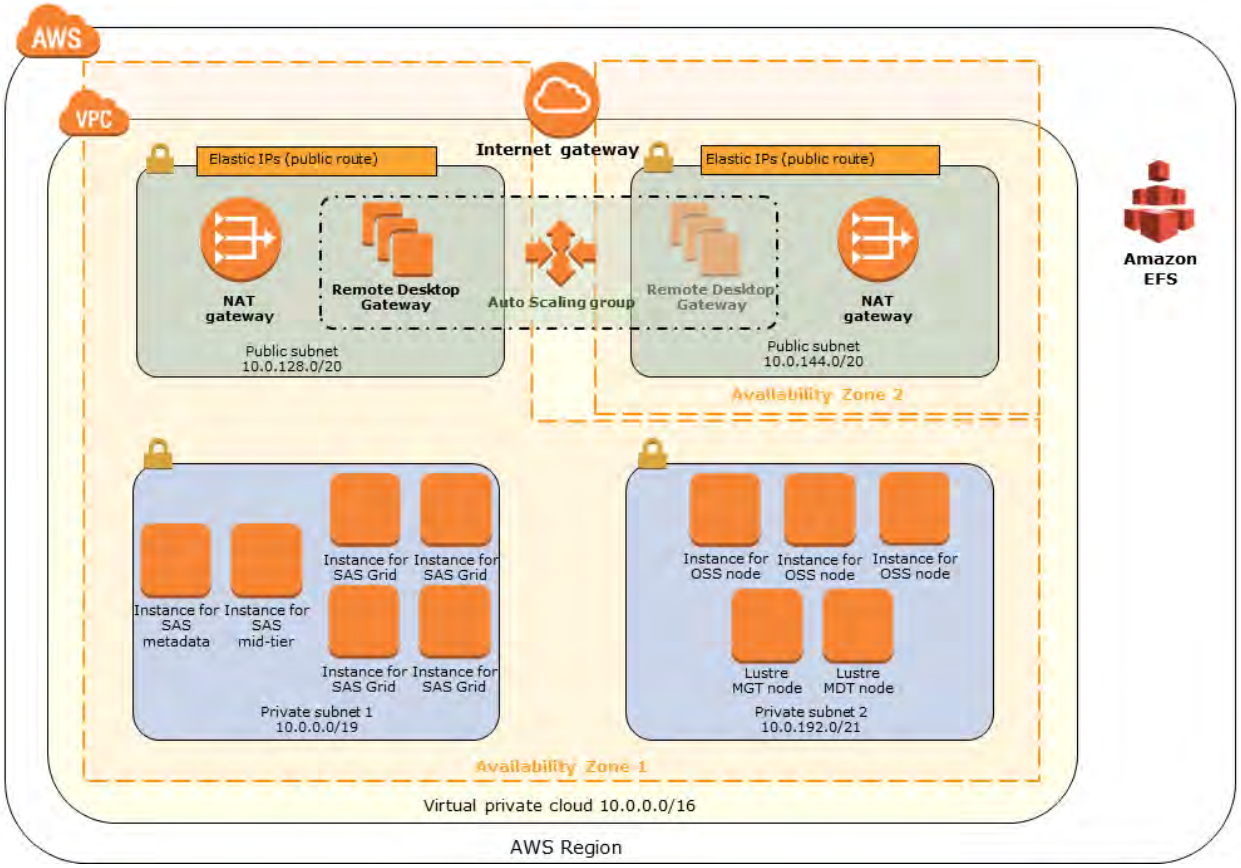
As part of the AWS quickstart for SAS grid we have built a cloudformation template, which launches the infrastructure for the SAS grid. This quickstart is for IT infrastructure architects, administrators, and DevOps professionals who are planning to implement or extend their SAS workloads on the AWS Cloud. It deploys the infrastructure for implementing SAS Grid Computing and related SAS components on Amazon Elastic Compute Cloud (Amazon EC2) instances and uses security groups, a virtual private cloud (VPC), subnets, and AWS Elastic Load Balancing to provide security and availability. A SAS grid environment in the cloud provides the elasticity and agility to scale your resources as needed. The quickstart automatically builds and configures the required infrastructure for SAS Grid Computing application installation, thereby reducing the dependency on your IT team. The effort required to plan, design, and implement the infrastructure is eliminated.

The Quick Start includes AWS CloudFormation templates and a guide that provides step-by-step instructions to help you get the most out of your deployment. This Quick Start was created by AWS in partnership with CoreCompete. CoreCompete is a big data analytics consulting organization, SAS Gold Partner, and AWS Advanced Consulting Partner.

WHAT THE QUICKSTART WILL BUILD

- A virtual private cloud (VPC) configured with public and private subnets according to AWS best practices. This provides the network infrastructure for your SAS Grid deployment.
- An Internet gateway to provide access to the Internet.
- A Remote Desktop Gateway instance in an Auto Scaling group, acting as a jump host.
- Managed NAT gateways to allow outbound Internet access for resources in the private subnets.
- In the private subnets, 2-6 EC2 instances for SAS Grid.

- In the private subnets, EC2 instances acting as Lustre MGT (Management), MDT (Metadata), and OSS nodes.
- Security groups for the Remote Desktop Gateway, SAS Grid, and Lustre stacks.
- Amazon Elastic File System (Amazon EFS) scalable file storage, to share the bootstrap information with the SAS Grid and Lustre nodes.
- Your choice to create a new VPC or deploy into your existing VPC on AWS. The template that deploys the Quick Start into an existing VPC skips the components marked by asterisks above.



Deploy the infrastructure for SAS Grid on AWS in a few simple steps:

1. Sign up for an AWS account.
2. Subscribe to the Intel Cloud Edition for Lustre AMI in AWS Marketplace.
3. Launch the Quick Start into a new VPC, if you want to build a new AWS infrastructure for your deployment.
   -or-
   Launch the Quick Start into an existing VPC, if you already have your AWS environment set up.
4. Validate the deployment by checking for mounted disks.

5. Obtain a SAS license by contacting A3@CoreCompete.com and install SAS Grid.

To customize your deployment, you can customize CIDR blocks for the subnets, choose the number of SAS Grid instances, and customize the number of Lustre OSS nodes and volume sizes.

DEPLOYMENT DETAILS

Deploy the infrastructure for SAS Grid on AWS in a few simple steps:

1. Sign up for an AWS account.
2. Subscribe to the Intel Cloud Edition for Lustre AMI in AWS Marketplace.
3. Launch the Quick Start into a new VPC, if you want to build a new AWS infrastructure for your deployment.
   -or-
   Launch the Quick Start into an existing VPC, if you already have your AWS environment set up.
4. Validate the deployment by checking for mounted disks.
5. Obtain a SAS license by contacting A3@CoreCompete.com and install SAS Grid.

To customize your deployment, you can customize CIDR blocks for the subnets, choose the number of SAS Grid instances, and customize the number of Lustre OSS nodes and volume sizes.

COST AND LICENCES

The client/customer are responsible for the cost of the AWS services used while running this Quick Start reference deployment. There is no additional cost for using the Quick Start.

The AWS CloudFormation template for this Quick Start includes configuration parameters that you can customize. Some of these settings, such as number of nodes and instance type, will affect the cost of deployment. See the pricing pages for each AWS service you will be using for cost estimates.

The Quick Start requires a subscription to the Amazon Machine Image (AMI) for Intel Cloud Edition for Lustre, which is available from AWS Marketplace, and additional pricing, terms, and conditions may apply.

This Quick Start doesn't install SAS Grid software. After using this Quick Start to set up the infrastructure for SAS Grid, you can obtain a SAS license and install the software.

CONCLUSION

By exercising this activity, we were able to achieve the infrastructure setup of SAS Grid software in as little as 15 minutes. In a traditional datacenter infrastructure, the same activity would have taken approximately 1 week. By creating this AWS Quickstart we look to provide efficient and faster solution to SAS Grid infrastructure setup.

REFERENCES

SAS papers on Performance Best Practices and Tuning Guides: http://support.sas.com/kb/42/197.html

Developing High-Performance, Scalable, cost effective storage solutions with Intel Cloud Edition Lustre and Amazon Web Services.

http://www.intel.com/content/dam/www/public/us/en/documents/reference-architectures/ICEL-RA.pdf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Siddharth Prasad
siddharth.prasad@corecompete.com