

Paper 3669-2019
SAS® Grid Quick Start 2.0 on Amazon Web Services

Siddharth Prasad, Core Compete Inc.
Srikanth Mergu, Core Compete Inc.

ABSTRACT

SAS® Grid 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 Quick Start bootstraps the infrastructure for a SAS® Grid cluster and installs the SAS® Grid software, which includes SAS® Grid Control Server, SAS® Grid nodes, SAS® Metadata Server, and SAS® mid-tier components. For storage, you can choose Cloud Edition for Lustre or IBM Spectrum Scale.

INTRODUCTION

The Quick Start 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 and related SAS® components on Amazon Elastic Compute Cloud (Amazon EC2) instances and uses security groups, a virtual private cloud (VPC), subnets to provide security and availability. A SAS® Grid environment in the cloud provides the elasticity and agility to scale your resources as needed. The Quick Start automatically builds and configures the required infrastructure and installs the SAS® Grid software, thereby reducing the dependency on your IT team. The effort required to plan, design, and implement the infrastructure and SAS® Grid software is eliminated, so your business can start using the environment right away.

SAS® Grid Components:

SAS® Grid consists of the following components:

- SAS® Grid Control Server
- SAS® Grid nodes
- SAS® Metadata Server
- SAS® mid-tier components

This Quick Start bootstraps the infrastructure for your SAS® Grid cluster by provisioning single EC2 instances for SAS® Metadata Server and mid-tier components and provisioning multiple EC2 instances for SAS® Grid. SAS® Grid requires a network share that all computers on your cluster can access. This can be a Network File System (NFS) mount, a directory on a SAN, an SMBFS/CIFS mount, or any other method of creating a directory that is shared among all the machines in the grid. To meet this requirement, the Quick Start sets up Cloud Edition for Lustre or IBM Spectrum Scale, which are clustered parallel file system.

WHAT THE QUICKSTART WILL BUILD

This Quick Start deploys the following components:

- AWS infrastructure components for SAS® Grid
- Your choice of one of these two options for the storage stack:
 - Cloud Edition for Lustre components
 - IBM Spectrum Scale
- SAS® Grid software on SAS® Grid instances
- BASE SAS® and SAS® Office Analytics software stack on SAS® Metadata Server, SAS® Grid nodes, and SAS® mid-tier.

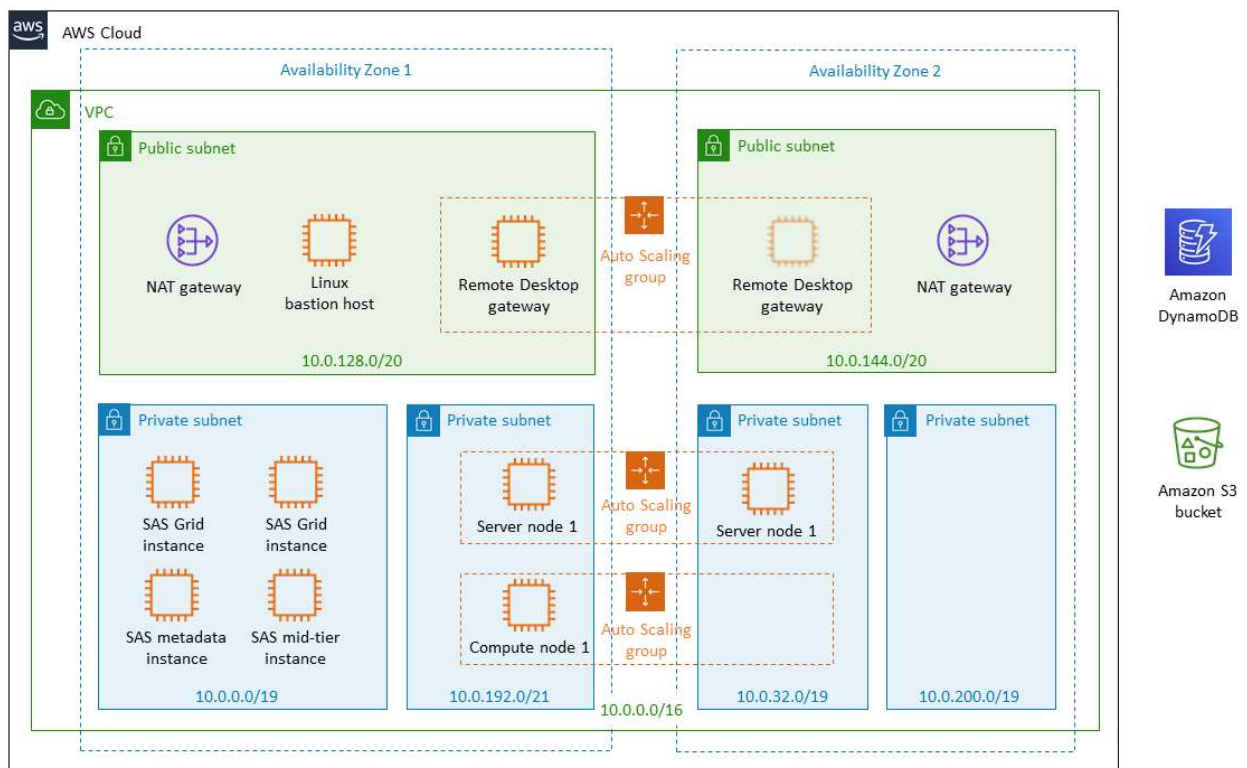


Figure 2: Quick Start architecture for SAS® Grid with the Spectrum Scale storage stack

DEPLOYMENT DETAILS

Step 1. Prepare Your AWS account

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

1. Sign up for an AWS account.
2. Use the region selector in the navigation bar to choose the AWS Region where you want to deploy the infrastructure for SAS® Grid on AWS.
3. Create an IAM user with administrator privileges, and enable multi-factor authentication (MFA) for both root and IAM user accounts.

4. Create up to three key pairs in your preferred region, for Remote Desktop Gateway, SAS® Grid, and Lustre or IBM Spectrum Scale instances. One key pair can be used for all three. You're responsible for managing and securing these EC2 key pairs .
5. Request a service limit increase for the following: – If the default limit of 5 Elastic IP addresses has already been used, increase the limit by 3. – Increase the limit for R4 and C4 instances to 50. – Increase the limit for i3.8xlarge instances to 6 (minimum).

Step 2. Request SAS® Licenses and Upload Files

1. Request SAS® software licenses. Your SAS® account team and the SAS® Enterprise Excellence Center can advise on the appropriate software licensing and sizing to meet your workload and performance needs. If you do not have an assigned SAS® account team, you can contact us at A3@CoreCompete.com. You will receive a SAS® software order confirmation email with license files for two different sites with the following products:
 - Site 1 license file:
 - Base SAS®
 - SAS® Metadata Server
 - Site 2 license file:
 - Base SAS®
 - SAS® Grid Manager Control Server
 - SAS®/CONNECT
 - SAS®/GRAPH
 - SAS®/STAT
 - Optional products
 - SAS®/ACCESS interface to Amazon Redshift
 - SAS®/ACCESS interface to ODBC
2. Use your preferred zip program to download the SAS® Software Depot into a .tar file. (The SAS® Software Depot is a collection of SAS® installation files.) For example, if you're using 7-Zip, you can right-click SAS® Software Depot and choose Add to archive it.
3. Upload the order and plan files to an S3 bucket. For the Quick Start deployment, your .tar file for the SAS® software order and plan files must be in an Amazon S3 location accessible from your installation account.
 - a. Upload the .tar file to the S3 bucket.
 - b. Upload the following plan files to the S3 bucket:
 - https://raw.githubusercontent.com/aws-quickstart/quickstart-SAS@grid/master/playbooks/templates/final_plan_meta_only.xml
 - https://raw.githubusercontent.com/aws-quickstart/quickstart-SAS@grid/master/playbooks/templates/final_plan_without_meta.xml

Step 3. Subscribe to the Lustre AMI

The Cloud Edition for Lustre software is available from AWS Marketplace. Before you deploy the Quick Start, you must subscribe to the AMI:

1. Log in to your AWS account.
2. Open the AWS Marketplace webpage for the Lustre AMI.
3. Choose Continue to Subscribe to view the license terms and launch information.

Step 4. Launch the Quick Start

Choose one of the following options to launch the AWS CloudFormation template into your AWS account. For help choosing an option, see deployment options earlier in this guide.

COST AND LICENCES

The client/customer is 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 templates for this Quick Start include configuration parameters that you can customize. Some of these settings, such as instance type, will affect the cost of deployment. For cost estimates, see the pricing pages for each AWS service you will be using. Prices are subject to change. The Quick Start requires SAS® software licenses. The SAS® account team and the SAS® Enterprise Excellence Center can advise on the appropriate software licensing and sizing to meet your workload and performance needs. If you do not have an assigned SAS® account team, contact us at A3@CoreCompete.com.

If user chooses Lustre for the storage stack, the Quick Start will require a subscription to the Amazon Machine Image (AMI) for DDN Cloud Edition for Lustre, which is available from AWS Marketplace, and additional pricing, terms, and conditions may apply.

If user chooses IBM Spectrum Scale for the storage stack, no AMI subscriptions are necessary. The IBM Spectrum Scale deployment launches an EC2 instance running the Red Hat Enterprise Linux (RHEL) version 7.4 operating system and deploys a trial version of the IBM Spectrum Scale software.

CONCLUSION

By exercising this activity, we were able to achieve the infrastructure setup, installation and configuration of SAS® Grid software in as little as 3 hours. In a traditional datacenter infrastructure, the same activity would have taken approximately 2 weeks. By creating this AWS quick start 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

Srikanth Mergu

Srikanth.mergu@corecompete.com