SAS 9.4M4 Grid on the AWS Cloud

Quick Start Reference Deployment

December 2018  
Last update: June 2020 ([revisions](#_Document_Revisions))

*Core Compete  
SAS Institute*

*AWS Quick Start Reference Team*

Contents

[Overview 2](#_Toc533164279)

[SAS Grid Components 3](#_Toc533164280)

[SAS Grid Features 3](#_Toc533164281)

[What the Quick Start Deploys 4](#_Toc533164282)

[Costs and Licenses 4](#_Toc533164283)

[Architecture 5](#_Toc533164284)

[Prerequisites 9](#_Toc533164285)

[Specialized Knowledge 9](#_Toc533164286)

[Deployment Options 9](#_Toc533164287)

[Planning Your Deployment 9](#_Toc533164288)

[Deployment Steps 10](#_Toc533164289)

[Step 1. Prepare Your AWS account 10](#_Toc533164290)

[Step 2. Request SAS Licenses and Upload Files 10](#_Toc533164291)

[Step 3. Subscribe to the Lustre AMI 11](#_Toc533164292)

[Step 4. Launch the Quick Start 14](#_Toc533164293)

[Step 5. Validate Your Deployment 21](#_Toc533164294)

[Step 6. Run SAS Post-Deployment Scripts 22](#_Toc533164295)

[FAQ 22](#_Toc533164296)

[Additional Resources 24](#_Toc533164297)

[Send Us Feedback 25](#_Toc533164298)

[Document Revisions 25](#_Toc533164299)

This Quick Start deployment guide was created by Core Compete and SAS Institute in collaboration with Amazon Web Services (AWS). Core Compete is a big data analytics consulting organization, SAS Gold Partner, and AWS Advanced Consulting Partner. SAS Institute is an AWS Advanced Technology Partner.

[Quick Starts](http://aws.amazon.com/quickstart/) are automated reference deployments that use AWS CloudFormation templates to launch, configure, and run the AWS compute, network, storage, and other services required to deploy a specific workload on AWS.

## Overview

This Quick Start reference deployment guide provides step-by-step instructions for launching and configuring the required IT infrastructure and installing SAS grid software in the AWS Cloud.

[SAS grid](https://www.sas.com/en_us/software/grid-manager.html) 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.

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, which is a parallel file system.

**Note** The Quick Start incorporates the prerequisites recommended by SAS for building the infrastructure for SAS grid. It also deploys SAS grid with BASE SAS and the SAS Office Analytics software stack.

### SAS Grid Features

SAS grid provides the following features for a SAS environment:

* Improves efficiency and utilization of computing resources by distributing jobs submitted by different users through dynamic, resource-based load balancing
* Enables you to run larger or more complex analyses by making computing resources available to multiple users and multiple applications
* Provides the ability to manage jobs, queues, hosts, and users across the enterprise
* Enables you to implement rules-based job queues and prioritization to govern the use of computing resources
* Automates identification, allocation, management, and optimization of computing resources and program flows
* Supports centralized policies for simplifying the administration of the SAS environment
* Adds high-availability capabilities for SAS Metadata Server and other critical SAS services
* Provides a hot-standby machine for failover by using the nodes in the grid
* Performs load balancing for all SAS servers to provide improved throughput and response time for all SAS clients
* Works with SAS Data Integration Studio and SAS Enterprise Guide, which can import SAS programs to the grid
* Accelerates the processing speeds of applicable SAS programs and applications, and provides more efficient computing resource utilization
* Supports easy configuration of user-written programs and many SAS solutions for easy submission to a grid of shared resources
* Provides high availability and high resilience for the SAS environment to support mission-critical applications

For additional information, see the [SAS documentation](http://support.sas.com/rnd/scalability/grid/) and [video](http://www.youtube.com/embed/x-qLpMyAF54?rel=0&fs=1&wmode=transparent&autoplay=1&showinfo=1&modestbranding=1&?theme=dark&color=white&cc_load_policy=0&hl=en&cc_lang_pref=en&?autohide=1#1280,720).

### What the Quick Start Deploys

This Quick Start deploys the following components:

* AWS infrastructure components for SAS grid (detailed in the [Architecture](#_Architecture) section)
* Your choice of one of these two options for the storage stack:
* Cloud Edition for Lustre components (see the [Lustre documentation](http://wiki.old.lustre.org/manual/LustreManual20_HTML/UnderstandingLustre.html#50438250_19307) for details)
* IBM Spectrum Scale (see the [IBM Spectrum Scale deployment guide](https://fwd.aws/rxqbD) for details)
* 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.

### Costs and Licenses

You 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 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, as detailed in [step 2](#_Step_2._Request). 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, contact us at [support@sas.com](mailto:support@sas.com).

If you choose Lustre for the storage stack, the Quick Start will require a subscription to the Amazon Machine Image (AMI) for  [DDN Cloud Edition for Lustre](https://aws.amazon.com/marketplace/pp/B07FTSL9VX), which is available from AWS Marketplace, and additional pricing, terms, and conditions may apply. We’ve provided subscription instructions in [step 3](#_Step_3._Subscribe) of the deployment steps.

If you choose 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. See the [IBM Spectrum Scale deployment guide](https://fwd.aws/rxqbD) for restrictions associated with trial evaluations.

The use of IBM Spectrum Scale on AWS (including all packages provided via the Quick Start offering, and packages derived from these) is only intended to be used for a maximum of 90 days, and is not intended for production use. IBM may decide to de-authorize access to the code, and the use of this code, at any time. After the trial period, you are responsible for acquiring the necessary licenses directly from IBM to use IBM Spectrum Scale. The [IBM](https://developer.ibm.com/storage/products/ibm-spectrum-scale/#evaluate) [Spectrum Scale evaluations page](https://developer.ibm.com/storage/products/ibm-spectrum-scale/#evaluate) will be updated with details on how to proceed with acquiring an IBM Spectrum Scale license after the 90-day trial expires.

## Architecture

Deploying this Quick Start for a new virtual private cloud (VPC) with **default parameters** builds the infrastructure illustrated in Figures 1 and 2 for SAS grid in the AWS Cloud.

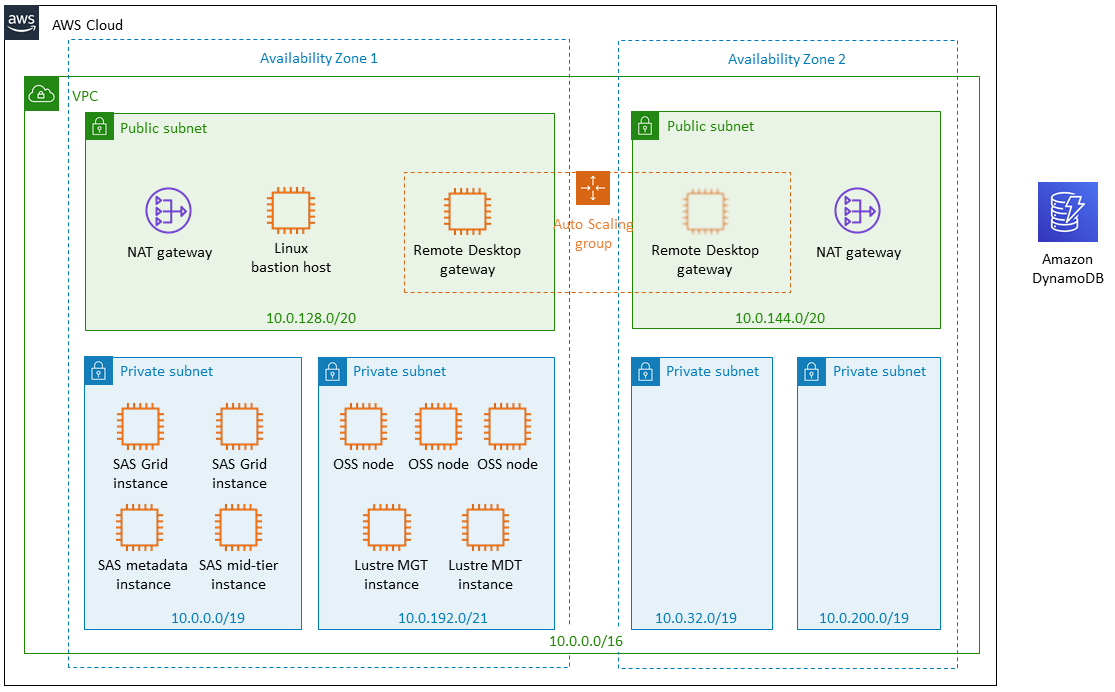


Figure 1: Quick Start architecture for SAS grid with the Lustre storage stack

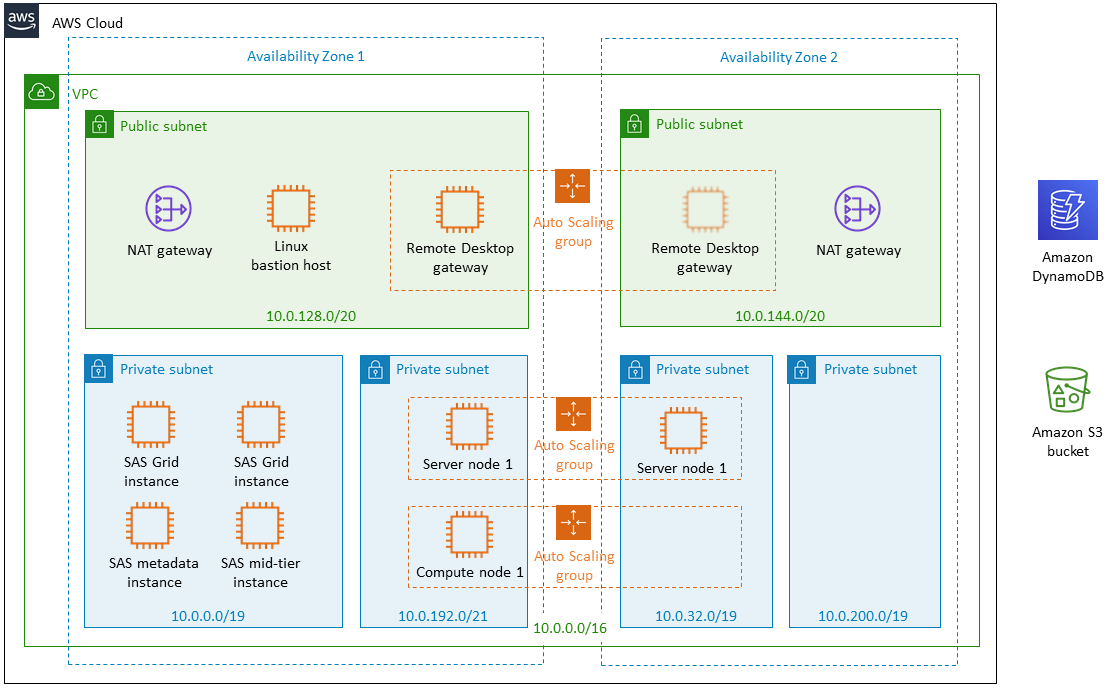


Figure 2: Quick Start architecture for SAS grid with the Spectrum Scale storage stack

You can choose the number of SAS grid compute nodes, based on your requirements, during deployment. (See the [Planning](#_Planning_Your_Deployment) section for details on the utilization choices.)

The Quick Start sets up the following:

* A VPC that spans two Availability Zones.\*

Following AWS best practices, the Quick Start uses two Availability Zones for high availability. For tips on handling failover scenarios in this architecture, see the [FAQ](#failover).

* Six subnets, two public and four private.\*

Following AWS best practices, the Quick Start uses two Availability Zones and sets up one public subnet and two private subnets in each Availability Zone. For the SAS grid infrastructure, the Quick Start uses the two private subnets in one of these Availability Zones. The private subnets in the second Availability Zone are used only if you choose Spectrum Scale for the storage stack. For tips on handling failover scenarios in this architecture, see the [FAQ](#failover).

* In a public subnet, a Remote Desktop Gateway in an Auto Scaling group with a default of one instance, acting as a jump host.\*
* In a public subnet, a Linux bastion host acting as an Ansible Controller Host.\*
* In the public subnets, managed NAT gateways to allow outbound internet access for resources in the private subnets.\*
* In the first private subnet of Availability Zone 1, four EC2 instances for SAS grid.

You can choose 2-20 instances; see the [Planning](#_Planning_Yyour_Ddeployment) section for details.

* The components provisioned in the second private subnet of Availability Zone 1 depend on the storage stack you’ve selected.

Lustre as storage stack:

* One EC2 instance acting as a Lustre MGT (Management) node
* One EC2 instance acting as a Lustre MDT (Metadata) node
* 3 EC2 instances for OSS nodes (you can choose up to 15 nodes)

IBM Spectrum Scale as storage stack:

* 2-64 EC2 instances acting as server nodes launched in an Auto Scaling group
* One EC2 instance acting as a compute (management) node launched in an Auto Scaling group
* Security groups for the following stacks:
* Remote Desktop Gateway: Allows port 3389 for RDP access (to the bastion host)
* SAS grid: Allows connectivity on all ports for all SAS grid hosts, Lustre hosts, and Remote Desktop Gateway host
* IBM Spectrum Scale (if selected): Allows connectivity on all ports for all SAS grid hosts, IBM Spectrum Scale server nodes, IBM Spectrum Scale compute nodes, and Remote Desktop Gateway host
* Lustre (if selected): Allows connectivity on all ports for all SAS grid hosts, Lustre hosts, and Remote Desktop Gateway host
* Amazon DynamoDB: Used for tracking AMI backups of EC2 instances launched by the Quick Start.

**\*** The template that deploys the Quick Start into an existing VPC skips the tasks marked by asterisks.

## Prerequisites

### Specialized Knowledge

Before you deploy this Quick Start, we recommend that you become familiar with the following AWS services. (If you are new to AWS, see [Getting Started with AWS](https://aws.amazon.com/getting-started/).)

* [Amazon DynamoDB](https://docs.aws.amazon.com/dynamodb/)
* [Amazon EBS](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html)
* [Amazon EC2](https://docs.aws.amazon.com/ec2/)
* [Amazon S3](https://docs.aws.amazon.com/s3/)
* [Amazon VPC](https://docs.aws.amazon.com/vpc/)
* [AWS Auto Scaling](http://docs.aws.amazon.com/autoscaling/)

## Deployment Options

This Quick Start provides two deployment options:

* **Deploy SAS grid into a new VPC** (end-to-end deployment). This option builds a new AWS environment consisting of the VPC, subnets, NAT gateway, and other components, and then deploys SAS grid into this new VPC.
* **Deploy SAS grid into an existing VPC**. This option provisions SAS grid in your existing AWS infrastructure.

The Quick Start provides separate templates for these options. It also lets you configure additional settings such as infrastructure size, CIDR blocks for the VPC and subnets, and EBS volume sizes, as discussed later in this guide.

**Note** For a list of supported regions for Lustre Software, see [DDN Cloud Edition for Lustre\* software](https://aws.amazon.com/marketplace/pp/B07FTSL9VX) in AWS Marketplace.

## Planning Your Deployment

Based on the number of CPU cores you have bought for the SAS software license, you can specify the number of EC2 instances for the Quick Start to launch. For example, if you have obtained a SAS license for 32 physical CPU cores, you should enter 2 SAS grid instances for the **Required number of SAS grid instances** parameter (using the i3.8xlarge instance type) when you launch the Quick Start.

## Deployment Steps

### Step 1. Prepare Your AWS account

1. If you don’t already have an AWS account, create one at <https://aws.amazon.com> by following the on-screen instructions.
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.

**Note** For a list of supported regions for Lustre Software, see [DDN Cloud Edition for Lustre\* software](https://aws.amazon.com/marketplace/pp/B07FTSL9VX) in AWS Marketplace.

1. Create an IAM user with administrator privileges, and enable multi-factor authentication (MFA) for both root and IAM user accounts.
2. Create up to three [key pair](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-key-pairs.html)s 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 .
3. [Request a service limit increase](https://console.aws.amazon.com/support/home#/case/create?issueType=service-limit-increase&limitType=service-code-) 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

**Important** Your deployment will fail if you deviate from these order requirements.

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 [support@sas.com](mailto:support@sas.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

1. 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.
2. 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.
3. Upload the.tar file to the S3 bucket.
4. 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>

To upload the files, see [How Do I Upload Files and Folders to an S3 Bucket?](http://docs.aws.amazon.com/AmazonS3/latest/user-guide/upload-objects.html) in the AWS documentation.

**Note** Please make sure that the S3 bucket is **not** publicly accessible.

### Step 3. Subscribe to the Lustre AMI

**Note** Skip this step if you want to use IBM Spectrum Scale instead of Lustre for storage.

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](https://aws.amazon.com/marketplace/pp/B07FTSL9VX).
3. Choose **Continue** **to Subscribe** to view the license terms and launch information (Figure 3).

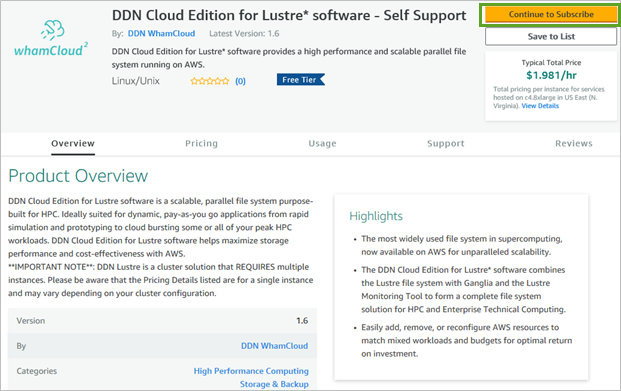
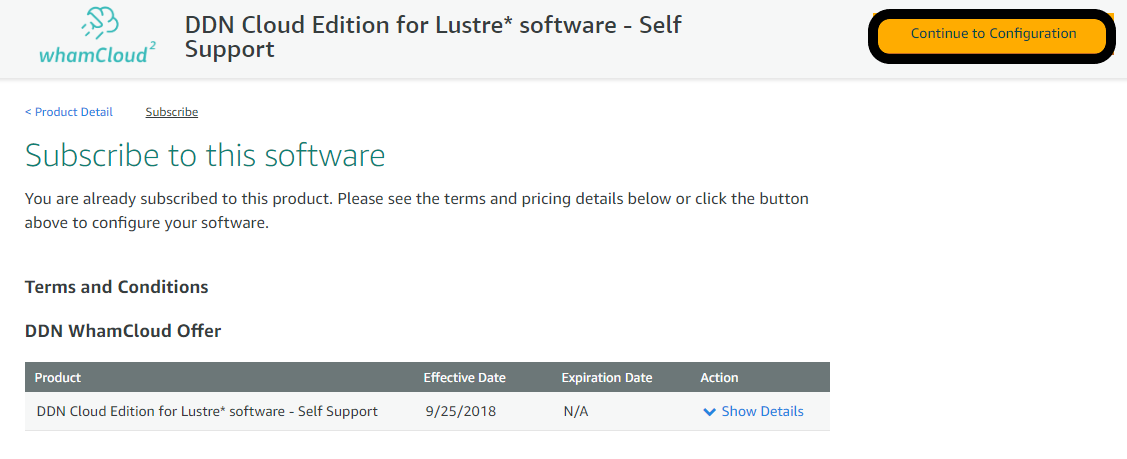


Figure 3: DDN Cloud Edition for Lustre AMI page in AWS Marketplace

1. Choose **Continue to Configuration** (Figure 4).



**Figure 4: Lustre AMI subscription page in AWS Marketplace**

1. Choose **Continue to Launch** to select the instance type and the AWS Region   
   (Figure 5).

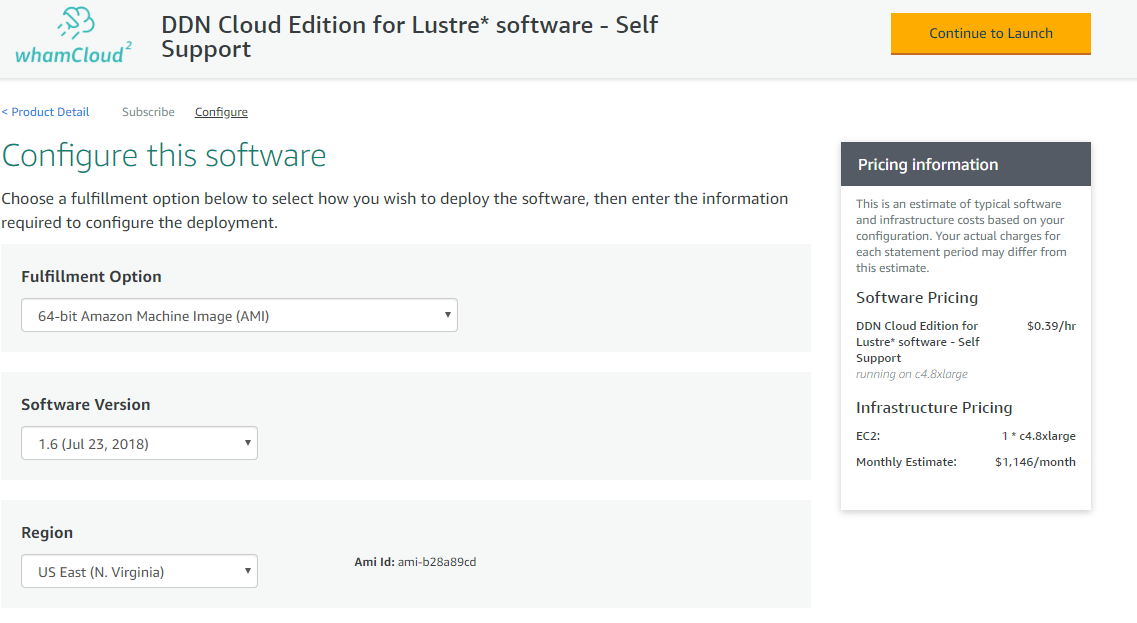


Figure 5: Choosing the AWS Region and accepting license terms for the AMI

1. When the subscription process is complete, exit out of AWS Marketplace without further action. **Do not** provision the software from AWS Marketplace—the Quick Start will deploy the AMI for you.

### Step 4. Launch the Quick Start

**Note** You are responsible for the cost of the AWS services used while running this Quick Start reference deployment. There is no additional cost for using this Quick Start. For full details, see the pricing pages for each AWS service you will be using in this Quick Start. Prices are subject to change. In addition, you are responsible for managing and securing the EC2 key pairs used in the Quick Start.

1. Choose one of the following options to launch the AWS CloudFormation template into your AWS account. For help choosing an option, see [deployment options](#_Deployment_Options) earlier in this guide.

|  |  |
| --- | --- |
| [Option 1](#_Scenario_1:_Deploy_1)  [**Launch**](https://fwd.aws/vA89R)  Deploy into a  new VPC on AWS | [Option 2](#_Scenario_2:_Extending_1)  [**Launch**](https://fwd.aws/k6bVB)  Deploy into an  existing VPC on AWS |

**Important** If you’re deploying the infrastructure for SAS grid into an existing VPC, make sure that your VPC is set up with public and private subnets, and Linux bastion host, NAT gateway, and Remote Desktop Gateway instances, as described in the [Architecture](#_Architecture) section. Note that this Quick Start doesn’t support [shared subnets](https://docs.aws.amazon.com/vpc/latest/userguide/vpc-sharing.html). You’ll also need the domain name option configured in the DHCP options as explained in the [Amazon VPC documentation](http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_DHCP_Options.html). You’ll be prompted for your VPC settings when you launch the Quick Start.

Each deployment takes about 2 hours to complete.

1. Check the region that’s displayed in the upper-right corner of the navigation bar, and change it if necessary. (See the note in [step 1](#_Step_1._Prepare) for supported regions.) This is where the network infrastructure for SAS grid will be built. The template is launched in the US West (Oregon) Region by default.
2. On the **Select Template** page, keep the default setting for the template URL, and then choose **Next**.
3. On the **Specify Details** page, change the stack name if needed. Review the parameters for the template. Provide values for the parameters that require input. For all other parameters, review the default settings and customize them as necessary. When you finish reviewing and customizing the parameters, choose **Next**.

In the following tables, parameters are listed by category and described separately for the two deployment options:

* [Parameters for deploying the infrastructure for SAS grid into a new VPC](#sc1)
* [Parameters for deploying the infrastructure for SAS grid into an existing VPC](#sc2)
* **Option 1: Parameters for deployment into a new VPC**

[View template](https://fwd.aws/kJ9kP)

*SAS grid license information:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| SAS Software Depot tar file in S3 (SASSoftwareDepot) | *Requires input* | The S3 location of the SAS Software Depot .tar file you received in the software order confirmation email attachment. For example, if you uploaded this file to the S3 bucket location **sas-grid-deployment/sas-depot/SASSoftwareDepot\_Final.tar**, enter **sas-grid-deployment/sas-depot**. |
| SAS Software Depot name (SASSoftwareDepotName) | *Requires input* | The name of the SAS Software Depot, without the .tar file extension. For example, if you uploaded **SASSoftwareDepot\_Final.tar** in step 2, enter **SASSoftwareDepot\_Final** in this field. |
| SAS plan files folder in S3 (SASPlanFiles) | *Requires input* | The S3 bucket location where you uploaded the SAS plan files; for example, **sasgrid-quickstart-staging/sas-depot/PlanFiles**. |
| SAS Metadata Server license file (SASLicenseMeta) | *Requires input* | The name of the SAS Metadata Server license file; for example, **SAS94\_xxxxxx\_xxxxxxxx\_LINUX\_X86-64.txt**. You will find this file inside the SAS Software Depot, in the folder sid\_files. |
| SAS Application Server license file (SASLicenseApp) | *Requires input* | The name of the SAS Application Server license file;for example**, SAS94\_xxxxxx\_xxxxxxxx\_LINUX\_X86-64.txt**. You will find this file inside the SAS Software Depot, in the folder sid\_files. |

*Network configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Availability Zones (AvailabilityZones) | *Requires input* | The list of Availability Zones to use for the subnets in the VPC. You must select a minimum of two Availability Zones. |
| VPC CIDR (VPCCIDR) | 10.0.0.0/16 | The CIDR block for the VPC. |
| Private subnet 1 (AZ1) CIDR (PrivateSubnet1CIDR) | 10.0.0.0/19 | The CIDR block for the first private subnet in Availability Zone 1, where the SAS grid instances will be deployed. |
| Private subnet 2 (AZ1) CIDR (PrivateSubnet2CIDR) | 10.0.192.0/21 | The CIDR block for the second private subnet in Availability Zone 1, where the Lustre instances or IBM Spectrum Scale server and compute nodes will be deployed. |
| Private subnet 3 (AZ2) CIDR (PrivateSubnet3CIDR) | 10.0.32.0/19 | The CIDR block for the first private subnet in Availability Zone 2. |
| Private subnet 4 (AZ2) CIDR (PrivateSubnet4CIDR) | 10.0.200.0/21 | The CIDR block for the second private subnet in Availability Zone 2. |
| Public subnet 1 CIDR (PublicSubnet1CIDR) | 10.0.128.0/20 | The CIDR block for the public (DMZ) subnet located in Availability Zone 1. |
| Public subnet 2 CIDR (PublicSubnet2CIDR) | 10.0.144.0/20 | The CIDR block for the public (DMZ) subnet located in Availability Zone 2. |

*Microsoft RD Gateway configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| RDGW key pair name (RDGWKeyPairName) | *Requires input* | A public/private key pair, which allows you to securely connect to your instance after it launches. |
| RDGW instance type (RDGWInstanceType) | t2.large | The EC2 instance type for the Remote Desktop Gateway instances. |
| Admin user name (AdminUser) | StackAdmin | The user name (5-25 alphanumeric characters) for the new local administrator account. |
| Admin password (AdminPassword) | *Requires input* | The password for the administrative account. This must be 8-32 characters, including letters, numbers, and symbols. |
| Domain DNS name (DomainDNSName) | example.com | The fully qualified domain name (FQDN) of the forest root domain. |
| Allowed RDGW external access CIDR (RDGWCIDR) | *Requires input* | The CIDR block that is allowed to access the Remote Desktop Gateway instances. We recommend that you set this value to a trusted IP range. For example, you might want to grant access only to your corporate network. |

*Storage stack configuration:*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter label (name) | Default | | Description | |
| Select the storage stack (StorageStack) | *Requires input* | | The storage stack you want to use. The two choices are Lustre or IBM Spectrum Scale (GPFS). | |
| Storage node count (StorageNodeCount) | 2 | | The number of server instances you want to provision, if you’re using IBM Spectrum Scale for storage. You can select 2-64 instances. For more information, see the [Planning](#_Planning_Yyour_Ddeployment) section. | |
| EBS volume type for storage instances (StorageEBSVolumeType) | gp2 | | The EBS volume type you want to use for the storage instances. You can choose General Purpose SSD (gp2), Provisioned IOPS SSD (io1), Cold HDD (sc1), Throughput Optimized HDD (st1), or standard. | |
| Storage disk size for each node (StorageEBSVolumeSize) | 500 | | The EBS volume size (in GiB) for storage nodes. You can specify 100-1638400 GiB, in multiples of 100. | |
| Storage key pair name (StorageKeyPairName) | | *Requires input* | | A public/private key pair for storage instances. |

*Personal configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Operator email (OperatorEmail) | *Requires input* | The email address that notifications of any scaling operations should be sent to. |

*License information:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| License agreement terms (LicenseAgreementTerms) | *Requires input* | Review the [licensing terms for IBM Spectrum Scale](http://spectrumscale-license.s3-website-us-west-2.amazonaws.com/), and then choose **Accept** to indicate your acceptance. |

*SAS grid EC2 configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| SAS Grid key pair name (SASGridKeyPairName) | *Requires input* | A public/private key pair, which allows you to securely connect to your SAS grid instances after launch. |
| Required number of SAS Grid instances (NoOfSASGridInstances) | 2 | The number of SAS grid instances you want to provision. You can select 2-20 instances, depending on the number of cores you are licensed for. For example, if you are licensed for 32 cores, you can keep the default setting of 2 in the text field.  For more information, see the [Planning](#_Planning_Yyour_Ddeployment) section. |
| Instance type for SAS Grid instances (SASGridInstanceType) | i3.8xlarge | The EC2 instance type you want to use for the SAS grid instances. |

*Lustre EC2 configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Number of Lustre OSS nodes (NumberOfOSSNodes) | 3 | The number of Lustre OSS nodes you want to deploy. You can specify 3-15 nodes. |
| EBS volume size for Lustre OSS (LustreOSSEBSVolumeSize) | 100 | The EBS volume size (in GiB) for Lustre OSS nodes. You can specify 100-9900 GiB, in multiples of 100. |

*AWS Quick Start configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Quick Start S3 bucket name (QSS3BucketName) | aws-quickstart | The S3 bucket you have created for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. The bucket name can include numbers, lowercase letters, uppercase letters, and hyphens, but should not start or end with a hyphen. |
| Quick Start S3 key prefix (QSS3KeyPrefix) | quickstart-sas-grid/ | The [S3 key name prefix](https://docs.aws.amazon.com/AmazonS3/latest/dev/UsingMetadata.html) used to simulate a folder for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. This prefix can include numbers, lowercase letters, uppercase letters, hyphens, and forward slashes. |

* **Option 1: Parameters for deployment into an existing VPC**

[View template](https://fwd.aws/nAR63)

*SAS Grid license information:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| SAS Software Depot tar file in S3 (SASSoftwareDepot) | *Requires input* | The S3 location of the SAS Software Depot .tar file you received in the software order confirmation email attachment. For example, if you uploaded this file to the S3 bucket location **sas-grid-deployment/sas-depot/SASSoftwareDepot\_Final.tar**, enter **sas-grid-deployment/sas-depot**. |
| SAS Software Depot name  (SASSoftwareDepotName) | *Requires input* | The name of the SAS Software Depot, without the .tar file extension. For example, if you uploaded **SASSoftwareDepot\_Final.tar** in step 2, enter **SASSoftwareDepot\_Final** in this field. |
| SAS plan files folder in S3  (SASPlanFiles) | *Requires input* | The S3 bucket location where you uploaded the SAS plan files; for example, **sasgrid-quickstart-staging/sas-depot/PlanFiles**. |
| SAS Metadata Server license file (SASLicenseMeta) | *Requires input* | The name of the SAS Metadata Server license file; for example, **SAS94\_xxxxxx\_xxxxxxxx\_LINUX\_X86-64.txt**. You will find this file inside the SAS Software Depot, in the folder sid\_files. |
| SAS Application Server license file (SASLicenseApp) | *Requires input* | The name of the SAS Application Server license file;for example, **SAS94\_xxxxxx\_xxxxxxxx\_LINUX\_X86-64.txt**. You will find this file inside the SAS Software Depot, in the folder sid\_files. |

*Network configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| VPC ID (VPCID) | *Requires input* | The ID of your existing VPC (e.g., vpc-0343606e). |
| Private subnet 1 ID (PrivateSubnet1ID) | *Requires input* | The ID of the first private subnet in Availability Zone 1 of your existing VPC (e.g., subnet-a0246dcd). |
| Private subnet 2 ID (PrivateSubnet2ID) | *Requires input* | The ID of the second private subnet in Availability Zone 1 of your existing VPC (e.g., subnet-b58c3d67). |
| Private subnet 3 ID (PrivateSubnet3ID) | *Requires input* | The ID of the private subnet in Availability Zone 2 of your existing VPC (e.g., subnet-d25a4e53). |
| Public subnet 1 ID (PublicSubnet1ID) | *Requires input* | The ID of the first public subnet in your existing VPC (e.g., subnet-e24f1g23). |
| VPC CIDR (VPCCIDR) | 10.0.0.0/16 | The CIDR block of your existing VPC. |

*Linux bastion configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Admin password (AdminPassword) | *Requires input* | The password for the administrative account. This must be 8-32 characters, including letters, numbers, and symbols. |
| Admin ingress location (AdminIngressLocation) | *Requires input* | The CIDR block that is allowed to access the Linux bastion host instance. |
| Domain DNS name (DomainDNSName) | example.com | The fully qualified domain name (FQDN) of the forest root domain. |
| RDGW security group ID (RDGWSG) | *Requires input* | The ID of the security group for Remote Desktop Gateway instances (e.g., sg-0343606e). |

*Storage stack configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Select the storage stack (StorageStack) | *Requires input* | The storage stack you want to use. The two choices are Lustre or IBM Spectrum Scale (GPFS). |
| Server node count (StorageNodeCount) | 2 | The number of server instances you want to provision, if you’re using IBM Spectrum Scale for storage. You can select 2-64 instances. For more information, see the [Planning](#_Planning_Yyour_Ddeployment) section. |
| EBS volume type for storage instances (StorageEBSVolumeType) | gp2 | The EBS volume type you want to use for the storage instances. You can choose General Purpose SSD (gp2), Provisioned IOPS SSD (io1), Cold HDD (sc1), Throughput Optimized HDD (st1), or standard. |
| Storage disk size for each node (StorageEBSVolumeSize) | 500 | The EBS volume size (in GiB) for storage nodes. You can specify 100-1638400 GiB, in multiples of 100. |
| Storage key pair name (StorageKeyPairName) | *Requires input* | A public/private key pair for storage instances. |

*Personal configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Operator email (OperatorEmail) | *Requires input* | The email address that notifications of any scaling operations should be sent to. |

*License information:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| License agreement terms (LicenseAgreementTerms) | *Requires input* | Review the [licensing terms for IBM Spectrum Scale](http://spectrumscale-license.s3-website-us-west-2.amazonaws.com/), and then choose **Accept** to indicate your acceptance. |

*Lustre EC2 configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Number of Lustre OSS nodes (NumberOfOSSNodes) | 3 | The number of Lustre OSS nodes you want to deploy. You can specify 3-15 nodes. |
| EBS volume size for Lustre OSS (LustreOSSEBSVolumeSize) | 100 | EBS volume size (in GiB) for Lustre OSS nodes. You can specify 1-9999 GiB, in multiples of 100. |

*SAS grid EC2 configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Instance type for SAS Grid instances (SASGridInstanceType) | i3.8xlarge | The EC2 instance type you want to use for the SAS grid instances. |
| Required number of SAS Grid instances (NoOfSASGridInstances) | 2 | The number of SAS grid instances you want to provision. You can select 2-20 instances, depending on the number of cores you are licensed for. For example, if you are licensed for 32 cores, you can keep the default setting of 2 in the text field.  For more information, see the [Planning](#_Planning_Yyour_Ddeployment) section. |
| SAS Grid key pair Name (SASGridKeyPairName) | *Requires input* | A public/private key pair, which allows you to securely connect to your SAS grid instances after launch. |

*AWS Quick Start configuration:*

|  |  |  |
| --- | --- | --- |
| Parameter label (name) | Default | Description |
| Quick Start S3 Bucket Name (QSS3BucketName) | aws-quickstart | The S3 bucket you have created for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. The bucket name can include numbers, lowercase letters, uppercase letters, and hyphens, but should not start or end with a hyphen. |
| Quick Start S3 Key Prefix (QSS3KeyPrefix) | quickstart-sas-grid/ | The [S3 key name prefix](https://docs.aws.amazon.com/AmazonS3/latest/dev/UsingMetadata.html) used to simulate a folder for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. This prefix can include numbers, lowercase letters, uppercase letters, hyphens, and forward slashes. |

### Step 5. Validate Your Deployment

If you selected IBM Spectrum Scale (GPFS) for the storage stack, follow the instructions in step 5 of the [IBM Spectrum Scale deployment guide](https://fwd.aws/rxqbD).

If you selected Lustre for the storage stack, follow these steps:

1. Check if disks are mounted on the following mount points on Lustre instances by using the df -h command.

* /lsf\_config on MGT and MDT nodes
* /lsf\_data on OSS nodes

1. Check if disks are mounted on the following SAS grid instances by using the df -h command.

* /sas on Metadata, mid-tier and SAS grid nodes
* /saswork on SAS grid nodes

The command output should include references to these disks. If it doesn’t, see the [FAQ](#_FAQ) for instructions on mounting the disks manually.

1. Validate your SAS grid installation by following the instructions on the [SAS website](https://support.sas.com/rnd/scalability/grid/gridinstall.html):.

* Choose the link for [Program to Verify Grid Environment Setup](https://support.sas.com/rnd/scalability/grid/gridtest.pdf).
* Save the file as gridtest.sas.
* Edit the gridtest.sas file and remove the title at the top as well as the page breaks.
* On any SAS grid node, submit the job as a valid SAS Metadata user (for example, sasdemo):

sas gridtest.sas

* Check the log file:
  + If successful, the log lists the number of grid nodes and reports a call to each node as successful.
  + If unsuccessful, the log reports “Sign on failed to node X” or “grid not enabled.”

### Step 6. Run SAS Post-Deployment Scripts

To save on costs, use the utility scripts to stop or start EC2 instances on nights or weekends. The following utility scripts are in the /tmp directory of the Linux bastion host.

* stop\_instances.sh – This script will shut down the Lustre and SAS grid nodes.  The Remote Desktop Gateway and Linux bastion host instances will remain up and running.
* start\_instances.sh – This script will start up the instances in the correct order, along with the Lustre and SAS services.

## FAQ

**Q.** I encountered a CREATE\_FAILED error when I launched the Quick Start.

**A.** If AWS CloudFormation fails to create the stack, we recommend that you relaunch the template with **Rollback on failure** set to **No**. (This setting is under **Advanced** in the AWS CloudFormation console, **Options** page.) With this setting, the stack’s state will be retained and the instance will be left running, so you can troubleshoot the issue. (You'll want to look at the log files in %ProgramFiles%\Amazon\EC2ConfigService and C:\cfn\log.)

**Important** When you set **Rollback on failure** to **No**, you’ll continue to incur AWS charges for this stack. Please make sure to delete the stack when you’ve finished troubleshooting.

For additional information, see [Troubleshooting AWS CloudFormation](https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/troubleshooting.html) on the AWS website.

**Q.** I encountered a size limitation error when I deployed the AWS Cloudformation templates.

**A.** We recommend that you launch the Quick Start templates from the location we’ve provided or from another S3 bucket. If you deploy the templates from a local copy on your computer or from a non-S3 location, you might encounter template size limitations when you create the stack. For more information about AWS CloudFormation limits, see the [AWS documentation](http://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/cloudformation-limits.html).

**Q.** I can’t find disk references in the output provided by the df –h command. How can I mount these disks?

**A.** You can mount the disks manually.

* For /lsf\_config, make sure that /lsf\_config is mounted on the MGT node, and run the following command in the MDT node to mount the disk manually:

# mount /lsf\_config

* For /lsf\_data, make sure that /lsf\_config is mounted on the MGT and MDT nodes, and run the following command in any of the OSS nodes to mount the disk manually:

# mount /lsf\_data

* For /sas, make sure that /lsf\_config and /lsf\_data are mounted on the MGT/MDT and OSS nodes, respectively. Run this command to mount /sas:

# /usr/local/bin/lustre\_mount

This will mount the Lustre file system.

* For /saswork, check if /dev/md0 is configured properly by using the command  
   disk "/dev/xvd[b-i]". Run one of these commands to mount /dev/md0 on /saswork:

# mount /saswork

 or:

# mount /dev/md0 /saswork

**Q.** What should I do if my SAS grid instances fail?

**A.** The Quick Start automatically sets up a VPC with two Availability Zones, but it deploys SAS grid instances into only one of these zones. In the case of a disruption or failure, you can use the [Quick Start workload template](https://fwd.aws/nAR63) to deploy the SAS grid and Lustre components into private subnets in the second Availability Zone, and restore the application data and metadata using backups. Alternatively, you can set up a separate, parallel VPC environment using the [Quick Start master template](https://fwd.aws/kJ9kP), and keep it on standby in case of a failure.

**Q.** What should I do if my CloudFormation stack fails with the error, “Your stack already exists”?

**A.** You cannot have more than one stack with the same name in the same AWS Region. If you stack fails for some reason, please delete the failed stack and then deploy the Quick Start again.

Make sure to delete any failed stacks in an AWS Region before re-deploying the Quick Start.

## Additional Resources

**AWS services**

* Amazon EC2  
  <https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/>
* AWS CloudFormation  
  <https://docs.aws.amazon.com/cloudformation/>
* Amazon VPC  
  <https://docs.aws.amazon.com/vpc/>
* Amazon EBS  
  <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonEBS.html>
* Auto Scaling  
  <https://docs.aws.amazon.com/autoscaling/latest/userguide/>

**SAS Grid**

* SAS grid documentation  
  <http://support.sas.com/software/products/gridmgr/>
* SAS grid installation  
  <https://support.sas.com/rnd/scalability/grid/gridinstall.html>

**Lustre**

* Lustre documentation  
  <http://wiki.old.lustre.org/manual/LustreManual20_HTML/UnderstandingLustre.html#50438250_19307>

**IBM Spectrum Scale**

* IBM Spectrum Scale evaluations <https://developer.ibm.com/storage/products/ibmspectrum-scale/#evaluate>
* IBM Spectrum Scale introduction: [https://www.ibm.com/support/knowledgecenter/en/STXKQY\_4.2.0/com.ibm.spectru m.scale.v4r2.ins.doc/bl1in\_IntroducingIBMSpectrumScale.htm](https://www.ibm.com/support/knowledgecenter/en/STXKQY_4.2.0/com.ibm.spectru%20m.scale.v4r2.ins.doc/bl1in_IntroducingIBMSpectrumScale.htm)
* IBM Spectrum Scale architecture: [https://www.ibm.com/support/knowledgecenter/en/STXKQY\_4.2.0/com.ibm.spectru m.scale.v4r2.ins.doc/bl1ins\_architr.htm](https://www.ibm.com/support/knowledgecenter/en/STXKQY_4.2.0/com.ibm.spectru%20m.scale.v4r2.ins.doc/bl1ins_architr.htm)
* IBM Spectrum Scale on AWS evaluation survey <https://www.surveygizmo.com/s3/3795938/IBM-Spectrum-Scale-on-AWS>
* IBM Spectrum Scale forum: <https://www.ibm.com/developerworks/community/forums/html/forum?id=11111111-0000-0000-0000-000000000479> (or you can email scale@us.ibm.com for support)

**Quick Start reference deployments**

* AWS Quick Start home page  
  <https://aws.amazon.com/quickstart/>

## Send Us Feedback

You can visit our [GitHub repository](https://fwd.aws/y5J6J) to download the templates and scripts for this Quick Start, to post your comments, and to share your customizations with others.

## Document Revisions

|  |  |  |
| --- | --- | --- |
| Date | Change | In sections |
| June 2020 | Changed Title.  Changed SAS Grid to SAS grid.  Updated contact email to support@sas.com | On title page  Changes in template and guide  Changed in licensing section and steps. |
| March 2019 | Added IBM Spectrum Scale (GPFS) as a storage stack option | Changes throughout templates and guide |
| December 2018 | Added SAS Grid nodes to the deployment | Changes throughout templates and guide |
| August 2017 | Initial publication | — |

© 2019, Amazon Web Services, Inc. or its affiliates, and Core Compete. All rights reserved.

**Notices**

This document is provided for informational purposes only. It represents AWS’s current product offerings and practices as of the date of issue of this document, which are subject to change without notice. Customers are responsible for making their own independent assessment of the information in this document and any use of AWS’s products or services, each of which is provided “as is” without warranty of any kind, whether express or implied. This document does not create any warranties, representations, contractual commitments, conditions or assurances from AWS, its affiliates, suppliers or licensors. The responsibilities and liabilities of AWS to its customers are controlled by AWS agreements, and this document is not part of, nor does it modify, any agreement between AWS and its customers.

The software included with this paper is licensed under the Apache License, Version 2.0 (the "License"). You may not use this file except in compliance with the License. A copy of the License is located at <http://aws.amazon.com/apache2.0/> or in the "license" file accompanying this file. This code is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.