DataSunrise with Amazon Redshift on the AWS Cloud

Quick Start Reference Deployment

November 2019
Last update: February 2020 (revisions)

Eduardo D. Benzecri, Radik Chumaren, DataSunrise, Inc.
Saunak Chandra, Sr. Solutions Architect, AWS
Tony Bulding, Sr. Solutions Architect, AWS

Visit our GitHub repository for source files and to post feedback, report bugs, or submit feature ideas for this Quick Start.

Contents

This Quick Start was created by DataSunrise, Inc. in collaboration with Amazon Web Services (AWS).

Quick Starts are automated reference deployments that use AWS CloudFormation templates to deploy key technologies on AWS, following AWS best practices.

Overview

This Quick Start reference deployment guide provides step-by-step instructions for deploying DataSunrise on the AWS Cloud.

This Quick Start is for users who want to protect access to the data in an Amazon Redshift cluster by using DataSunrise deployed as a high availability (HA) cluster.

DataSunrise on AWS

DataSunrise Database Security Suite is cross-platform, high-performance software that secures databases and data in real time. DataSunrise helps to protect companies' sensitive data from outside threats and internal security breaches.

DataSunrise supports databases and data warehouses that include Amazon Aurora MySQL, Amazon Aurora PostgreSQL, Amazon Redshift, Amazon DynamoDB, MariaDB, MongoDB, MySQL, PostgreSQL, Apache Hive, and others.

DataSunrise features components such as an intelligent Database Firewall, Database Activity Monitoring, Static Data Masking and Dynamic Data Masking capabilities, and Sensitive Data Discovery, among others.

DataSunrise attests that DataSunrise can be used in compliance with database-specific regulations, such as SOX, PCI DSS, GDPR, and HIPAA. DataSunrise helps you monitor, track, and report on database activity to meet compliance requirements.

Running DataSunrise on AWS provides many benefits. For example, because of the proxy-based nature of DataSunrise, database activity monitoring of databases that are deployed on AWS (both SQL and no-SQL) is fully supported. DataSunrise Database Firewall helps protect not only Amazon databases such as Amazon Redshift, but also Amazon Simple Storage Service (Amazon S3) buckets. Also, DataSunrise can be integrated with Amazon CloudWatch.

Cost 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.

> **Tip** After you deploy the Quick Start, we recommend that you enable the AWS Cost and Usage Report to track costs associated with the Quick Start. This report delivers billing metrics to an Amazon Simple Storage Service (Amazon S3) bucket in your account. It provides cost estimates based on usage throughout each month and finalizes the data at the end of the month. For more information about the report, see the AWS documentation.

This Quick Start requires a license for DataSunrise. To use the Quick Start in your production environment, sign up for a license and get a key.

If you're downloading DataSunrise for the first time, you can get a two-week trial license. The license is included in an email you receive when you download the product. If the license has expired, all DataSunrise rules are disabled. In this case, all user queries are sent directly to the database, bypassing the DataSunrise proxy.

Architecture

Deploying this Quick Start for a new virtual private cloud (VPC) with **default parameters** builds the following DataSunrise environment in the AWS Cloud.

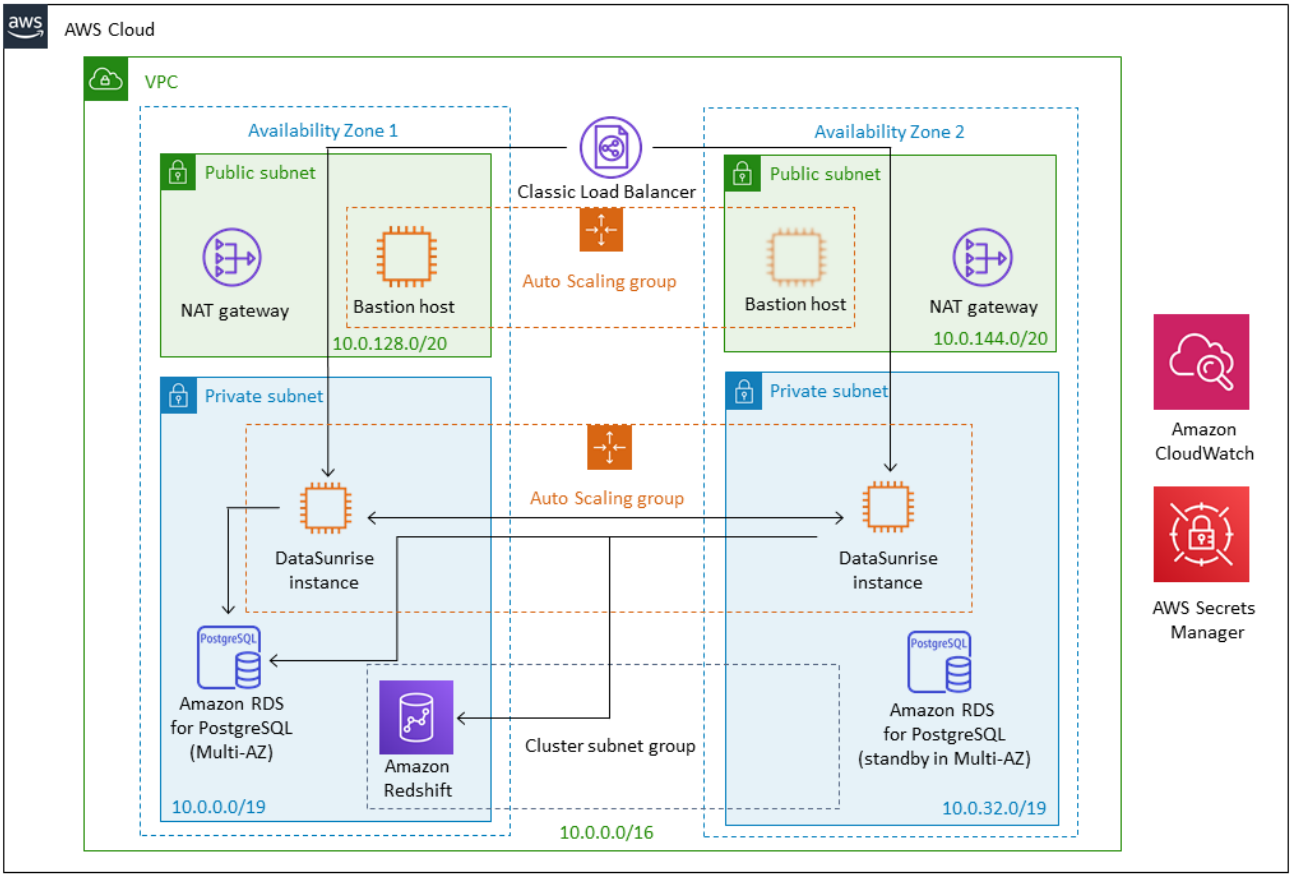


Figure 1: Quick Start architecture for DataSunrise on AWS

The Quick Start sets up the following environment:

- A highly available architecture that spans two Availability Zones.*
- A VPC configured with public and private subnets according to AWS best practices, to provide you with your own virtual network on AWS.*
- A Classic Load Balancer configured in the public subnets as the connection endpoint, to serve incoming database traffic to an available and healthy DataSunrise instance.
- In the public subnets:
  - Managed network address translation (NAT) gateways to allow outbound internet access for resources in the private subnets.*

  - A Linux bastion host in an Auto Scaling group to allow inbound Secure Shell (SSH) access to Amazon Elastic Compute Cloud (Amazon EC2) instances in public and private subnets.*

- In the private subnets:
  - A DataSunrise host in an Auto Scaling group to allow an inbound database connection via the Classic Load Balancer and SSH access from a Linux bastion host.
  - Two Amazon Relational Database Service (Amazon RDS) for PostgreSQL instances to store DataSunrise configuration and audit data.
  - An Amazon Redshift cluster inside a subnet group. The subnet group spans two subnets to enable restoring from a snapshot, in case a disaster occurs in the Region.
- Amazon CloudWatch connected to DataSunrise instances to receive custom metrics.
- AWS Secrets Manager to store sensitive DataSunrise data.

* The template that deploys the Quick Start into an existing VPC skips the components marked by asterisks and prompts you for your existing VPC configuration.

Planning the deployment

Specialized knowledge

This Quick Start assumes familiarity with basic networking concepts.

This deployment guide also requires a moderate level of familiarity with AWS services. If you're new to AWS, visit the Getting Started Resource Center and the AWS Training and Certification website for materials and programs that can help you develop the skills to design, deploy, and operate your infrastructure and applications on the AWS Cloud.

AWS account

If you don't already have an AWS account, create one at https://aws.amazon.com by following the on-screen instructions. Part of the sign-up process involves receiving a phone call and entering a PIN using the phone keypad.

Your AWS account is automatically signed up for all AWS services. You are charged only for the services you use.

Technical requirements

Before you launch the Quick Start, your account must be configured as specified in the following table. Otherwise, deployment might fail.

Resources	If necessary, request service limit increases for the following resources. You might need to do this if you already have an existing deployment that uses these resources, and you think you might exceed the default limits with this deployment. For default limits, see the AWS documentation. AWS Trusted Advisor offers a service limits check that displays your usage and limits for some aspects of some services.

Resource	This deployment uses
VPCs	**1**
Elastic IP addresses	**1**
EC2 security groups	**6**
AWS Identity and Access Management (IAM) roles	**3**
Auto Scaling groups	**2**
Classic Load Balancers	**1**
EC2 instances	**>=2**
Amazon RDS for PostgreSQL instances	**2**
Secrets	**6**

Key pair	Make sure that at least one Amazon EC2 key pair exists in your AWS account in the Region where you are planning to deploy the Quick Start. Make note of the key pair name. You'll be prompted for this information during deployment. To create a key pair, follow the instructions in the AWS documentation. If you're deploying the Quick Start for testing or proof-of-concept purposes, we recommend that you create a new key pair instead of specifying a key pair that's already being used by a production instance.
IAM permissions	To deploy the Quick Start, you must log in to the AWS Management Console with IAM permissions for the resources and actions the templates will deploy. The *AdministratorAccess* managed policy within IAM provides sufficient permissions, although your organization may choose to use a custom policy with more restrictions.

Deployment options

This Quick Start provides two deployment options:

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

The Quick Start provides separate templates for these options. It also lets you configure CIDR blocks, instance types, and DataSunrise settings, as discussed later in this guide.

Deployment steps

Step 1. Sign in to your AWS account

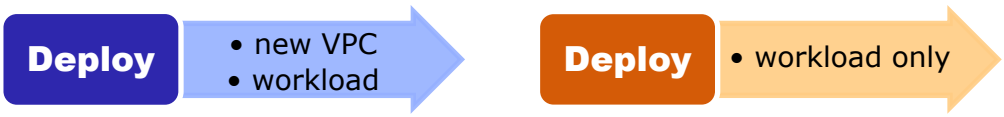
1. Sign in to your AWS account at https://aws.amazon.com with an IAM user role that has the necessary permissions. For details, see Planning the deployment earlier in this guide.
2. Make sure that your AWS account is configured correctly, as discussed in the Technical requirements section.

Step 2. Launch the Quick Start

Notes The instructions in this section reflect the older version of the AWS CloudFormation console. If you're using the redesigned console, some of the user interface elements might be different.

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.

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



Deploy DataSunrise into a new VPC on AWS

Deploy DataSunrise into an existing VPC on AWS

Important If you're deploying DataSunrise into an existing VPC, make sure that your VPC has two private subnets in different Availability Zones for the workload instances and database instances, and that the subnets aren't shared. This Quick Start doesn't support shared subnets. These subnets require NAT gateways in their route tables, to allow the instances to download packages and software without exposing them to the internet. You will also need the domain name option configured in the DHCP options as explained in the Amazon VPC documentation. You will be prompted for your VPC settings when you launch the Quick Start.

Each deployment takes about 20 minutes to complete.

2. Check the AWS Region that's displayed in the upper-right corner of the navigation bar, and change it if necessary. This is where the network infrastructure for DataSunrise will be built. The template is launched in the US East (Ohio) Region by default.

3. On the **Select Template** page, keep the default setting for the template URL, and then choose **Next**.
4. 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.

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

   - Parameters for deploying DataSunrise into a new VPC
   - Parameters for deploying DataSunrise into an existing VPC

   When you finish reviewing and customizing the parameters, choose **Next**.

OPTION 1: PARAMETERS FOR DEPLOYING DATASUNRISE INTO A NEW VPC

View template

VPC network configuration:

Parameter label (name)	Default	Description
Availability Zones (`VPCAvailabilityZones`)	*Requires input*	Two Availability Zones to be used for the subnets in the VPC. The logical order will be preserved.
VPC CIDR (`VPCCIDR`)	10.0.0.0/16	The CIDR block for the VPC.
Private subnet 1A CIDR (`VPCPrivateSubnet1 CIDR`)	10.0.0.0/19	The CIDR block for the private subnet 1 located in Availability Zone 1.
Private subnet 2A CIDR (`VPCPrivateSubnet2 CIDR`)	10.0.32.0/19	The CIDR block for the private subnet 2 located in Availability Zone 2.
Public subnet 1 CIDR (`VPCPublicSubnet1CIDR`)	10.0.128.0/20	The CIDR block for the public DMZ subnet 1 located in Availability Zone 2.
Public subnet 2 CIDR (`VPCPublicSubnet2CIDR`)	10.0.144.0/20	The CIDR block for the public DMZ subnet 2 located in Availability Zone 2.
VPC tenancy (`VPCTenancy`)	default	Allowed tenancy of instances launched into the VPC.
Allowed external access CIDR (`VPCRemoteAccessCIDR`)	*Requires input*	Allowed CIDR block for external access. If you want to allow external access from everywhere, use 0.0.0.0/0

Linux bastion - basic configuration:

Parameter label (name)	Default	Description
Bastion key pair (`LinuxBastionKeyPair`)	*Requires input*	The EC2 key pair to be attached.

Amazon Redshift - basic configuration:

Parameter label (name)	Default	Description
Node type for Redshift cluster (`RedshiftNodeType`)	dc2.large	Type of node to be provisioned.
Number of nodes in Redshift cluster (`RedshiftNodesNumber`)	2	The number of nodes in the cluster. For multinode clusters, this parameter's value should be greater than 1.
Redshift database name (`RedshiftDBName`)	rsdsqsdev01	The name of the first database to be created when the cluster is deployed.
Redshift master user name (`RedshiftMasterUserName`)	rsadmin	The user name associated with the master user account for the cluster to be deployed.
Redshift master user password (`RedshiftMasterUserPassword`)	*Requires input*	The user password associated with the master user account for the cluster to be deployed.

Amazon Redshift - tag identifiers:

Parameter label (name)	Default	Description
Environment (`RedshiftTagEnvironment`)	none	The environment tag used to designate the environment stage of the cluster.
Unique friendly name (`RedshiftTagName`)	rsqs-DataSunrise	The unique friendly name required by your company's tagging strategy document, which will be added to the environment tag.
Project cost center (`RedshiftTagProjectCostCenter`)	12345	The cost center associated with the cluster.
Confidentiality classifier	—	The confidentiality classification of the data associated with the cluster.

Parameter label (name)	Default	Description
(RedshiftTag Confidentiality)		
Compliance classifier (RedshiftTag Compliance)	—	The compliance level for the cluster.

Amazon Redshift - advanced configuration:

Parameter label (name)	Default	Description
Redshift cluster port (RedshiftClusterPort)	8200	The port number to be used by the cluster for incoming connections.
Enable Redshift logging to S3 (RedshiftEnable LoggingToS3)	false	Enables (true) or disables (false) logging to an S3 bucket.
Maximum number of concurrent clusters (RedshiftMaximum ConcurrentCluster)	1	The maximum number of concurrency scaling Amazon Redshift clusters.
Encryption at rest (RedshiftEncryptionAt Rest)	false	Enables (true) or disables (false) encryption at rest of the Amazon Redshift database.
KMS key ID (RedshiftKMSKey)	*Optional*	(Optional if RedshiftEncryptionAtRest = false) The confidentiality classification of the data associated with the cluster.
Maintenance window (RedshiftMaintenance Window)	sat:05:00-sat:05:30	The maintenance window for the Amazon Redshift cluster.
Amazon S3 bucket for Redshift IAM role (RedshiftS3BucketFor RedshiftIAMRole)	*Optional*	(Optional) The existing Amazon S3 bucket. An IAM role will be created and associated with the Amazon Redshift cluster with GET and LIST access to this bucket.
Redshift glue catalog DB (RedshiftGlueCatalog DB)	*Optional*	(Optional) The name of your AWS Glue Data Catalog database.

DataSunrise - basic cluster configuration:

Parameter label (name)	Default	Description
Cluster member minimum size (`DSClusterMinimumSize`)	1	The minimum size of the DataSunrise cluster.
Cluster member maximum size (`DSClusterMaximumSize`)	3	The maximum size of the DataSunrise cluster.
Cluster member instance type (`DSClusterMember InstanceType`)	t2.micro	The EC2 instance type to be used by a cluster member. By default, t2.micro is free-tier eligible.
Cluster member key pair (`DSClusterMemberKey Pair`)	*Requires input*	The EC2 key pair to be attached.
DataSunrise license (`DSClusterLicense`)	yourdatasunrise licensehere	The license string. If you don't have a license key, go to: https://www.datasunrise.com/activation-key/?Type=AWS&Os=linux

DataSunrise cluster - dictionary configuration:

Parameter label (name)	Default	Description
Dictionary database name (`DSDictionaryDBName`)	dsdictionary	The DataSunrise Dictionary is used to store configurations. Must begin with a letter and contain only alphanumeric characters.
Dictionary database username (`DSDictionaryDBUser Name`)	dsuser	The database user name with administrator privileges. Should begin with a letter and contain only alphanumeric characters.
Dictionary database size (`DSDictionaryDB StorageSize`)	20	The Dictionary database size (GB).
Dictionary database instance type (`DSDictionaryDBClass`)	db.t2.micro	The Amazon RDS instance type for the Dictionary database.

DataSunrise cluster - audit configuration:

Parameter label (name)	Default	Description
Audit database name (DSAuditDBName)	dsaudit	The DataSunrise Audit Storage database used to store audit logs. Should begin with a letter and contain only alphanumeric characters.
Audit database username (DSAuditDBUserName)	dsauser	The database user name with administrator privileges. Should begin with a letter and contain only alphanumeric characters.
Audit database size (DSAuditDBStorageSize)	200	Audit Storage database size (GB).
Audit database instance type (DSAuditDBClass)	db.t2.micro	The Amazon RDS instance type for the Audit Storage database.

DataSunrise - advanced cluster configuration:

Parameter label (name)	Default	Description
Average CPU utilization (DSAutoScalingPolicy AverageCPUBusy)	50	The average CPU utilization (%).
Health check target (DSLoadBalancerHealth CheckTarget)	HTTPS:11000/ healthcheck/ general	Specifies which health check type to use.

DataSunrise: user configuration

Parameter label (name)	Default	Description
Admin email (DSAdminEmail)	admin-email-here@example.domain	The administrator email address.
Administrator password (DSAdminPassword)	*Requires input*	The administrator password. Must contain 8 to 41 printable ASCII characters (excluding / @ " ' \).
User name (DSUserName)	mydatasunriseusernamehere	Regular DataSunrise user. Must begin with a letter and contain only alphanumeric characters.
User email (DSUserEmail)	user-email-here@example.domain	The user email address.

Parameter label (name)	Default	Description
User password (DSUserPassword)	*Requires input*	The user password. Must contain 8 to 41 printable ASCII characters (excluding: /, @, “, ‘ \)

AWS Quick Start configuration:

Note We recommend that you keep the default settings for the following two parameters, unless you are customizing the Quick Start templates for your own deployment projects. Changing the settings of these parameters will automatically update code references to point to a new Quick Start location. For additional details, see the AWS Quick Start Contributor’s Guide.

Parameter label (name)	Default	Description
Quick Start S3 bucket name (QSS3BucketName)	aws-quickstart	The S3 bucket you 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-datasunrise/	The S3 key prefix for the Quick Start assets. AWS Quick Start key prefix can include numbers, lowercase letters, uppercase letters, hyphens (-), and forward slashes (/). It cannot start or end with a forward slash (/) because they are automatically appended.

OPTION 2: PARAMETERS FOR DEPLOYING DATASUNRISE INTO AN EXISTING VPC

View template

VPC network configuration:

Parameter label (name)	Default	Description
VPC ID (VPCID)	*Requires input*	The ID of your existing VPC you want to use for deploying DataSunrise into.
Private subnet 1 ID (VPCPrivateSubnet1ID)	*Requires input*	The ID of the private subnet in Availability Zone 1 in your existing VPC.
Private subnet 2 ID (VPCPrivateSubnet2ID)	*Requires input*	The ID of the private subnet in Availability Zone 2 in your existing VPC.
Public subnet 1 ID (VPCPublicSubnet1ID)	*Requires input*	The ID of the public subnet in Availability Zone 1 in your existing VPC.

Parameter label (name)	Default	Description
Public subnet 2 ID (VPCPublicSubnet2ID)	*Requires input*	The ID of the public subnet in Availability Zone 2 in your existing VPC.
Allowed external access CIDR (VPCRemoteAccessCIDR)	*Requires input*	The allowed CIDR block for external access. If you want to allow external access from everywhere, use 0.0.0.0/0

Linux bastion - basic configuration:

Parameter label (name)	Default	Description
Bastion key pair (LinuxBastionKeyPair)	*Requires input*	The EC2 key pair to be attached.

Amazon Redshift - basic configuration:

Parameter label (name)	Default	Description
Node type for the Redshift cluster (RedshiftNodeType)	dc2.large	The type of node to be provisioned.
Number of nodes in the Redshift cluster (RedshiftNodesNumber)	2	The number of nodes in the cluster. For multinode clusters, this parameter's value should be greater than 1.
Redshift database name (RedshiftDBName)	rsdsqsdev01	The name of the first database to be created when the cluster is deployed.
Redshift master user name (RedshiftMasterUserName)	rsadmin	The user name associated with the master user account for the cluster to be deployed.
Redshift master user password (RedshiftMasterUserPassword)	*Requires input*	The user password associated with the master user account for the cluster to be deployed. Must have at least 8 characters and no more than 64 characters. Must include 1 uppercase letter, 1 lowercase letter, 1 number and 1 symbol (excluding /, @, \, ').

Amazon Redshift - tag identifiers:

Parameter label (name)	Default	Description
Environment (RedshiftTag Environment)	none	The environment tag used to designate the environment stage of the cluster.
Unique friendly name (RedshiftTagName)	rsqs-DataSunrise	The unique friendly name required by your company's tagging strategy document, and which will be added to the Environment tag.
Project cost center (RedshiftTagProject CostCenter)	12345	The cost center associated with the cluster.
Confidentiality classifier (RedshiftTag Confidentiality)	—	The confidentiality classification of the data associated with the cluster.
Compliance classifier (RedshiftTag Compliance)	—	The compliance level for the cluster.

Amazon Redshift - advanced configuration:

Parameter label (name)	Default	Description
Redshift cluster port (RedshiftClusterPort)	8200	The port number to be used by the cluster for incoming connections.
Enable Redshift logging to S3 (RedshiftEnable LoggingToS3)	false	Enables (true) or disables (false) logging to an S3 bucket.
Maximum number of concurrent clusters (RedshiftMaximum ConcurrentCluster)	1	The maximum number of concurrency scaling Amazon Redshift clusters.
Encryption at rest (RedshiftEncryptionAt Rest)	false	Enables or disables encryption at rest of the Amazon Redshift database.
KMS key ID (RedshiftKMSKey)	*Optional*	(Optional if RedshiftEncryptionAtRest = false) Confidentiality classification of the data associated with the cluster.
Maintenance window (RedshiftMaintenance Window)	sat:05:00-sat:05:30	Maintenance window for the Amazon Redshift cluster.

Parameter label (name)	Default	Description
Amazon S3 bucket for Redshift IAM role (RedshiftS3BucketForRedshiftIAMRole)	*Optional*	(Optional) Existing Amazon S3 bucket. An IAM role will be created and associated with the Amazon Redshift cluster with GET and LIST access to this bucket.
Redshift glue catalog DB (RedshiftGlueCatalogDB)	*Optional*	(Optional) The name of your AWS Glue Data Catalog database.

DataSunrise - basic cluster configuration:

Parameter label (name)	Default	Description
Cluster member minimum size (DSClusterMinimumSize)	1	The minimum size of the DataSunrise cluster.
Cluster member maximum size (DSClusterMaximumSize)	3	The maximum size of the DataSunrise cluster.
Cluster member instance type (DSClusterMemberInstanceType)	t2.micro	The EC2 instance type to be used by a cluster member. By default, t2.micro is free-tier eligible.
Cluster member key pair (DSClusterMemberKeyPair)	*Requires input*	The EC2 key pair to be attached.
DataSunrise license (DSClusterLicense)	yourdatasunrise licensehere	The license string. If you don't have a license key, go to https://www.datasunrise.com/activation-key/?Type=AWS&Os=linux

DataSunrise cluster - dictionary configuration:

Parameter label (name)	Default	Description
Dictionary database name (DSDictionaryDBName)	dsdictionary	The DataSunrise Dictionary is used to store configurations. Must begin with a letter and contain only alphanumeric characters.
Dictionary database username (DSDictionaryDBUserName)	dsuser	The database user name with administrator privileges. Should begin with a letter and contain only alphanumeric characters

Parameter label (name)	Default	Description
Dictionary database size (`DSDictionaryDBStorageSize`)	20	The Dictionary database size (GB)
Dictionary database instance type (`DSDictionaryDBClass`)	db.t2.micro	The Amazon RDS instance type for the Dictionary database

DataSunrise cluster - audit configuration:

Parameter label (name)	Default	Description
Audit database name (`DSAuditDBName`)	dsaudit	The DataSunrise Audit Storage database used to store audit logs. Should begin with a letter and contain only alphanumeric characters.
Audit database username (`DSAuditDBUserName`)	dsauser	The database user name with administrator privileges. Should begin with a letter and contain only alphanumeric characters.
Audit database size (`DSAuditDBStorageSize`)	200	The Audit Storage database size (GB).
Audit database instance type (`DSAuditDBClass`)	db.t2.micro	The Amazon RDS instance type for the Audit Storage database.

DataSunrise - advanced cluster configuration:

Parameter label (name)	Default	Description
Average CPU utilization (`DSAutoScalingPolicyAverageCPUBusy`)	50	The average CPU utilization.
Healthcheck target (`DSLoadBalancerHealthCheckTarget`)	HTTPS:11000/healthcheck/general	Specifies which health check type to use.

DataSunrise - user configuration:

Parameter label (name)	Default	Description
Admin email (DSAdminEmail)	admin-email-here@example.domain	The administrator email address.
Administrator password (DSAdminPassword)	*Requires input*	The administrator password. Must contain 8 to 41 printable ASCII characters (excluding / @ " ' \)
User name (DSUserName)	mydatasunriseusernamehere	Regular DataSunrise user. Must begin with a letter and contain only alphanumeric characters
User email (DSUserEmail)	user-email-here@example.domain	The user email address
User password (DSUserPassword)	*Requires input*	The user password. Must contain 8 to 41 printable ASCII characters (excluding: /, @, “, ‘, \)

AWS Quick Start configuration:

Note We recommend that you keep the default settings for the following two parameters, unless you are customizing the Quick Start templates for your own deployment projects. Changing the settings of these parameters will automatically update code references to point to a new Quick Start location. For additional details, see the AWS Quick Start Contributor’s Guide.

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-datasunrise/	The S3 key name prefix 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.

5. On the **Options** page, you can specify tags (key-value pairs) for resources in your stack and set advanced options. When you’re done, choose **Next**.

6. On the **Review** page, review and confirm the template settings. Under **Capabilities**, select the two check boxes to acknowledge that the template will create IAM resources and that it might require the capability to auto-expand macros.
7. Choose **Create** to deploy the stack.
8. Monitor the status of the stack. When the status is **CREATE_COMPLETE**, the DataSunrise cluster is ready.
9. Use the URLs displayed in the **Outputs** tab for the stack to view the resources that were created.

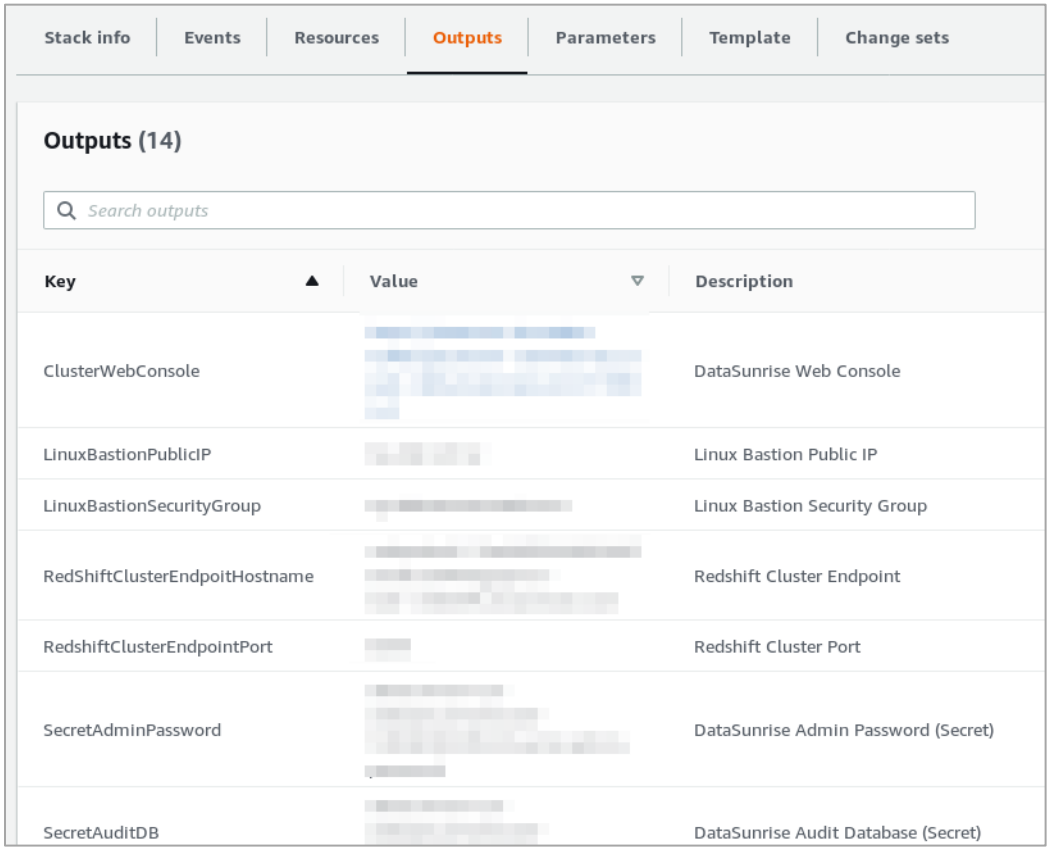


Figure 2: DataSunrise outputs after successful deployment

Step 3. Test the deployment

The easiest way to test the deployment is using the web console. After the stack is successfully deployed, you can see the web console's URL. Just use DataSunrise credentials created in the process. If you get access to the web console, take a quick look at the dashboard, and check if there's any error. Also, you can check DataSunrise logs.

Another way to check if the deployment is working is to connect via SSH to a Linux bastion host and, from there, connect via SSH to any DataSunrise instance. You can check the logs created in /tmp because they will provide information related to the installation and configuration process.

Best practices for using DataSunrise with Amazon Redshift on AWS

Best practices include:

- Use two EC2 key pairs: one for the Linux bastion host and the other for the DataSunrise cluster.
- At least once per month, check if there any updated versions of DataSunrise available.
- If you want to update DataSunrise to the latest version available, create a configuration backup before proceeding.

For more information:

- Download the general DataSunrise best practices guide at https://www.datasunrise.com/download-the-datasunrise-security-best-practices/
- Download the DataSunrise AWS best practices guide at https://www.datasunrise.com/download-the-datasunrise-aws-security-best-practices/
- See Frequently Asked Questions at https://www.datasunrise.com/documentation/faq/

Security

Keep in mind the following security guidelines:

- When you allow external access to the Linux bastion and the DataSunrise cluster, avoid using "0.0.0.0/0." If you want to allow "0.0.0.0/0" temporarily, add it manually to each related security group.
- Each password is stored inside an AWS secret.
- Keep the operating system of each EC2 instance updated.

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. For Linux, look at `/var/log/cfn-init.log, /var/log/cloud-init.log, /var/log/cfn-init-output.log and /var/log/cloud-init-output.log`.

> **Important** When you set **Rollback on failure** to **No**, you will continue to incur AWS charges for this stack. Please make sure to delete the stack when you finish troubleshooting.

For additional information, see Troubleshooting AWS CloudFormation 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 links in this guide 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.

Q. I'm trying to access the DataSunrise web console immediately after the deployment is done, but I can't access it. What should I do?

A. Once the stack deployment is completed, you need to wait for 5-10 minutes until DataSunrise is installed and properly configured. This is an average time because we update each EC2 instance used with the latest operating system updates.

Q. I already waited several minutes after the deployment and I still can't connect to the DataSunrise web console. What is the problem?

A. Inability to connect to the web console means that the DataSunrise installation failed. In that case, you must access the Linux bastion host using SSH, create a file with a proper private key, and then connect with SSH to the EC2 instance where the failed installation occurred. Once you've logged in, retrieve all the log files located in /tmp for troubleshooting.

Q. What is the default Administrator user account for DataSunrise?

A. The default administrator user account is `admin`.

Send us feedback

To post feedback, submit feature ideas, or report bugs, use the **Issues** section of the GitHub repository for this Quick Start. If you'd like to submit code, please review the Quick Start Contributor's Guide.

Additional resources

AWS resources

- Getting Started Resource Center
- AWS General Reference
- AWS Glossary

AWS services

- AWS CloudFormation
- Amazon Cloud Watch
- Amazon EC2
- IAM
- Amazon RDS
- Amazon Redshift
- AWS Secrets Manager
- Amazon VPC

DataSunrise documentation

- DataSunrise Administration Guide for Linux
- DataSunrise User Guide
- DataSunrise Database and Data Security and Compliance
- DataSunrise Security

Other Quick Start reference deployments

- AWS Quick Start home page

Document revisions

Date	Change	In sections
February 2020	FAQ item added	FAQ
November 2019	Initial publication	

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.