Using Customer Managed Keys for Encrypting EBS Volumes and RDS on AWS (Preview)

Date published: 2021-12-03
Date modified: 2021-12-17

Legal Notice

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Revision history

2021-12-17 Added support for encrypting RDS.

Introduction

By default, Data Lake and FreeIPA are encrypted using the default key from Amazon's KMS, but you can optionally configure encryption using Customer Managed Keys (CMK) for Amazon Elastic Block Store (EBS) volumes and Relational Database Service (RDS) used by the Data Lake and the FreeIPA.

Amazon offers the option to encrypt EBS volumes and RDS instances using the default key from Amazon's Key Management System (KMS) or using an external customer-managed KMS. By default, Data Lake and FreeIPA are encrypted using the default key from Amazon's KMS present in the region where the environment is running, but you can provide a customer-managed KMS key instead of the default key.

Encryption is configured for block devices and root devices. When encryption is configured for a given cluster, it is automatically applied to all the disk devices of any new VM instances added as a result of cluster scaling or repair.

Environment and Data Hub encryption options

Unlike Data Lake and FreeIPA, Data Hub's EBS volumes are not encrypted by default. If a customer-managed key was provided during environment registration, it is applied to Data Hubs, and, instead or in addition to that, encryption via the default key or customer-managed key can be configured per host group during Data Hub creation.

in order to provide backward compatibility with an existing Data Hub encryption functionality, configuring encryption in CDP works as follows:

- If no encryption key is provided during environment registration, Data Lake and FreeIPA are encrypted with the default KMS and Data Hubs are not encrypted.
- When an encryption key is provided during environment registration, encryption is applied to Data Lake, FreeIPA, and Data Hub cluster's VM instances.
- Once an environment is running and using an encryption key, it is possible to provide a different encryption key during Data Hub creation.
- Or, if no encryption key was provided during environment registration, you can still encrypt a specific Data Hub by providing an encryption key during Data Hub creation.

These scenarios are summarized in the following table:

Encryption key during environment registration	**Encryption key during Data Hub creation**	**Result**
Absent	Absent	● EBS and RDS encryption for Data Lake and FreeIPA is with the default regional encryption key. ● No EBS encryption is configured for Data Hubs running in the environment.
Present	Absent	● EBS and RDS encryption for Data Lake, FreeIPA, and all Data Hubs is with the CMK provided during environment registration.
Present	Present	● EBS and RDS encryption for Data Lake and FreeIPA is with the CMK provided during environment registration. ● If a default key is provided for a Data Hub, then EBS encryption for the Data Hub is with the CMK provided during environment registration. ● If a CMK is provided for a Data Hub, then EBS encryption for the Data Hub is with the CMK provided per host group during the Data Hub's creation.
Absent	Present	● No EBS and RDS encryption is configured for Data Lale and FreeIPA. ● EBS encryption for the specific Data Hub is with the default or CMK provided per host group during Data Hub creation.

Overview of the setup steps

Configuring your environment to use a CMK involves the following steps:

1. Ensure that your provisioning credential has the minimum access permissions.
2. Ensure that your existing encryption key fulfills the required criteria or create a new encryption key according to the instructions provided in this document.
3. When creating an environment, specify the encryption key that should be used for encrypting the environment, including the Data Hubs running in it.

Note: Once your environment is running, if you would like to use a different key for encrypting a specific Data Hub, you can configure it as described in Encryption for Data Hub's EBS volumes on AWS.

Permissions for using encryption

If you are planning to use encryption, ensure that the cross-account IAM role used for the provisioning credential includes the following permissions:

EC2 permissions

```
{
  "Version": "2012-10-17",
  "Statement": {
    "Effect": "Allow",
    "Action": [
      "ec2:CopyImage",
      "ec2:CreateSnapshot",
      "ec2:DeleteSnapshot",
      "ec2:DescribeSnapshots",
      "ec2:CreateVolume",
      "ec2:DeleteVolume",
      "ec2:DescribeVolumes",
      "ec2:DeregisterImage",
    ],
    "Resource": "*"
  }

}
```

KMS permissions

```
{
  "Version": "2012-10-17",
  "Statement": {
    "Effect": "Allow",
    "Action": [
      "kms:DescribeKey",
      "kms:ListKeys",
      "kms:ListAliases"
    ],
    "Resource": "*"
  }

}
```

Encryption key requirements

If planning to use encryption, ensure that your encryption key can be used or create a new encryption key.

For more information on both options, refer to the following documentation:

- Ensuring that an existing encryption key can be used
- Create a new encryption key on AWS

Ensuring that an existing encryption key can be used

If you already have an existing encryption key, make sure that the key fulfills the following requirements.

If you have an existing encryption key that you would like to use with Data Hub, make sure that:

- The following are attached as key user:
  - The *AWSServiceRoleForAutoScaling* built-in role.
  - Your IAM role or IAM user used for the cloud credential.
- To check that these are attached, in the AWS Management Console, navigate to the KMS console > Customer managed keys, select your encryption key, and scroll to Key Users.
- The encryption key is located in the same region where you would like to create clusters with encrypted volumes.

Create a new encryption key on AWS

If you don't have an existing encryption key, use the following instructions to create one.

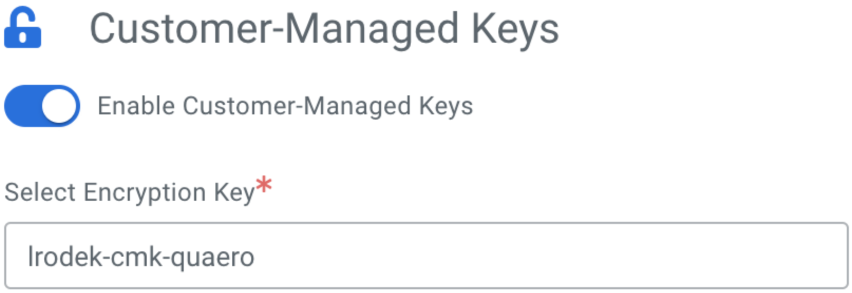
1. In the AWS Management Console, navigate to KMS console.
2. Select Customer managed keys.
3. From the Region dropdown, select the region in which you would like to create and use the encryption key.
4. Click Create key.
5. In Step 1: Configure Key:
    1. Under Key type, choose Symmetric.
    2. Expand Advanced Options and under Key Material Origin, select “KMS” or “External”.
6. In Step 2: Create Alias and Description:
    1. Enter an Alias for your key.
    2. Defining Tags is optional.
7. In Step 3: Define Key Administrative Permissions, select the following:
    1. Choose your own IAM user / role used for logging into the AWS Management Console. Do not set *AWSServiceRoleForAutoScaling* or the cross-account IAM role as the key admin.
8. In Step 4: Define Key Usage Permissions:
    1. Select the *AWSServiceRoleForAutoScaling* built-in role.
    2. Select the cross-account IAM role.
9. In Step 5: Review and edit key policy, you may optionally tweak the key policy as desired, or simply leave it as generated by AWS.
10. Navigate to the last page of the wizard and then click Finish to create an encryption key.

Create an environment with a CMK

You can register your environment as described in Register an AWS environment from CDP UI, just make sure to specify the CMK that should be used to encrypt data, as described in the below steps. Alternatively, you can use Beta CDP CLI.

Steps - CDP UI

1. Log in to the CDP web interface.
2. Navigate to the **Management Console > Environments**, and click **Register environment.**
3. Provide an **Environment Name.**
4. Select a **provisioning credential.**
5. Click **Next.**
6. Provide a **Data Lake Name.**
7. In the **Data Access and Audit** section, provide your data storage location and IAM resources created for minimal setup for cloud storage.
8. Click **Next.**
9. Select your region.
10. Under **Customer-Managed Keys**, click **Enable Customer-Managed Keys.**
11. In the same section, select the CMK:



Customer-Managed Keys

Enable Customer-Managed Keys

Select Encryption Key*

lrodek-cmk-quaero

12. Select network, security groups, and provide an SSH key. If required, add tags.
13. Click **Next.**
14. In the **Logs** section, provide your logs storage location and managed identities created for minimal setup for cloud storage.
15. Click **Register Environment.**

Steps - CDP CLI

In order to use this feature you must install CDP Beta CLI. See Installing Beta CDP CLI.

You can use your usual CDP CLI command to create an environment with a CMK, just add the **`--encryption-key-arn`** parameter and provide the encryption key created earlier. The easiest way to obtain the correct CLI template for creating an environment is by obtaining it from CDP wen UI as described in Obtain CLI commands from the register environment wizard.

For example:

This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.

```
cdp environments create-aws-environment \
--environment-name <ENVIRONMENT-NAME> \
--credential-name <EXISTING_CREDENTIAL> \
--region "<REGION>" \
--security-access cidr=<CIDR> \
--authentication publicKeyId="<SSH_KEY>" \
--log-storage storageLocationBase=<BUCKET_URL>,instanceProfile=<IDBROKER_IP>
\
--vpc-id <VPC_ID> \
--subnet-ids <SUBNETS \
--encryption-key-arn <ENCRYPTION_KEY_ARN>
```

The ARN of the encryption key created earlier should be passed in the parameter **--encryption-key-arn**

If the customer-managed encryption key ARN is not passed, then the AWS region-specific default encryption key is used for encrypting EBS volumes and RDS instances.

This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.