HashiCorp Vault on the AWS Cloud

Quick Start Reference Deployment

November 2016
Last update: April 2017 ([revisions](#))

Cameron Stokes, HashiCorp, Inc.
Tony Vattathil and Brandon Chavis, Amazon Web Services

Contents

This Quick Start deployment guide was created by Amazon Web Services (AWS) in partnership with HashiCorp, Inc.

Overview

This Quick Start reference deployment guide provides step-by-step instructions for deploying [HashiCorp](#) Vault on the Amazon Web Services (AWS) Cloud. [Quick Starts](#) 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.

HashiCorp Vault secures, stores, and tightly controls access to secrets across distributed infrastructure and applications from a central location. Vault handles leasing, key revocation, key rolling, and auditing. Through a unified API, users can access an encrypted key/value store and network encryption-as-a-service, or generate AWS Identity and Access Management (IAM) and AWS Security Token Service (AWS STS) credentials, SQL and NoSQL databases, X.509 certificates, Secure Shell (SSH) credentials, and more. Vault provides several key features:

- Secret sprawl and management: The growing popularity of microservices, infrastructure automation, and dynamic cloud environments has increased the number of secrets required to connect services and infrastructure. This secret sprawl expands the surface area for an attack, both in terms of potential infiltration points and internal damage in the event of a compromise. Vault uses time-bound, limited permissioned, dynamic secrets to reduce the potential impact of a secret compromise.

- Auditability: The volume of secrets in a modern infrastructure makes it difficult for security teams to organize, distribute, and secure secrets. Vault gives security operations certainty in when, where, and how secrets are being used across a system with a detailed audit log.

- Usability: Often the biggest hurdle to proper security is the complexity of implementing the security solution. With simple installation and setup, Vault lowers the barrier to entry for organizations to use responsible secret management across their infrastructure.

Vault is designed for both DevOps professionals and application developers, making it perfect for modern, elastic infrastructures.

This Quick Start is for users who looking for a service discovery solution, monitoring solution, or a key/value store. The Quick Start is built using the open-source version of Vault, but is also compatible with Vault Enterprise.

Additional details about Vault are available on the HashiCorp [Vault](#) and [Vault Enterprise](#) websites.

For additional solutions from HashiCorp and AWS, see the [AWS Quick Start for HashiCorp Consul](#).

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

This Quick Start uses the open-source version of HashiCorp Vault, which doesn't require a license.

Architecture

Deploying this Quick Start with the **default parameters** builds the following Consul and Vault environment in its own virtual private cloud (VPC) in the AWS Cloud. For details about the VPC architecture, see the [Amazon VPC Quick Start Guide](#).)
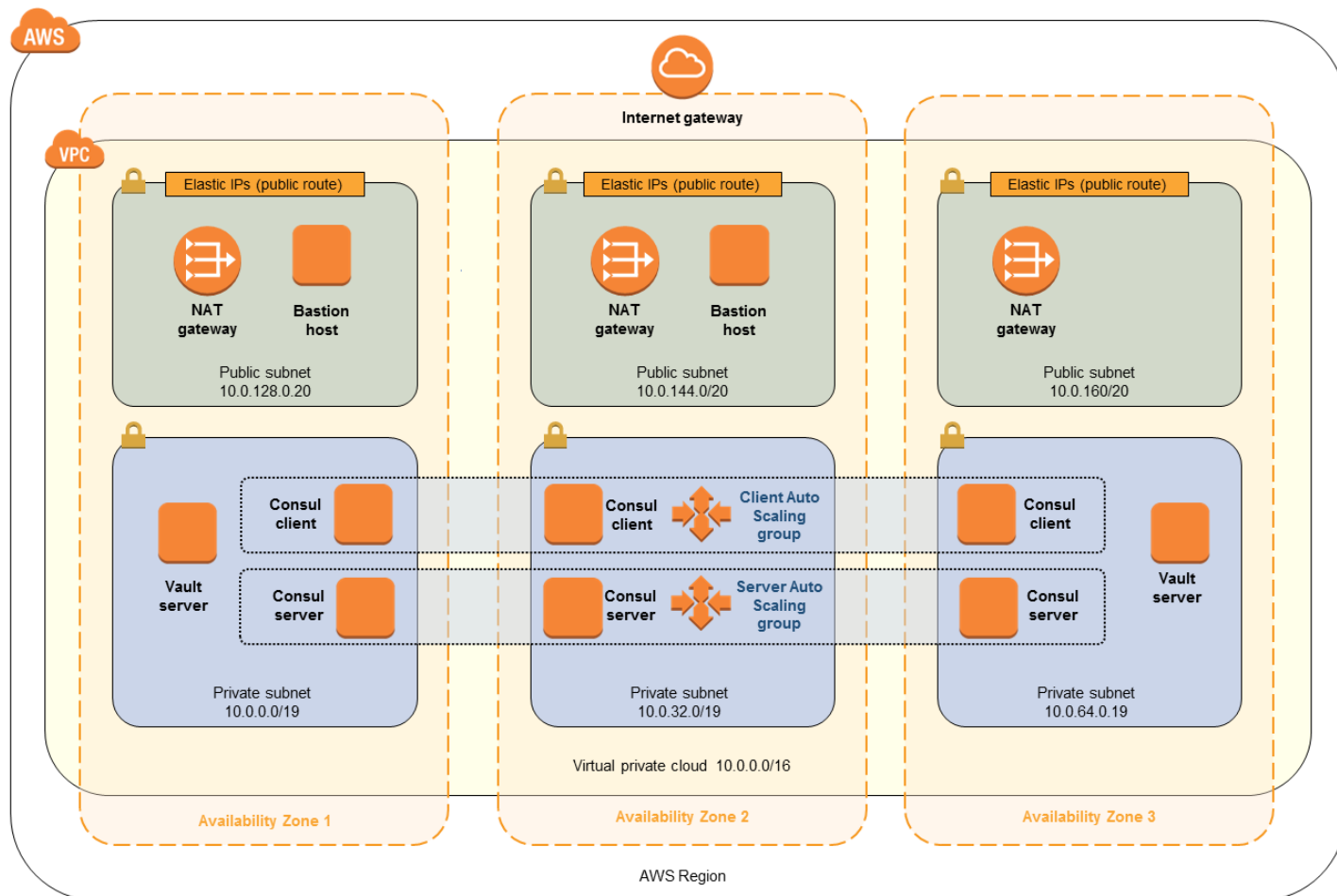
Figure 1: Quick Start Architecture for Consul and Vault on AWS

The Quick Start provides two deployment options:

- **Deployment of HashiCorp Vault into a new VPC** (end-to-end deployment) builds a new VPC with public and private subnets, and then deploys HashiCorp Vault into that infrastructure.

- **Deployment of HashiCorp Vault into an existing VPC** provisions HashiCorp Vault into your existing infrastructure.

If you use the deployment option to create a new VPC, the AWS CloudFormation template included with the Quick Start creates the following components:

- A VPC with public and private subnets across three Availability Zones.

- Linux bastion hosts in the public subnets to allow inbound Secure Shell (SSH) access to EC2 instances in the private subnets.

- A Consul environment, as described in the [HashiCorp Consul Quick Start deployment guide](#). Vault uses Consul DNS to discover and integrate with Consul.

- Two Vault server nodes in the private subnets.

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](#).)

- [Amazon VPC](#)
- [Amazon EC2](#)

Deployment Steps

Step 1. Prepare an AWS Account

1. If you don't already have an AWS account, create one at http://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 HashiCorp Vault on AWS.

3. Create a [key pair](#) in your preferred region.

4. If necessary, [request a service limit increase](#) for the Amazon EC2 **t2.medium** and **m4.large** instance types. You might need to do this if you already have an existing deployment that uses these instance types, and you think you might exceed the [default limit](#) with this reference deployment.

Step 2. Launch the Quick Start

1. Choose one of the following options to deploy the AWS CloudFormation template into your AWS account.

Launch Quick Start (for new VPC)   **Launch Quick Start (for existing VPC)**

The templates are launched in the US West (Oregon) region by default. You can change the region by using the region selector in the navigation bar.

Each stack takes approximately 10 minutes to create.

> **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. See the pricing pages for each AWS service you will be using for full details.

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

3. On the **Specify Details** page, review the parameters for the template. Enter values for the parameters that require your input. For all other parameters, you can customize the default settings provided by the template.

In the following tables, parameters are listed and described separately for deploying HashiCorp Vault into a new VPC or an existing VPC.

The deployment option for an existing VPC requires a Consul Quick Start environment to be present in your VPC. For more information, see the Quick Start deployment guide for HashiCorp Consul, or use the standalone (non-VPC) Consul Quick Start template.

> **Note**   The templates for the two scenarios share most, but not all, of the same parameters. For example, the template for an existing VPC prompts you for the VPC and private subnet IDs in your existing VPC environment. You can also download the templates and edit them to create your own parameters based on your specific deployment scenario.

- **Parameters for deployment into a new VPC:**

  View template

  VPC Network Configuration:

Parameter (name)	Default	Description
Availability Zones (AvailabilityZones)	*Requires input*	Choose three Availability Zones that will be used to deploy the components for HashiCorp Vault. The Quick Start preserves the logical order you specify.
VPC CIDR (VPCCIDR)	10.0.0.0/16	CIDR block for the VPC.
Private Subnet 1 CIDR (PrivateSubnet1CIDR)	10.0.0.0/19	CIDR block for the private subnet located in Availability Zone 1.

Parameter (name)	Default	Description
Private Subnet 2 CIDR (PrivateSubnet2CIDR)	10.0.32.0/19	CIDR block for the private subnet located in Availability Zone 2.
PrivateSubnet 3 CIDR (PrivateSubnet3CIDR)	10.0.64.0/19	CIDR block for the private subnet located in Availability Zone 3.
Public Subnet 1 CIDR (PublicSubnet1CIDR)	10.0.128.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 1.
Public Subnet 2 CIDR (PublicSubnet2CIDR)	10.0.144.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 2.
PublicSubnet 3 CIDR (PublicSubnet3CIDR)	10.0.160.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 3.
Permitted IP range (AccessCIDR)	*Requires input*	The CIDR IP range that is permitted to access the Vault environment. We recommend that you use a constrained CIDR range to reduce the potential of inbound attacks from unknown IP addresses.

Vault Setup:

Parameter	Default	Description
Email Address (EmailAddress)	*Requires input*	Email address for SNS topic. Alarms for Vault instance memory utilization.
Key Name (KeyPairName)	*Requires input*	Public/private key pair, which allows you to connect securely to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred region.
Vault node instance type (VaultInstanceType)	m4.large	EC2 instance type for the Vault node.

Consul Setup:

Parameter	Default	Description
Consul client nodes (ConsulClientNodes)	3	The number of client nodes that will be created.
Consul server nodes (ConsulServerNodes)	3	The number of Consul server nodes that will be created. You can choose 3, 5, or 7 nodes.
Consul cluster node instance type (ConsulInstanceType)	t2.medium	The EC2 instance type for the Consul instance.

AWS Quick Start Configuration:

Parameter	Default	Description
Quick Start S3 Bucket Name (QSS3BucketName)	quickstart-reference	S3 bucket name for the Quick Start assets. This bucket name can include numbers, lowercase letters, uppercase letters, and hyphens (-), but should not start or end with a hyphen. You can specify your own bucket if you copy all of the assets and submodules into it, if you want to override the Quick Start behavior for your specific implementation.
Quick Start S3 Key Prefix (QSS3KeyPrefix)	hashicorp/vault/latest	S3 key prefix for the Quick Start assets. This prefix can include numbers, lowercase letters, uppercase letters, hyphens (-), and forward slashes (/), but should not start or end with a forward slash (which is automatically added). This parameter enables you to override the Quick Start behavior for your specific implementation.

• **Parameters for deployment into an existing VPC:**

Parameter label	Default	Description
AccessCIDR	*Requires input*	The CIDR IP range that is permitted to access the Consul environment. We recommend that you set this value to a trusted CIDR block. For example, you might want to restrict access to your corporate network. A value of 0.0.0.0/0 will allow access from any IP address.
BastionSecurityGroupID	*Requires input*	The ID of the bastion host security group for enabling SSH connections (e.g., sg-7f16e910).
ConsulEC2RetryTagKey	*Requires input*	The EC2 instance tag key to filter on when joining to other Consul nodes.
ConsulEC2RetryTagValue	*Requires input*	The EC2 instance tag value to filter on when joining to other Consul nodes.
EmailAddress	*Requires input*	Email address for the Amazon SNS topic, which is triggered by Vault instance memory utilization alarms.
KeyPair	*Requires input*	Public/private key pair, which allows you to connect securely to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred region.

Parameter label	Default	Description
PrivateSubnet1ID	*Requires input*	ID of the private subnet in Availability Zone 1 (e.g., subnet-a0246dcd) where the first Vault server instance will be deployed.
PrivateSubnet2ID	*Requires input*	ID of the private subnet in Availability Zone 2 (e.g., subnet-e3246d8e) where the second Vault instance will be deployed.
QSS3BucketName	quickstart-reference	S3 bucket name for the Quick Start assets. This bucket name can include numbers, lowercase letters, uppercase letters, and hyphens (-), but should not start or end with a hyphen. You can specify your own bucket if you copy all of the assets and submodules into it, if you want to override the Quick Start behavior for your specific implementation.
QSS3KeyPrefix	hashicorp/vault/latest	S3 key prefix for the Quick Start assets. This prefix can include numbers, lowercase letters, uppercase letters, hyphens (-), and forward slashes (/), but should not start or end with a forward slash (which is automatically added). This parameter enables you to override the Quick Start behavior for your specific implementation.
QuickStartS3URL	https://s3.amazonaws.com	Used to dynamically generate URLs for sub-templates.
VPCCIDR	*Requires input*	CIDR block for your existing VPC.
VPCID	*Requires input*	ID of your existing VPC (e.g., vpc-0343606e).
VaultDownloadURL	https://releases.hashicorp.com/vault/0.7.0/vault_0.7.0_linux_amd64.zip	The URL used to download the Vault zip file.
VaultInstanceType	m4.large	EC2 instance type for the Vault node.

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

4. 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**.

5. On the **Review** page, review and confirm the template settings. Under **Capabilities**, select the check box to acknowledge that the template will create IAM resources.

6. Choose **Create** to deploy the stack.

7. Monitor the status of the stack. When the status is **CREATE_COMPLETE**, the deployment is complete.

8. You can use the URL displayed in the **Outputs** tab for the stack to view the resources that were created.

Step 3. Access Consul via SSH

To access the Vault environment, first connect to one of the bastion host instances. Use an SSH agent to forward your private key on connection.

> **Important**   Do not copy your private key to the bastion host.

For more information on SSH agents, see the [GitHub documentation.](#)

To use an SSH agent to access the Consul environment on Mac or Linux:

1. Use the command:

```
ssh-add ~/.ssh/id_rsa
```

2. At the prompt, type your passphrase or press **Enter** for no passphrase.

```
Enter passphrase (empty for no passphrase): [Hit Enter Again or
Enter passphrase]
Enter same passphrase again: [Hit Enter Again or Enter passphrase]
```

3. In the Amazon EC2 console, select one of the two bastion hosts and note its Elastic IP address.

Figure 2: Finding the Elastic IP address for the bastion host instance

In the example in Figure 2, the Elastic IP for `LinuxBastion1` is **52.14.140.157**.

4. Log in, and type **yes** when prompted to continue connecting:

```
ssh –A ubuntu@ 52.14.140.157
```


5. In the Amazon EC2 console, select one of the Vault hosts and note its private IP address.

Figure 3: Finding the private IP address for the Vault host

In the example in Figure 2, the private IP for Consul-Server is **172.31.13.203**.

6. From the bastion host, connect to the Vault host, using Ubuntu as the user:

7. View Consul members:

```
consul members
```

Step 4. Initialize Vault

Open an SSH tunnel from your local workstation and Linux bastion host:

```
export VAULT_ADDR=http://127.0.0.1:8200
vault init
```

You will see output similar to the following:

Warning   If you lose the keys shown in the previous output screen, your Vault will be permanently sealed.

Step 5. Unseal Vault

1. To unseal your Vault, use the command:

   ```
   vault unseal
   ```

   You will be prompted for the key. Cut and paste the keys shown in the output in step 4.

2. Repeat the `unseal` command three times.

   ```
   4. ubuntu@ip-172-31-13-203: ~ (ssh)
   ubuntu@ip-172-31-13-203:~$ vault unseal
   Key (will be hidden): Cut and paste Unseal Key 1
   Sealed: true
   Key Shares: 5
   Key Threshold: 3
   Unseal Progress: 1
   Unseal Nonce: 10c6d500-cc7f-92f6-c5d6-211642d6fcb3
   ubuntu@ip-172-31-13-203:~$ vault unseal
   Key (will be hidden): Cut and paste Unseal Key 2
   Sealed: true
   Key Shares: 5
   Key Threshold: 3
   Unseal Progress: 2
   Unseal Nonce: 10c6d500-cc7f-92f6-c5d6-211642d6fcb3
   ubuntu@ip-172-31-13-203:~$ vault unseal
   Key (will be hidden): Cut and paste Unseal Key 3
   Sealed: false
   Key Shares: 5
   Key Threshold: 3
   Unseal Progress: 0
   Unseal Nonce:
   ubuntu@ip-172-31-13-203:~$
   ```

3. Use the command `vault status` to check the status of Vault. If you've unsealed Vault successfully, it should output `Sealed: false` similar to the following:

   ```
   4. ubuntu@ip-172-31-13-203: ~ (ssh)
   ubuntu@ip-172-31-13-203:~$ vault status
   Sealed: false  <---
   Key Shares: 5
   Key Threshold: 3
   Unseal Progress: 0
   Unseal Nonce: Version: 0.6.5
   Cluster Name: vault-cluster-962622ba
   Cluster ID: 06b60c18-a7a3-20f6-0f83-daaea9f22d6e

   High-Availability Enabled: true
           Mode: active
           Leader: http://172.31.13.203:8200
   ubuntu@ip-172-31-13-203:~$
   ```

4. Repeat the `vault unseal` command for the second Vault server node (using **VaultNode2PrivateIp**) to unseal the second Vault server and to activate Vault's high availability mode.

Step 6. Enable Audit Logging

1. Authenticate by using the initial root token, which is provided as part of the Vault initialization output.

2. Enable the Vault audit logs:

```
vault audit-enable file file_path=/var/log/vault_audit.logstatus
```

3. This Quick Start is configured to ship Vault audit logs to Amazon CloudWatch. To see your logs, open the Amazon CloudWatch console at https://console.aws.amazon.com/cloudwatch/. In the navigation pane, choose **Logs**, and then choose **Vault-Audit-Logs**. You'll see a screen similar to Figure 4.

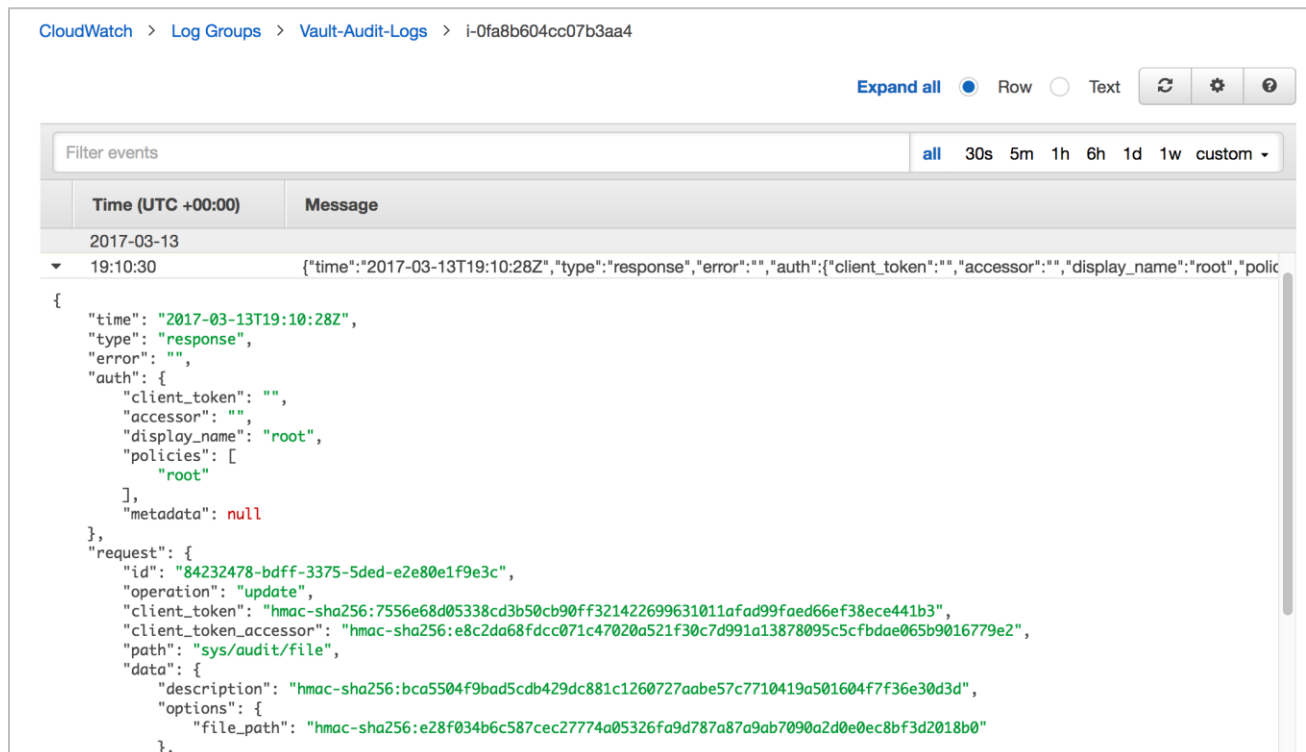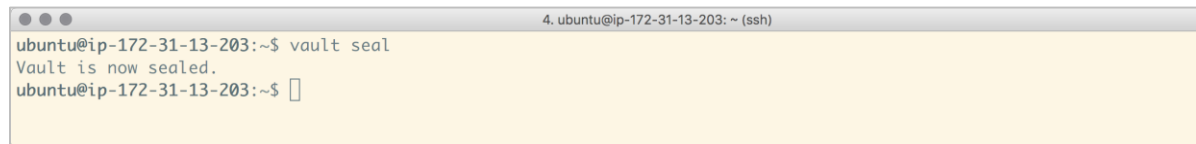**Figure 4: Viewing Vault audit logs**

Step 7. Seal Vault

1. To seal your Vault, use the command:

    ```
    vault seal
    ```

    Expected output:

    ```
    4. ubuntu@ip-172-31-13-203: ~ (ssh)
    ubuntu@ip-172-31-13-203:~$ vault seal
    Vault is now sealed.
    ubuntu@ip-172-31-13-203:~$
    ```

2. Check the status of your Vault:

    ```
    vault status
    ```

    Expected output:

    ```
    4. ubuntu@ip-172-31-13-203: ~ (ssh)
    ubuntu@ip-172-31-13-203:~$ vault status
    Sealed: true
    Key Shares: 5
    Key Threshold: 3
    Unseal Progress: 0
    Unseal Nonce: Version: 0.6.5

    High-Availability Enabled: true
            Mode: sealed
    ubuntu@ip-172-31-13-203:~$
    ```

Step 8. Get Started with Vault

To create your first secret and integrate Vault with your environment, see the getting started section of the HashiCorp Vault website.

Troubleshooting

Q. I encountered a CREATE_FAILED error when I launched the Quick Start. What should I do?

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](#) on the AWS website or contact us on the [AWS Quick Start Discussion Forum](#).

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](#).

Additional Resources

AWS services

- Amazon EC2
  http://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/
- Amazon VPC
  http://aws.amazon.com/documentation/vpc/

HashiCorp Vault

- HashiCorp
  https://www.hashicorp.com
- Vault
  https://www.vaultproject.io
- Vault Enterprise
  https://www.hashicorp.com/vault.html

Quick Start reference deployments

- AWS Quick Start home page
  https://aws.amazon.com/quickstart/

- AWS Quick Start for HashiCorp Consul
  https://s3.amazonaws.com/quickstart-reference/hashicorp/consul/latest/doc/hashicorp-consul-on-the-aws-cloud.pdf

Send Us Feedback

We welcome your questions and comments. Please post your feedback on the AWS Quick Start Discussion Forum.

You can visit our GitHub repository to download the templates and scripts for this Quick Start, and to share your customizations with others.

Document Revisions

Date	Change	In sections
April 2017	Added Linux bastion hosts; updated Consul to version 0.8.0; removed Seed server; added Amazon EC2 retry functionality	Changes in templates and throughout guide
November 2016	Initial publication	—

Notices