# HashiCorp Vault on the AWS Cloud

**Quick Start Reference Deployment**

*November 2016*

*Last update: October 2019 ([revisions](#))*

*Cameron Stokes, HashiCorp, Inc.*

*Tony Vattathil and Brandon Chavis, Amazon Web Services*

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>2</td>
</tr>
<tr>
<td>Costs and Licenses</td>
<td>3</td>
</tr>
<tr>
<td>Architecture</td>
<td>3</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>5</td>
</tr>
<tr>
<td>Specialized Knowledge</td>
<td>5</td>
</tr>
<tr>
<td>Deployment Steps</td>
<td>5</td>
</tr>
<tr>
<td>Step 1. Prepare an AWS Account</td>
<td>5</td>
</tr>
<tr>
<td>Step 2. Launch the Quick Start</td>
<td>5</td>
</tr>
<tr>
<td>Step 3. Access Consul via SSH</td>
<td>10</td>
</tr>
<tr>
<td>Step 4. Initialize Vault</td>
<td>13</td>
</tr>
<tr>
<td>Step 5. Unseal Vault</td>
<td>14</td>
</tr>
<tr>
<td>Step 6. Enable Audit Logging</td>
<td>15</td>
</tr>
<tr>
<td>Step 7. Seal Vault</td>
<td>16</td>
</tr>
<tr>
<td>Step 8. Get Started with Vault</td>
<td>16</td>
</tr>
<tr>
<td>Troubleshooting</td>
<td>16</td>
</tr>
<tr>
<td>Additional Resources</td>
<td>17</td>
</tr>
<tr>
<td>Send Us Feedback</td>
<td>18</td>
</tr>
<tr>
<td>Document Revisions</td>
<td>18</td>
</tr>
</tbody>
</table>
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](https://www.hashicorp.com/vault) and [Vault Enterprise](https://www.hashicorp.com/vault-enterprise) websites.

For additional solutions from HashiCorp and AWS, see the [AWS Quick Start for HashiCorp Consul](https://aws.amazon.com/quickstart/eks-hashicorp-consul/).

Please know that we may share who uses AWS Quick Starts with the AWS partner that collaborated with AWS on the content of the Quick Start.

**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](https://aws.amazon.com/quickstart/eks-aws-vpc/).
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.
• A Linux bastion host in the public subnet 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 quota 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 quota 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.
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:**

<table>
<thead>
<tr>
<th>Parameter (name)</th>
<th>Default</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Availability Zones</strong></td>
<td><strong>Requires input</strong></td>
<td>Choose three Availability Zones that will be used to deploy the components for HashiCorp Vault. The Quick Start preserves the logical order you specify.</td>
</tr>
<tr>
<td><em>(AvailabilityZones)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VPC CIDR</strong></td>
<td>10.0.0.0/16</td>
<td>CIDR block for the VPC.</td>
</tr>
<tr>
<td><em>(VPCCIDR)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Private Subnet 1 CIDR</strong></td>
<td>10.0.0.0/19</td>
<td>CIDR block for the private subnet located in Availability Zone 1.</td>
</tr>
<tr>
<td><em>(PrivateSubnet1CIDR)</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 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.
**Private Subnet 3 CIDR** (PrivateSubnet3CIDR) | 10.0.64.0/19 | CIDR block for the private subnet located in Availability Zone 3.
**Public Subnet 1 CIDR** (PublicSubnet1CIDR) | 10.0.128.0/19 | 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.
**Public Subnet 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:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quick Start S3 Bucket Name</strong></td>
<td><strong>aws-quickstart</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Quick Start S3 Key Prefix</strong></td>
<td><strong>quickstart-hashicorp-vault/</strong></td>
<td>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.</td>
</tr>
</tbody>
</table>

- Parameters for deployment into an existing VPC:

  **View template**

<table>
<thead>
<tr>
<th>Parameter label</th>
<th>Default</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AccessCIDR</strong></td>
<td><strong>Requires input</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>BastionSecurityGroupID</strong></td>
<td><strong>Requires input</strong></td>
<td>The ID of the bastion host security group for enabling SSH connections (e.g., sg-7f16e910).</td>
</tr>
<tr>
<td><strong>ConsulEC2RetryTagKey</strong></td>
<td><strong>Requires input</strong></td>
<td>The EC2 instance tag key to filter on when joining to other Consul nodes.</td>
</tr>
<tr>
<td><strong>ConsulEC2RetryTagValue</strong></td>
<td><strong>Requires input</strong></td>
<td>The EC2 instance tag value to filter on when joining to other Consul nodes.</td>
</tr>
<tr>
<td><strong>EmailAddress</strong></td>
<td><strong>Requires input</strong></td>
<td>Email address for the Amazon SNS topic, which is triggered by Vault instance memory utilization alarms.</td>
</tr>
<tr>
<td><strong>KeyPair</strong></td>
<td><strong>Requires input</strong></td>
<td>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.</td>
</tr>
<tr>
<td>Parameter label</td>
<td>Default</td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PrivateSubnet1ID</td>
<td>Requires input</td>
<td>ID of the private subnet in Availability Zone 1 (e.g., subnet-a0246dcd) where the first Vault server instance will be deployed.</td>
</tr>
<tr>
<td>PrivateSubnet2ID</td>
<td>Requires input</td>
<td>ID of the private subnet in Availability Zone 2 (e.g., subnet-e3246d8e) where the second Vault instance will be deployed.</td>
</tr>
<tr>
<td>QSS3BucketName</td>
<td>aws-quickstart</td>
<td>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.</td>
</tr>
<tr>
<td>QSS3KeyPrefix</td>
<td>quickstart-hashicorp-vault/</td>
<td>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.</td>
</tr>
<tr>
<td>QuickStartS3URL</td>
<td><a href="https://aws-quickstart.s3.amazonaws.com">https://aws-quickstart.s3.amazonaws.com</a></td>
<td>Used to dynamically generate URLs for sub-templates.</td>
</tr>
<tr>
<td>VPCCIDR</td>
<td>Requires input</td>
<td>CIDR block for your existing VPC.</td>
</tr>
<tr>
<td>VPCID</td>
<td>Requires input</td>
<td>ID of your existing VPC (e.g., vpc-0343606e).</td>
</tr>
<tr>
<td>VaultDownloadURL</td>
<td><a href="https://releases.hashicorp.com/vault/0.7.0/vault_0.7.0_linux_amd64.zip">https://releases.hashicorp.com/vault/0.7.0/vault_0.7.0_linux_amd64.zip</a></td>
<td>The URL used to download the Vault zip file.</td>
</tr>
<tr>
<td>VaultInstanceType</td>
<td>m4.large</td>
<td>EC2 instance type for the Vault node.</td>
</tr>
</tbody>
</table>

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](https://github.com).

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

1. Use the command:

   ```bash
   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.
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:

```
[ec2-user@ip-172-31-48-10 ~]$ ssh ubuntu@172.31.13.203
```

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:

```
Vault initialized with 5 keys and a key threshold of 3. Please securely distribute the above keys. When the Vault is re-sealed, restarted, or stopped, you must provide at least 3 of these keys to unsalve it again.

Vault does not store the master key. Without at least 3 keys, your Vault will remain permanently sealed.
```

**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.

   ![Unseal Key 1](image1)
   ![Unseal Key 2](image2)
   ![Unseal Key 3](image3)

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:

   ![Sealed False](image4)

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.

```
$ vault auth
Token (will be hidden): Copy and paste the Initial Root Token
Successfully authenticated! You are now logged in.
token: 27122110-ed06-e817-d923-6e63c0b2ccfc
token_duration: 0
token_policies: [root]
```

2. Enable the Vault audit logs:

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

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:

   ```bash
   vault seal
   ```

   Expected output:

   ```
   4. 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:

   ```bash
   vault status
   ```

   Expected output:

   ```
   4. 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
- Amazon VPC

HashiCorp Vault
- HashiCorp
  [https://www.hashicorp.com](https://www.hashicorp.com)
- Vault
  [https://www.vaultproject.io](https://www.vaultproject.io)
- Vault Enterprise
  [https://www.hashicorp.com/vault.html](https://www.hashicorp.com/vault.html)

Quick Start reference deployments
- AWS Quick Start home page
  [https://aws.amazon.com/quickstart/](https://aws.amazon.com/quickstart/)
• AWS Quick Start for HashiCorp Consul

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

<table>
<thead>
<tr>
<th>Date</th>
<th>Change</th>
<th>In sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>October 2019</td>
<td>Updated to reflect only one bastion host being deployed.</td>
<td>Architecture</td>
</tr>
<tr>
<td>April 2017</td>
<td>Added Linux bastion hosts; updated Consul to version 0.8.0; removed Seed server; added Amazon EC2 retry functionality</td>
<td>Changes in templates and throughout guide</td>
</tr>
<tr>
<td>November 2016</td>
<td>Initial publication</td>
<td>—</td>
</tr>
</tbody>
</table>
© 2020, Amazon Web Services, Inc. or its affiliates, and HashiCorp, Inc. 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.