HP Vertica on Amazon Web Services Guide

HP Vertica Analytic Database

Software Version: 7.1.x

Document Release Date: 8/18/2015

Legal Notices

Warranty

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted Rights Legend

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright Notice

© Copyright 2006 - 2015 Hewlett-Packard Development Company, L.P.

Trademark Notices

Adobe® is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Contents

Overview of HP Vertica on Amazon Web Services (AWS) 5

Supported Instance Types 5

Understanding the AWS Procedure Components 6

Creating Evaluation or Development Instances 11

Enhanced Networking 12

Packages 12

Installing and Running HP Vertica on AWS 13

Configuring and Launching an Instance 13

Creating a Placement Group 15

Creating a Key Pair 15

Creating a Virtual Private Cloud (VPC) 16

Network ACL Settings 17

Creating and Assigning an Internet Gateway 18

Creating a Security Group 19

Security Group Settings 19

Adding Rules to a Security Group 20

Assigning an Elastic IP 23

Connecting to an Instance 24

Connecting to an Instance from Windows Using Putty 25

Preparing Instances 25

Configuring Storage 26

Determining Volume Names 26

Combining Volumes for Storage 26

Forming a Cluster 27

Combining Instances 27

Considerations When Using the install_vertica or update_vertica Scripts 28

After Your Cluster Is Up and Running 28

Initial Installation and Configuration 29
Using Management Console (MC) on AWS 29

Adding Nodes to a Running AWS Cluster 31
Launching New Instances to Add to an Existing Cluster 31
Including New Instances as Cluster Nodes 31
Adding Nodes and Rebalancing the Database 32

Removing Nodes From a Running AWS Cluster 33
Preparing to Remove a Node 33
Removing Hosts From the Database 34
Removing Nodes From the Cluster 35
Stopping the AWS Instances (Optional) 35

Migrating Data Between AWS Clusters 36

Migrating to HP Vertica 7.0 or later on AWS 39

Upgrading to the version 7.0 HP Vertica AMI on AWS 41
Preparing to Upgrade Your AMI 41
Upgrading HP Vertica Running on AWS 41

Troubleshooting: Checking Open Ports Manually 43
Using the Netcat (nc) Utility 43

Quick Start to Setting Up HP Vertica AWS 45

Send Documentation Feedback 47

Overview of HP Vertica on Amazon Web Services (AWS)

This document describes the steps to create an HP Vertica cluster on AWS. To run an HP Vertica cluster on AWS requires creating Amazon Machine Instances (AMIs). The instructions in this document apply to AMIs built with HP Vertica Version 7.1.x.

The HP Vertica Community Edition is installed on the AMI. Community Edition is limited to three nodes and up to 1 TB of data. Each AMI includes a Community Edition license. Once HP Vertica is installed , you can find the license at this location:

```
/opt/vertica/config/licensing/vertica_community_edition.license.key
```

Most of the remainder of this document describes the details of how to prepare your AWS environment, launch AMI instances, and combine instances to create a cluster. To set up your HP Vertica cluster on AWS, follow the detailed directions that follow, or use the summarized set of tasks in Quick Start to Setting Up HP Vertica AWS

Supported Instance Types

HP Vertica supports a range of Amazon Web Services (AWS) instance types, each optimized for different purposes. For more information about Amazon cluster instances and their limitations, see the Amazon documentation.

Sizing an instance involves estimating the hardware requirements for optimal HP Vertica performance in a typical scenario. Hewlett Packard recommends using its AMI with the instance type optimized for your requirements, as follows:

Optimization	Type	Description
Compute	c3.4xlarge c3.8xlarge	Compute optimized instances employ high performance CPUs for calculation-intensive applications.

Memory	r3.4xlarge r3.8xlarge	Memory optimized configured for memory intensive applications such as in-memory analytics, and enterprise-level collaborative applications.
Storage	i2.4xlarge i2.8xlarge	Storage optimized use solid-state drives (SSD) for rapid read and write performance.

Note: Data stored on the ephemeral drives of a storage-optimized instance exists only while that instance is powered on. After powering off a storage-optimized system, data on ephemeral drives is lost.

Understanding the AWS Procedure Components

The following illustration and table display and describe the AWS components that you must configure. These are preliminary steps in creating an HP Vertica cluster on Amazon AWS. You can skip this section and go right to the Installing and Running HP Vertica on AWS.

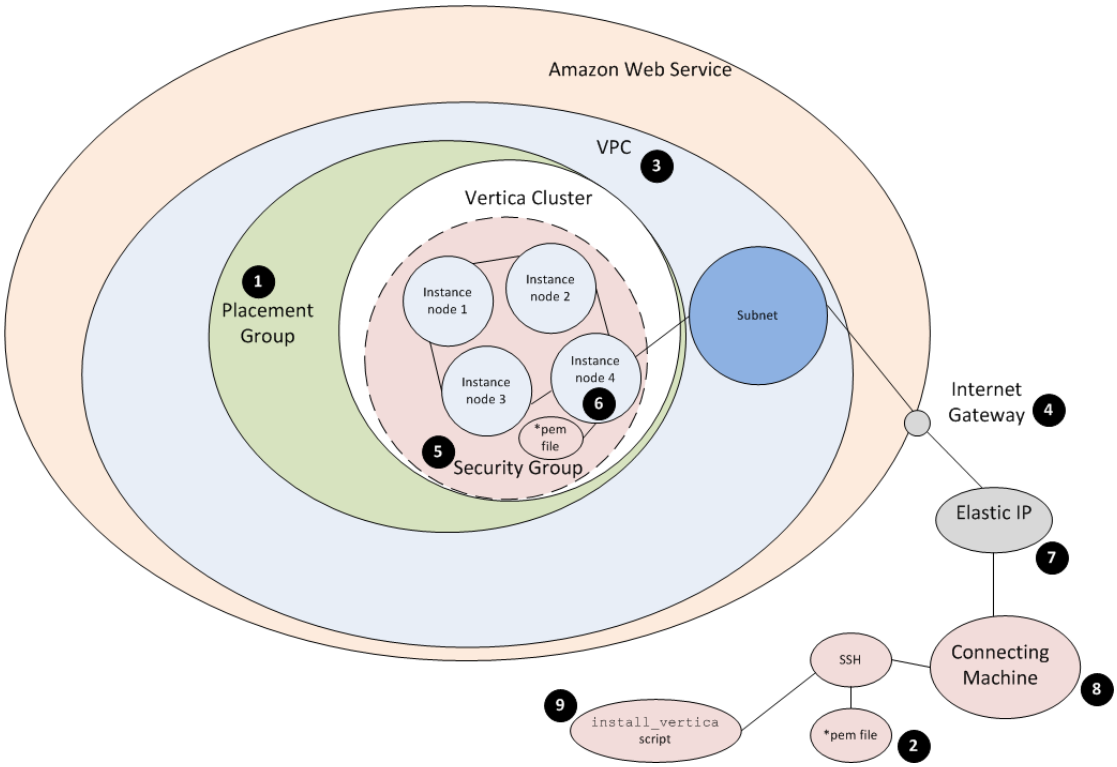


Illustration Reference and Step #	Component	Sample	Description
1	Placement Group	GroupVerticaP	You supply the name of a Placement Group when you create instances. You use a Placement Group to group instances together. A Placement Group for a cluster resides in one availability zone; your Placement Group cannot span zones. A Placement Group includes instances of the same type; you cannot mix types of instances in a Placement Group. You can choose one of two regions for a Placement Group; see Creating a Placement Group for information.

Illustration Reference and Step #	Component	Sample	Description
2	Key Pair	MyKey (MyKey.pem would be the resultant file.)	You need a Key Pair to access your instances using SSH. You create the Key Pair through the AWS interface, and you store a copy of your key (*.pem) file on your local machine. When you access an instance, you need to know the local path of your key, and you copy the key to your instance before you can run the `install_vertica` script.
3	VPC	Vpc-d6c18dbd Amazon assigns this name automatically.	You create a Virtual Private Cloud (VPC) on Amazon so that you can create a network of your EC2 instances. All instances within the VPC share the same network and security settings. A VPC uses the CIDR format for specifying a range of IP addresses.
4	Internet Gateway	Igw-d7c18dbc Amazon assigns and attaches an internet gateway automatically.	An internet gateway allows instances to access the internet. Note that, typically, a gateway is automatically assigned when you create a VPC. You can also create your own named internet gateway, and then attach that gateway to your VPC.

Illustration Reference and Step #	Component	Sample	Description
5	Security Group	"default" Amazon assigns a security group named "default" automatically, but you can create and name your own security group.	The security group includes firewall settings; its rules specify how traffic can get in and out of your instances. When you launch your instances, you choose a security group. If you use the default security group, you must add HP Vertica recommended rules to the group as you would if you created the group yourself.

Illustration Reference and Step #	Component	Sample	Description
6	Instances	10.0.3.158 (all samples) 10.0.3.159 10.0.3.157 10.0.3.160 (Amazon assigns IP addresses. You need to list the addresses when you run the `install_vertica` script to form a cluster of the instances.)	An instance is a running version of the AMI. You first choose an AMI and specify the number of instances. You then launch those instances. Once the instances are running, you can access one or more of them through SSH. Notes: • Once IPs are assigned to your instances, they must not change in order for HP Vertica to continue working. The IP addresses are crucial once you have run the `install_vertica` script; if you change them or they become reassigned, your cluster breaks. • When you stop one or more instances, your My Instances page may show blanks for the IP fields. However, by default Amazon retains the IPs and they should show up again in short order.
7	Elastic IP	107.23.104.78 (sample)	You associate an elastic IP to one of your instances so you can communicate with your cluster. An elastic IP is a static IP address that stays connected to your account until you explicitly release it.

Illustration Reference and Step #	Component	Sample	Description
8	Connecting	---	Once you have completed all of the procedures for setting up your cluster, you can connect to the instance attached to your elastic IP. You can use SSH. If connecting from a Windows machine, you can, for example, use Putty.
9	`install_vertica` script	---	Your instances are already grouped through Amazon and include the HP Vertica software, but you must then run the `install_vertica` script to combine instances through their IP addresses so that HP Vertica knows the instances are all part of the same cluster.

Creating Evaluation or Development Instances

You can create a very simple HP Vertica instance that you can use for testing, development, or evaluation. Although you can create multiple one-click instances, these instances do not share a placement group, do not have mounted volumes, and cannot form a cluster. If you want to create an enterprise or cluster deployment of HP Vertica, launch the AMI from the EC2 console by selecting the Manual Launch tab.

The 1-Click AMI is available from the Amazon Marketplace at https://aws.amazon.com/marketplace/library?ref_=gtw_navgno_library

1. Navigate to the AWS Console, and select **EC2**.
2. From the EC2 Console Dashboard, click **Launch Instance**.

3. Select the HP Vertica AMI.
    a. Select the **AWS Marketplace** tab and enter "HP/Vertica" in the search field.
    b. Choose your HP Vertica AMI. Click the **Select** button next to the AMI.
4. Click **Launch with 1-Click**. The AMI launches.

Enhanced Networking

HP Vertica AMIs support the AWS enhanced networking feature. See Enabling Enhanced Networking on Linux Instances in a VPC in the AWS documentation for details.

Packages

HP Vertica AMIs come with the following packages preinstalled:

- HP Vertica Place
- HP Vertica Pulse
- HP Distributed R Extensions for HP Vertica

Installing and Running HP Vertica on AWS

Use these procedures to install and run HP Vertica on AWS. Note that this document mentions basic requirements only, by way of a simple example. Refer to the AWS documentation for more information on each individual parameter.

For a high level overview of the installation process, refer to Appendix: Installing and Running HP Vertica on AWS - QuickStart

Configuring and Launching an Instance

Perform the following steps to configure and launch the instances that will become your cluster in a later procedure. A basic Elastic Compute Cloud (EC2) instance (without an HP Vertica AMI) is similar to a traditional host. When you create an EC2 instance using an HP Vertica AMI, the instance includes the HP Vertica software and a standard recommended configuration to ease creation of your cluster. HP Vertica recommends that you use the HP Vertica AMI as is – without modification. The HP Vertica AMI acts as a template, requiring fewer configuration steps.

1. Navigate to the AWS Console, and select **EC2**.
2. From the EC2 Console Dashboard, click **Launch Instance**.
3. Select the HP Vertica AMI.
    a. Select the **AWS Marketplace** tab and enter "HP Vertica" in the search field.
    b. Choose your HP Vertica AMI. Click the **Select** button next to the AMI.
4. Click Continue. The launch page opens.
5. Click Manual Launch, select a region, and click **Launch with EC2 Console**. The Choose Instance Type page opens.
6. Select a supported instance type and click **Next: Configure Instance Details**.

7. Click **Next: Configure Instance Details**.
    a. Choose the number of instances you want to launch. An HP Vertica cluster uses identically configured instances of the same type. You cannot mix instance types.
    b. From the **Network** drop-down, choose your VPC.

    > **Note:** Not all data centers support VPC. If you receive an error message that states "VPC is not currently supported...", choose a different region and zone (for example, choose us-east-1e rather than us-east-1c).

    c. From the **Placement group** drop-down, choose a placement group. Alternatively, you can Create a Placement Group.
8. Click **Next: Add Storage**.The Add Storage page opens.
9. Add storage to your instances based on your needs.

    > **Note:**
    >
    > - HP Vertica recommends that you add a number of drives equal to the number of cores in your instance. For example, for a c3.8xlarge, add eight drives. For an r3.4xlarge, add four drives.
    > - HP Vertica does not recommend that you store data on the root drive.
    > - For optimal performance with EBS volumes, Amazon recommends that you configure them in a RAID 0 array on each node in your cluster.

10. Click **Next: Tag Instance**. The Tag Instance page opens.
11. Create and add a key value pair if needed and click **Next: Configure Security Group**. The Configure Security Group page opens.
12. Create or select a security group and click **Review and Launch**. The Review Instance Launch page opens.
13. Review your instance details and click **Launch**. A key value pair dialog box opens.
14. Configure your key value pairs, and click **Launch Instances**.

15. Click **Launch Instances**. You receive a message saying that your instances are launching.

You can click **View Instances** from the Launch Status page. Check that your instances are running (show green as their state).

Note: You can stop an instance by right-clicking and choosing **Stop**. Use Stop rather than Terminate; the Terminate command deletes your instance.

Creating a Placement Group

Perform the following steps to create a Placement Group. A Placement Group ensures that your nodes are properly co-located.

1. Log in to your Amazon EC2 Console Dashboard.
2. Select **Placement Groups**. The Placement Group screen appears, listing your existing placement groups.
3. Click **Create Placement Group**.
4. Name your placement group.
5. Click **Create**. Your group is created.

Creating a Key Pair

Perform the following steps to create a Key Pair.

1. Select **Key Pairs** from the Navigation panel.

   Note: Depending upon which browser you are using, you may have to turn off the pop-up blocker in order to download the Key Pair.

2. Click **Create Key Pair**.

3. Name your Key Pair.
4. Click **Yes**.

   The system displays a message letting you know that the Key Pair has been created.
5. Save your key pair. Ensure that you keep the *.pem file; you need it to logon to your instances.

   Note that your Key Pair name now appears on the Key Pair list.

Creating a Virtual Private Cloud (VPC)

Perform the following steps to create a VPC.

1. From the AWS Management Console Home, navigate to the VPC console by selecting VPC.

   Note: If the VPC Console Dashboard shows that VPCs already exist, you can select Your VPCs to note the names of the existing VPCs. As the VPC IDs are very similar, noting the names of the existing VPCs helps you later in identifying the new one you are about to create.
2. Click **Start VPC Wizard**.
3. From the wizard that displays, select **VPC with a Single Public Subnet Only**.
4. Change the Public Subnet as desired. HP Vertica recommends that you secure your network with an Access Control List (ACL) that is appropriate to your situation. The default ACL does not provide a high level of security.
5. Choose an Availability Zone.

   Note: An HP Vertica cluster is operated within a single availability zone.
6. Click **Create VPC**. Amazon displays a message noting success.

7. Choose **Your VPCs** from the navigation pane, select your new VPC, and be sure that both **Enable DNS resolution** and **Enable DNS hostname support for instances launched in this VPC** are both checked.
8. Click **Close**. Your virtual private cloud is created.
9. Add the required network inbound and outbound rules to the VPC.

Network ACL Settings

HP Vertica requires the following network access control list (ACL) settings on an AWS instance running the HP Vertica AMI.

For detailed information on network ACLs within AWS, refer to Amazon's documentation.

Inbound Rules

Type	Protocol	Port Range	Source	Allow/Deny
SSH (22)	TCP (6)	22	0.0.0.0/0	Allow
Custom TCP Rule	TCP (6)	5450	0.0.0.0/0	Allow
Custom TCP Rule	TCP (6)	5433	0.0.0.0/0	Allow
Custom TCP Rule	TCP (6)	1024-65535	0.0.0.0/0	Allow
ALL Traffic	ALL	ALL	0.0.0.0/0	Deny

Outbound Rules

Type	Protocol	Port Range	Source	Allow/Deny
Custom TCP Rule	TCP (6)	0 - 65535	0.0.0.0/0	Allow

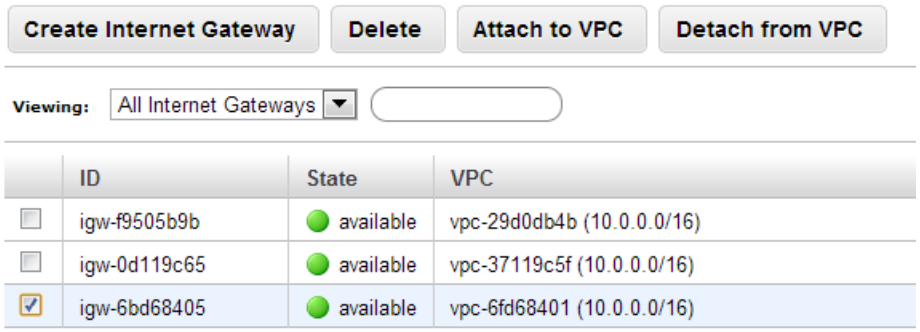
Creating and Assigning an Internet Gateway

When you create a VPC, an internet gateway is automatically assigned to the VPC. You can use that gateway, or you can assign your own. To create and assign your own, perform the following steps. If using the default, continue with the next procedure, Creating a Security Group.

1. From the navigation pane, choose **Internet Gateways**.
2. Click **Create Internet gateway**.
3. Choose **Yes, Create**.

   The Internet Gateway screen appears. Note that the new gateway is not assigned to a VPC.
4. Click **Attach to VPC**.
5. Choose your VPC.
6. Click **Yes, Attach**.

   Note that your gateway is now attached to your VPC.



Creating a Security Group

When you create a Virtual Private Cloud (VPC), AWS automatically creates a default security group assigns it to the VPC. You can use that security group, or you can name and assign your own. To create and assign your own, perform the following steps. If using the default, continue with the next procedure, Adding Rules to a Security Group. Note that you must add the HP Vertica rules as described in the next section. The HP Vertica AMI has specific security group requirements.

To create and name your own security group, perform the following steps.

1. From the Navigation pane, select **Security Groups**.
2. Click **Create Security Group**.
3. Enter a name for the group and provide a description.
4. Select a VPC to enable communication between nodes.
5. Click **Create**. The security group is created.

Security Group Settings

HP Vertica requires the following security group settings on an AWS instance running the HP Vertica AMI.

For detailed information on security groups within AWS, refer to Amazon's documentation.

Type	Protocol	Port Range	Source	IP
	TCP	22	My IP	0.0.0.0/0
	TCP	80	My IP	0.0.0.0/0
	TCP	443	My IP	0.0.0.0/0

Type	Protocol	Port Range	Source	IP
	TCP	4803	My IP	0.0.0.0/0
	TCP	5433	My IP	0.0.0.0/0
	TCP	5450	My IP	0.0.0.0/0
	TCP	5444	My IP	0.0.0.0/0
	TCP	8080	My IP	0.0.0.0/0
	TCP	48073	My IP	0.0.0.0/0
	TCP	50000	My IP	0.0.0.0/0
	UDP	53	My IP	0.0.0.0/0
	UDP	4803	My IP	0.0.0.0/0
	ICMP	0 to -1	My IP	0.0.0.0/0
	ICMP	-1	My IP	0.0.0.0/0
	ICMP	30 to -1	My IP	0.0.0.0/0

Adding Rules to a Security Group

Perform the following steps to add rules to the security group you plan to use (whether you plan to use the default group or have created your own). This section includes procedures for adding rules from the **Inbound** tab and adding rule from the **Outbound** tab. You perform both procedures to establish your security group.

Perform the following to add rules from the **Inbound** tab:

1. Ensure that you have checked the box for the security group you just created, and that the **Inbound** tab is selected.
2. Add the HTTP rule.

   a. Ensure that the **Inbound** tab is selected.
   b. From the **Create a new rule** dropdown, choose **HTTP**.
   c. Click **Add Rule**. The HTTP rule is added to your security group.
3. Add the Echo Reply rule.
   a. From the **Create a new rule** dropdown, select **Custom ICPM rule**.
   b. From the Type dropdown, select **Echo Reply**.
   c. Click **Add Rule**. The Echo Reply rule is added to your security group.
4. Add the Traceroute rule.
   a. Choose **Custom ICMP rule** once again.
   b. Select **Traceroute**.
   c. Click **Add Rule**. The Traceroute rule is added to your security group.
5. Add the SSH rule.
   a. From the **Create a new rule** dropdown, choose **SSH**.
   b. Click **Add Rule**. The SSH rule is added to your security group
6. Add the HTTPS rule.
   a. From the **Create a new rule** dropdown, choose **HTTPS**.
   b. Click **Add Rule**. The HTTPS rule is added to your security group.
7. Add a port range to your security group.
   a. From the **Create a new rule** dropdown, select **Custom TCP Rule**.
   b. Under **Port range**, enter **4803-4805**.
   c. Click **Add Rule**. The port range is added to your security group.
   d. From the **Create a new rule** dropdown, select **Custom UDP Rule**.

e. Under **Port range**, enter **4803-4805**.

f. Click **Add Rule**. The port range is added to your security group.

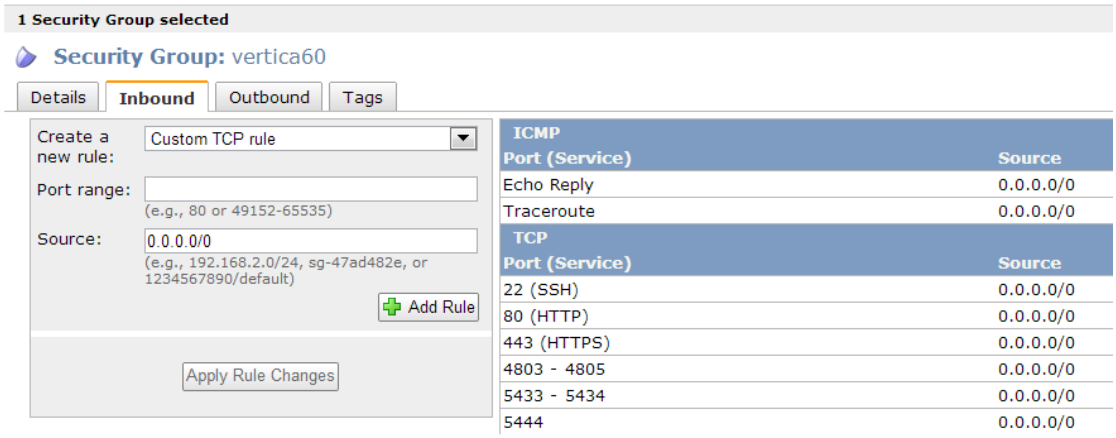
8. Add individual ports.

   a. Also under **Custom TCP rule**, enter the following ports under **Port Range**, one by one: **5433**, **5434**, **5444**, and **5450**. You enter the ports one by one to ensure the port assignments are sequential. HP Vertica uses these ports for internode communication.

   b. Click **Add Rule** as you enter each number. Each port is added to your security group.

9. Click **Apply Rule Changes**.

> **Note:** You must click **Apply Rule Changes** or your rules will not be applied to your security group. With the **Inbound** tab selected, your security group screen should look similar to the following.



Perform the following to add rules from the **Outbound** tab.

> **Note:** You want to ensure that all outbound traffic is allowed.

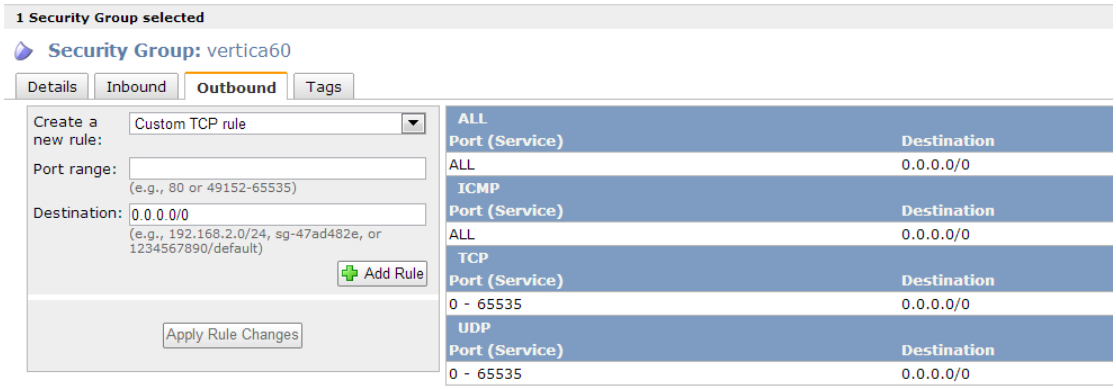
1. Select the **Outbound** tab.

   a. Choose **All TCP** rule from the **Create a new rule** dropdown.

   b. Click **Add Rule**. The All TCP rule is added to your security group.

2. Add the **All ICMP** rule.
    a. Choose **All ICMP** rule from the **Create a new rule** dropdown.
    b. Click **Add Rule**. The All ICMP rule is added to your security group.
3. Add the **ALL UDP** rule.
    a. Choose **ALL UDP** rule from the **Create a new rule** dropdown.
    b. Click **Add Rule**. The ALL UDP rule is added to your security group.
4. Click **Apply Rule Changes**.

Note: You must click **Apply Rule Changes** or your rules will not be applied to your security group. With the **Outbound** tab selected, your screen should look similar to the following.



Assigning an Elastic IP

The elastic IP is an IP address that you attach to an instance; you communicate with your cluster through the instance that is attached to this IP address. An elastic IP is a static IP address that stays connected to your account until you explicitly release it.

Note the following on IP addresses:

- You can stop your instances, but you must ensure that the IP addresses assigned are not released.

- By default, Amazon keeps your IP addresses assigned until you release them.

1. From the Navigation menu, select **Elastic IPs**.
2. Click **Allocate New Address**.
3. On the **Allocate New Address** screen, choose **VPC** from the dropdown.
4. Click **Yes, Allocate**. The elastic IP is created and appears in the list of available addresses.
5. Select the address that you want to assign and click **Associate Address**.
6. Choose one of the instances you created.
7. Click **Associate**. Your instance is associated with your elastic IP.

Connecting to an Instance

Perform the following procedure to connect to an instance within your VPC.

1. As the dbadmin user, type the following command, substituting your ssh key:

```
# ssh --ssh-identity <ssh key> dbadmin@ipaddress
```

2. Select **Instances** from the Navigation panel.
3. Select the instance that is attached to the Elastic IP.
4. Click **Connect**.
5. On **Connect to Your Instance**, choose one of the following options:
   - **A Java SSH Client directly from my browser—**Add the path to your private key in the field **Private key path**, and click **Launch SSH Client**.
   - **Connect with a standalone SSH client—**Follow the steps required by your standalone SSH client.

Connecting to an Instance from Windows Using Putty

If you connect to the instance from the Windows operating system, and plan to use Putty:

1. Convert your key file using PuTTYgen.
2. Connect with Putty or WinSCP (connect via the elastic IP), using your converted key (i.e., the *ppk file).
3. Move your key file (the *pem file) to the root dir using Putty or WinSCP.

Preparing Instances

After you create your instances, you need to prepare them for cluster formation. Prepare your instances by adding your AWS .pem key and your HP Vertica license.

1. As the root user, copy your *pem file (from where you saved it locally) onto your primary instance.

   Depending upon the procedure you use to copy the file, the permissions on the file may change. If permissions change, the `install_vertica` script fails with a message similar to the following:

   ```
   FATAL (19): Failed Login Validation 10.0.3.158, cannot resolve
   or connect to host as root.
   ```

   If you receive a failure message, enter the following command to correct permissions on your *pem file:

   ```
   chmod 600 /<name-of-pem>.pem
   ```

2. Copy your HP Vertica license over to your primary instance, and also place it in /root.

Configuring Storage

As a best practice, use dedicated volumes for node storage.

Important: Hewlett-Packard recommends that you do *not* store your data on the root drive. When configuring your storage, make sure to use a supported file system.

For best performance, you can combine multiple EBS volumes into RAID-0. HP Vertica provides a shell script, which automates the storage configuration process.

Note: To take advantage of bursting, limit EBS volumes to 1TB or less.

Determining Volume Names

Before you combine volumes for storage, make note of your volume names so that you can alter the `configure_aws_raid.sh` shell script. You can find your volumes with the following commands:

```
cd /dev
ls
```

Your volumes start with xvd.

Important: Ignore your root volume. Do not include any of your root volumes in the RAID creation process.

Combining Volumes for Storage

Follow these sample steps to combine your EBS volumes into RAID 0 using the `configure_aws_raid.sh` shell script.

1. Edit the `/opt/vertica/sbin/configure_aws_raid.sh` shell file as follows:
   a. Comment out the safety `exit` command at the beginning .
   b. Change the sample volume names to your own volume names, which you noted previously. Add more volumes, if necessary.
2. Run the `/opt/vertica/sbin/configure_aws_raid.sh` shell file. Running this file creates a RAID 0 volume and mounts it to /vertica/data.
3. Change the owner of the newly created volume to dbadmin with `chown`.
4. Repeat steps 1-3 for each node on your cluster.

For more information about EBS storage, refer to the Amazon documentation.

Forming a Cluster

Use the `install_vertica` script to combine two or more individual instances and create a cluster.

Check the **My Instances** page for a list of current instances and their associated IP addresses. You need these IP addresses when you run the `install_vertica` script.

Combining Instances

The following example combines instances using the `install_vertica` script.

> **Important:** Before you run `install_vertica`, be sure to stop any running databases on your nodes. The `install_vertica` script cannot complete successfully if any databases are running.

While connected to your primary instance, enter the following command to combine your instances into a cluster. Substitute the IP addresses for your instances and include your root *pem file name.

```
sudo /opt/vertica/sbin/install_vertica --hosts
10.0.11.164,10.0.11.165,10.0.11.166 --dba-user-password-disabled
--point-to-point --data-dir /vertica/data --ssh-identity ~/<name-
of-pem>.pem --license <license.file>
```

Note: If you are using Community Edition, which limits you to three instances, you can simply specify `-L CE` with no license file.

When you issue `install_vertica` or `update_vertica` on an AMI, always use the `--point-to-point` parameter.This parameter configures spread to use direct point-to-point communication between all HP Vertica nodes, which is a requirement for clusters on AWS. If you do not use the parameter, you receive an error telling you that you must use point-to-point communication on AWS.

Considerations When Using the install_vertica or update_vertica Scripts

- By default, the installer assumes that you have mounted your storage to `/vertica/data`. To specify another location, use the `--data-dir` argument. Hewlett-Packard does not recommend that you store your data on the root drive.
- Password logons present a security risk on AWS. Include the parameter `--dba-user-password-disabled` so that the installer does not prompt for a password for the database user.

For complete information on the install_vertica script and its parameters, see the Installation Guide, specifically the section, *About the install_vertica Script*.

After Your Cluster Is Up and Running

Stop or reboot instances using the Amazon AWS console, but you must stop the database before doing so. Once your cluster is up and running, if you need to stop or reboot:

1. Stop the database.
2. Stop or reboot one or more instances.

Caution: If you stop or reboot an instance (or the cluster) without shutting the database down first, disk or database corruption could result. Shutting the database down first ensures that HP Vertica is not in the process of writing to disk when you

shutdown. Refer to the HP Vertica Administrator's Guide for information on stopping a database.

Once your cluster is configured and running:

1. Create a database. When HP Vertica was installed for you, an HP Vertica database administrator was created, dbadmin. You can use this pre-created dbadmin user to create and start a database. Refer to the HP Vertica Installation Guide for information on the dbadmin administrator.
2. Configure a database. Refer to the HP Vertica Administrator's Guide for information on configuring a database.
3. Refer to the full documentation set for HP Vertica for other tasks.

Initial Installation and Configuration

Once you have created your cluster on AWS, you can log on to your nodes and perform an HP Vertica installation.

1. Log on to your cluster using the following command:

   ```
   # ssh -i <ssh key> dbadmin@ipaddress
   ```

2. Run the HP Vertica installer. Be sure to include all of the nodes in your cluster.
3. Create a database.

Using Management Console (MC) on AWS

MC is a database management tool that provides a way for you to view and manage aspects of your HP Vertica cluster. If you are running HP Vertica Release 6.1.2 or later, you can install and run MC.

This release of MC on AWS includes restrictions.

- You cannot create a cluster on AWS using the MC. You cannot import a cluster into AWS using the MC.

- You cannot monitor an AWS cluster using MC on a node that is outside of your AWS cluster. You must install MC on an instance within the AWS cluster itself.

Note: Each version of HP Vertica Management Console (MC) is compatible only with the matching version of the HP Vertica server. Version numbers must match to three digits; for example, HP Vertica 6.1.2 server is supported with HP Vertica 6.1.2 MC only. This is a general MC requirement and is not specific to MC on AWS.

What follows is a reading path for learning more about MC:

- For an overview of MC, where you can get it, and what you can do with it, refer to the Concepts Guide, specifically, *Management Console*.
- For information on installing and configuring MC, refer to the Installation Guide, specifically, *Installing and Configuring Management Console (MC)*.
- For information on the differences between what you can do with MC versus what you can do with the Administration Tools, refer to the Administration Guide, specifically the section, *Administration Tools and Management Console*.
- For information on creating a database using MC, refer to the Getting Started Guide, specifically the section, *Create the Example Database Using Management Console*.

Keep the following in mind concerning user accounts and the MC.

- When you first configure MC, during the configuration process you create an MC superuser (a Linux account). Issuing a Factory Reset on the MC does not create a new MC superuser, nor does it delete the existing MC superuser. When initializing after a Factory Reset, you must logon using the original MC superuser account.

  For information on setting MC to its original state (Factory Reset), and why you might implement a Factory Reset, refer to the Administration Guide, specifically the section, Resetting MC to Pre-configured state.

- Note that, once MC is configured, you can add users that are specific to MC. Users created through the MC interface are MC specific. When you subsequently change a password through the MC, you only change the password for the specific MC user. Passwords external to MC (i.e., system Linux users and HP Vertica database passwords) remain unchanged.

  For information on MC users, refer to the Administration Guide, specifically the sections, Creating an MC User and MC configuration privileges.

Adding Nodes to a Running AWS Cluster

Use these procedures to add instances/nodes to an AWS cluster. The procedures assume you have an AWS cluster up and running and have most-likely accomplished each of the following.

- Created a database.
- Defined a database schema.
- Loaded data.
- Run the database designer.
- Connected to your database.

Launching New Instances to Add to an Existing Cluster

Perform the procedure in Configuring and Launching an Instance to create new instances that you then will add to your existing cluster. Be sure to choose the same details you chose when you created the original instances (e.g., VPC and Placement group).

Including New Instances as Cluster Nodes

The **Instances** page lists the instances and their associated IP addresses. You need the IP addresses when you run the `install_vertica` script.

If you are configuring EBS volumes, be sure to configure the volumes on the node before you add the node to your cluster.

To add the new instances as nodes to your existing cluster:

1. Connect to the instance that is assigned to the Elastic IP. See Connecting to an Instance if you need more information.
2. Enter the following command to add the new instances as nodes to your cluster. The following is an example. Substitute the IP addresses for your instances and include your *pem file name. Your instances are added to your existing cluster.

   ```
   sudo /opt/vertica/sbin/install_vertica --add-hosts 10.0.11.166
   --dba-user-password-disabled --point-to-point --data-dir
   /vertica/data --ssh-identity ~/<name-of-pem>.pem
   ```

Adding Nodes and Rebalancing the Database

Once you have added the new instances to your existing cluster, you add them as nodes to your cluster, and then rebalance the database.

Follow the procedure given in the Administration Guide, Adding Nodes to a Database.

Removing Nodes From a Running AWS Cluster

Use these procedures to remove instances/nodes from an AWS cluster.

Preparing to Remove a Node

Removing one or more nodes consists of the following general steps. The first two steps, backing up a database and lowering the k-safety before node removal, are prerequisites for the subsequent steps.

1. Back up the Database. See the section, *Creating Full and Incremental Snapshots (vbr.py)* in the Administrator's Guide.

   HP recommends that you back up the database before performing this significant operation because it entails creating new projections, deleting old projections, and reloading data.

2. Lower the K-safety of your database if the cluster will not be large enough to support its current level of K-safety after you remove nodes. See the section, *Lowering the K-safety Level to Allow for Node Removal* in the Administrator's Guide.

   Note: You cannot remove nodes if your cluster would not have the minimum number of nodes required to maintain your database's current K-safety level (3 nodes for a database with a K-safety level of 1, and 5 nodes for a K-safety level of 2). To remove the node or nodes from the database, you first must reduce the K-safety level of your database.

3. Remove the hosts from the database.
4. Remove the nodes from the cluster if they are not used by any other databases.
5. Optionally, stop the instances within AWS that are no longer included in the cluster.

Removing Hosts From the Database

Before performing the procedure in this section, you must have completed the tasks referenced in Preparing to Remove a Node. The following procedure assumes that you have both backed up your database and lowered the k-safety.

Note: Do not stop the database.

Perform the following to remove a host from the database.

1. While logged on as dbadmin, launch Administration Tools.

   ```
   $ /opt/vertica/bin/admintools
   ```

   Note: Do not remove the host that is attached to your EIP.

2. From the **Main Menu**, select **Advanced Tools Menu**.
3. From **Advanced Menu**, select Cluster Management. Select **OK**.
4. From **Cluster Management**, select **Remove Host(s)**. Select **OK**.
5. From **Select Database**, choose the database from which you plan to remove hosts. Select **OK**.
6. Select the host(s) to remove. Select **OK**.
7. Click **Yes** to confirm removal of the hosts.

   Note: Enter a password if necessary. Leave blank if there is no password.

8. Select **OK**. The system displays a message letting you know that the hosts have been removed. Automatic re-balancing also occurs.
9. Select **OK** to confirm. Administration Tools brings you back to the **Cluster Management** menu.

Removing Nodes From the Cluster

To remove nodes from a cluster, run the `install_vertica` script, specifying the IP addresses of the nodes you are removing and the location and name of your *pem file. (The following example removes only one node from the cluster.)

```
sudo /opt/vertica/sbin/install_vertica --remove-hosts 10.0.11.165
--point-to-point --ssh-identity ~/<name-of-pem>.pem --dba-user-
password-disabled
```

Stopping the AWS Instances (Optional)

Once you have removed one or more nodes from your cluster, to save costs associated with running instances, you can choose to stop or terminate the AWS instances that were previously part of your cluster. This step is optional because, once you have removed the node from your HP Vertica cluster, HP Vertica no longer sees the instance/node as part of the cluster even though it is still running within AWS.

To stop an instance in AWS:

1. On AWS, navigate to your **Instances** page.
2. Right-click on the instance, and choose **Stop**.

Migrating Data Between AWS Clusters

This section provides guidance for copying (importing) data from another AWS cluster, or exporting data between AWS clusters.

There are three common issues that occur when exporting or copying on AWS clusters. The issues are listed below. Except for these specific issues as they relate to AWS, copying and exporting data works as documented in the Administrator's Guide section, Copying and Exporting Data.

Issue 1. Ensure that all nodes in source and destination clusters have their own elastic IPs (or public IPs) assigned.

Each node in one cluster must be able to communicate with each node in the other cluster. Thus, each source and destination node needs an elastic IP (or public IP) assigned.

Issue 2. Set the parameter DontCheckNetworkAddress to true.

On AWS, when creating a network interface, you receive an error if you attempt to assign the elastic IP to an AWS node (example uses a sample elastic IP address):

```
dbadmin=> CREATE NETWORK INTERFACE eipinterface ON v_tpch_
node0001 with '107.23.151.10';

ERROR 4125: No valid address found for [107.23.151.10] on this
node
```

This error occurs because the elastic IP is the public IP and not the private IP of the target node. To resolve this issue, first set the parameter `DontCheckNetworkAddress` to true:

```
select set_config_parameter('DontCheckNetworkAddress','1');
```

You can find information on the CREATE NETWORK INTERFACE statement and SET_CONFIG_PARAMETER in the SQL Reference Manual.

Issue 3. Ensure your security group allows the AWS clusters to communicate.

Check your security groups for both your source and destination AWS clusters. Ensure that ports 5433 and 5434 are open.

If one of your AWS clusters is on a separate VPC, ensure that your network access control list (ACL) allows communication on port 5434.

Note: This communication method exports and copies (imports) data through the internet. You can alternatively use non-public IPs and gateways, or VPN to connect the source and destination clusters.

Migrating to HP Vertica 7.0 or later on AWS

Note: If you had an HP Vertica installation running on AWS prior to Release 6.1.x, you can migrate to HP Vertica 7.1.x or later using a new preconfigured AMI.

For more information, see the Solutions tab of the myVertica portal.

Upgrading to the version 7.0 HP Vertica AMI on AWS

Use these procedures for upgrading to the latest HP Vertica AMI. The procedures assume that you have a 6.1.x or later cluster successfully configured and running on AWS. If you are setting up an HP Vertica cluster on AWS for the first time, follow the detailed procedure for installing and running HP Vertica on AWS.

Note: Both `install_vertica` and `update_vertica` use the same parameters.

Preparing to Upgrade Your AMI

Perform this procedure to prepare for the upgrade to the latest HP Vertica AMI.

1. Back up your existing database. See *Backing Up and Restoring the Database* in the Administrator's Guide.
2. Download the HP Vertica install package. See *Download and Install the HP Vertica Install Package* in the Installation Guide.

Upgrading HP Vertica Running on AWS

HP Vertica supports upgrades of Vertica Server running on AWS instances created from the Vertica AMI. To upgrade HP Vertica, follow the instructions provided in the HP Vertica upgrade documentation.

Troubleshooting: Checking Open Ports Manually

You originally configured your security group through the AWS interface. Once your cluster is up and running, you can check ports manually through the command line using the netcat (nc) utility. What follows is an example using the utility to check ports.

Before performing the procedure, choose the private IP addresses of two nodes in your cluster.

The examples given below use nodes with the private IPs:

```
10.0.11.60 10.0.11.61
```

Using the Netcat (nc) Utility

After installing the nc utility on your nodes, you can issue commands to check the ports on one node from another node.

1. To check a TCP port:

   a. Put one node in listen mode and specify the port. In the following sample, we're putting IP `10.0.11.60` into listen mode for port `4804`.

   ```
   [root@ip-10-0-11-60 ~]# nc -l 4804
   ```

   b. From the other node, run nc specifying the IP address of the node you just put in listen mode, and the same port number.

   ```
   [root@ip-10-0-11-61 ~]# nc 10.0.11.60 4804
   ```

   c. Enter sample text from either node and it should show up on the other. To cancel once you have checked a port, enter **Ctrl+C**.

Note: To check a UDP port, use the same nc commands with the –u option.

```
[root@ip-10-0-11-60 ~]# nc -u -l 4804
[root@ip-10-0-11-61 ~]# nc -u 10.0.11.60 4804
```

Quick Start to Setting Up HP Vertica AWS

This topic presents a summary of the detailed procedures included in this document. The procedures require basic knowledge of the AWS Management Console. For information about working with AWS Management Console, refer to the AWS documentation, or use the detailed procedures.

From the AWS Management Console:

1. Choose a region, and create and name a Placement Group.
2. Create and name a Key Pair.
3. Create a VPC.
    a. Edit the public subnet; change according to your planned system set-up.
    b. Choose an availability zone.
4. Create an Internet Gateway and assign it to your VPC (or use the default gateway).
5. Create a Security Group for your VPC (or use the default security group, but add rules).

    Inbound rules to add:

    - HTTP
    - Custom ICMP Rule: Echo Reply
    - Custom ICMP Rule: Traceroute
    - SSH
    - HTTPS
    - Custom TCP Rule: 4803-4805
    - Custom UDP Rule: 4803-4805
    - Custom TCP Rule: (Port Range) 5433

- Custom TCP Rule: (Port Range) 5434
- Custom TCP Rule: (Port Range) 5434
- Custom TCP Rule: (Port Range) 5450

Outbound rules to add:

- All TCP: (Destination) 0.0.0.0/0
- All ICMP: (Destination) 0.0.0.0/0

6. Create instances.
   a. Select **Launch Instance**.
   b. From Community AMIs, choose an HP Vertica AMI.
   c. Select the **Compute optimized** tab and select a supported instance type.
   d. Choose the number of instances, network, and placement group.
   e. Select a security group.
   f. Click **Launch** and choose your key pair.
   g. Click **Launch Instances**.
7. Assign an elastic IP to an instance.
8. Connect to the instance that is attached to the elastic IP.
9. Run the `install_vertica` script after placing the *pem file on your Amazon instance.

Send Documentation Feedback

If you have comments about this document, you can contact the documentation team by email. If an email client is configured on this system, click the link above and an email window opens with the following information in the subject line:

Feedback on HP Vertica on Amazon Web Services Guide (Vertica Analytic Database 7.1.x)

Just add your feedback to the email and click send.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to vertica-docfeedback@hp.com.

We appreciate your feedback!