Progress®
Kemp® LoadMaster®

Installation Guide LoadMaster HA for AWS

Generated on: 18 December 2024

Copyright

Visit the following page online to see Progress Software Corporation's current Product Documentation Copyright Notice/Trademark Legend: Product Documentation Copyright Notice & Trademarks | Progress

Table of Contents

Chapter 1: Introduction 5
Document Purpose 6
Intended Audience 6
Prerequisites 6

Chapter 2: AWS Network Load Balancing Service Architecture 7

Chapter 3: Terminology Differences 11

Chapter 4: Using LoadMaster HA for AWS 12

Chapter 5: Configure GEO Clusters with HA 14
Configure GEO Clusters with HA IP Addresses 14

Chapter 6: Creating AWS HA Pairs 19
Create the Network Load Balancer in AWS 19
Configure the LoadMaster 23
Virtual Service Restrictions 25

Chapter 7: LoadMaster Firmware Upgrades/Downgrades 26

Chapter 8: Best Practices for Backups 28

Chapter 9: First/Second Unconnected 29

Chapter 10: References 30

1

Introduction

The LoadMaster system can be deployed as a single unit or in an active/standby dual-unit configuration (High Availability (HA)). HA allows two physical or virtual machines to become one logical device. Only one of these units is active and handling traffic at any one time. One unit is active and the other is a hot standby (passive). This provides redundancy and resiliency, meaning if one LoadMaster goes down for any reason, the hot standby can become active, therefore minimizing any downtime.

The AWS Network Load Balancer is used to achieve HA in AWS when using LoadMasters. The Network Load Balancer does not leverage multiplexing therefore the LoadMaster persistence options can be enabled.

Network Load Balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances in the cloud. Network Load Balancing ensures that only healthy Amazon EC2 instances receive traffic by detecting unhealthy instances and re-routing traffic across the remaining healthy instances.

Placing the LoadMasters behind the Network Load Balancer enables advanced application delivery functionality.

Related Links

For details, see the following topics:

- Document Purpose
- Intended Audience
- Prerequisites

Document Purpose

The purpose of this document is to provide information and step-by-step instructions on how to configure HA when using the LoadMaster in AWS.

Intended Audience

This document is intended to be read by anyone who is interested in finding out how to configure HA when using the LoadMaster in an AWS environment.

Prerequisites

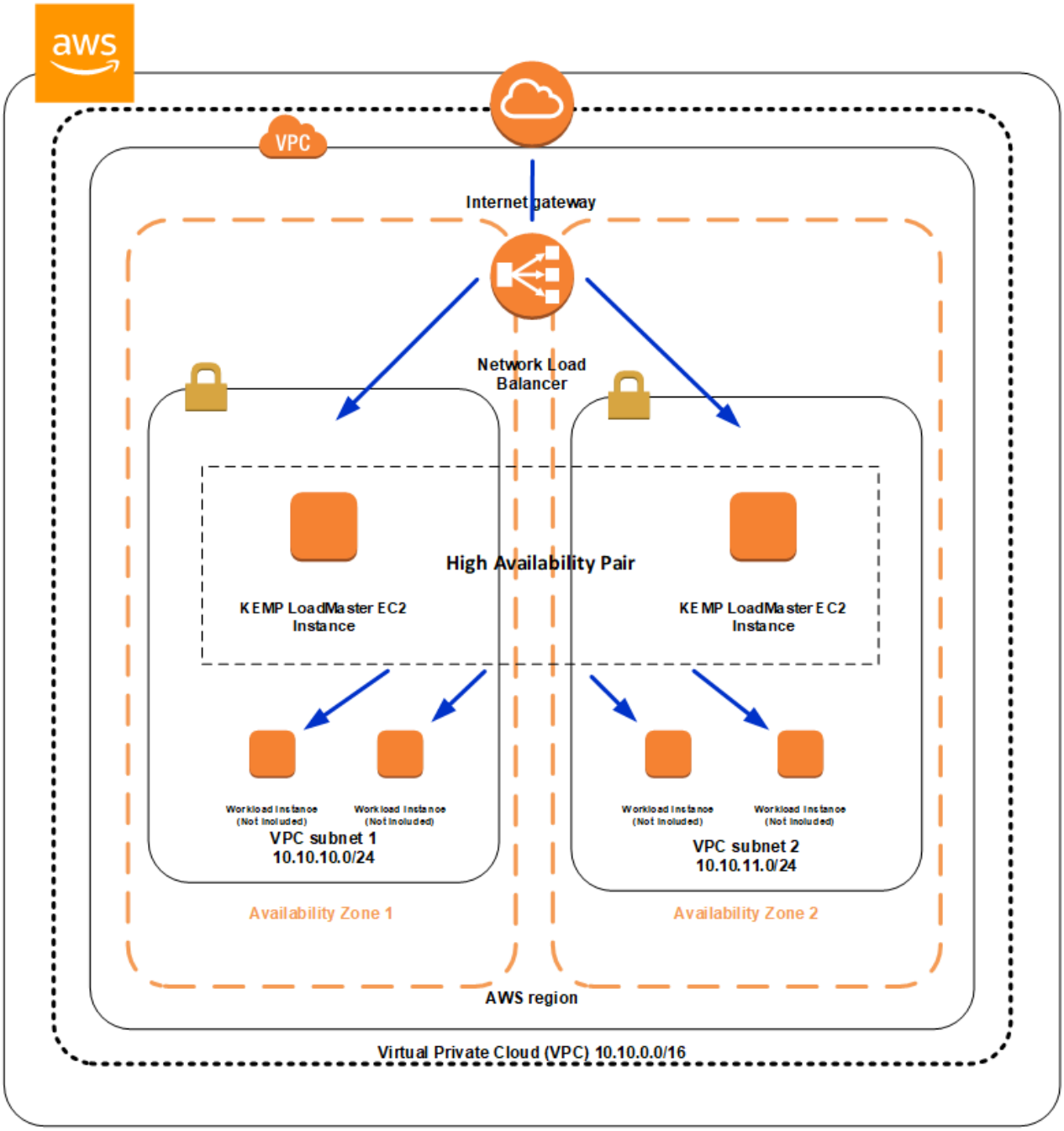
This document assumes that you already have two LoadMaster HA instances which are configured and accessible using the User Interface (UI). For instructions on how to do this, refer to the **LoadMaster for AWS Feature Description** on the Documentation Page. One should be designated as active and the other as a standby.

Note: Due to AWS limitations, it is not possible to bond interfaces on AWS LoadMasters.

For step-by-step instructions on how to deploy a LoadMaster in AWS, refer to the LoadMaster for AWS Installation Guide.

2

AWS Network Load Balancing Service Architecture



There are two logical components in the Network Load Balancing service architecture:

- Load balancers
- A controller service

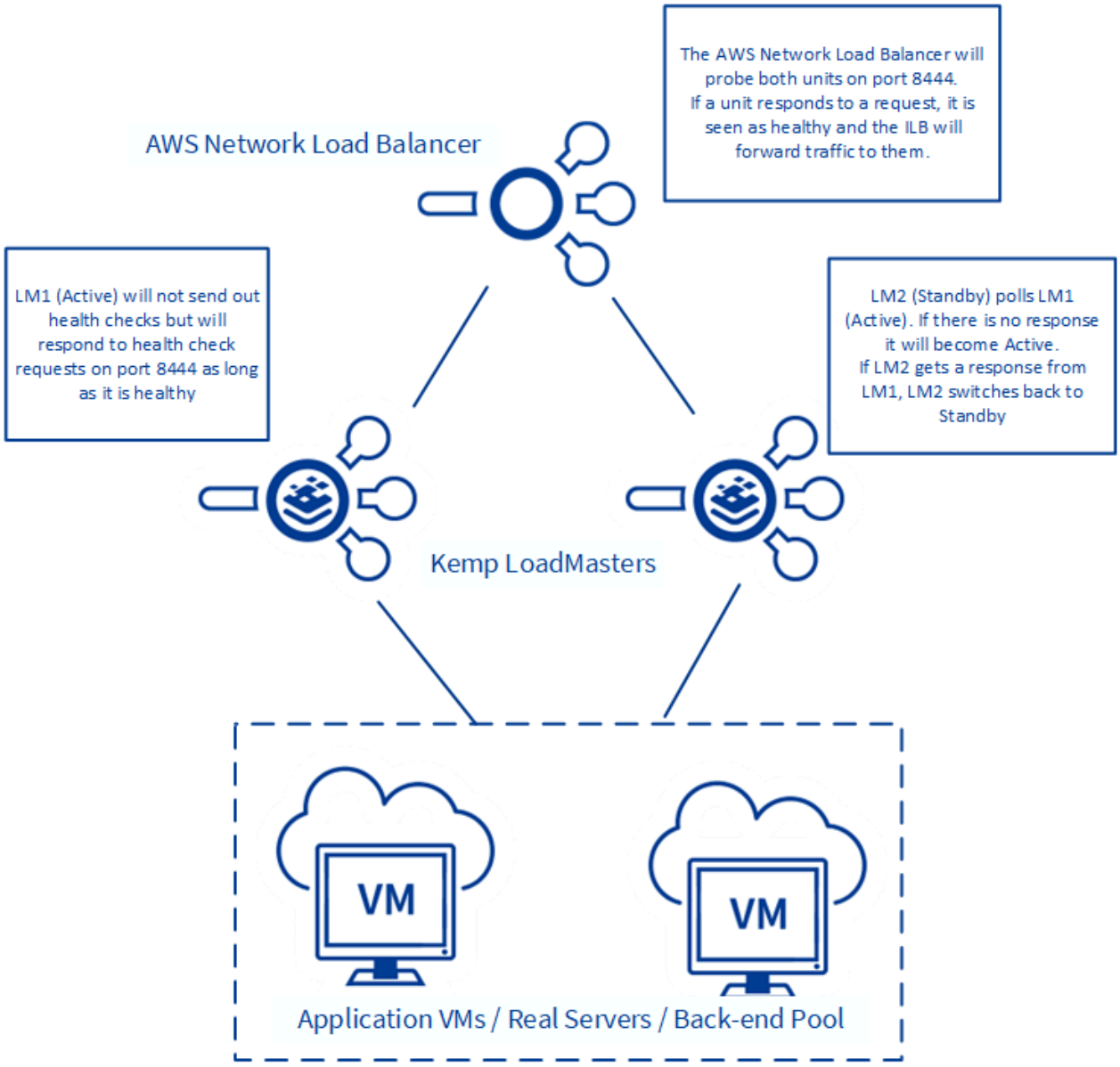
The load balancers are resources that monitor traffic and handle requests that come in through the Internet, that is, the LoadMaster.

The controller service monitors the load balancers and verifies that load balancers are behaving properly.

Once you create a network load balancer, you must configure it to accept incoming traffic and route requests to your EC2 instances. These configuration parameters are stored by the controller, and the controller ensures that all of the load balancers are operating with the correct configuration.

Network Load Balancing will perform health checks on back-end instances, using the configuration that you supply.

To discover the availability of your EC2 instances, the load balancer periodically sends pings, attempts connections, or sends requests to test the EC2 instances. These tests are called health checks. Instances that are healthy at the time of the health check are marked as **InService** and the instances that are unhealthy at the time of the health check are marked as **OutOfService**. The load balancer performs health checks on all registered instances, whether the instance is in a healthy state or an unhealthy state. When using AWS VLMs in HA mode – one unit is active and in service, the other is stand-by and out-of-service.



The load balancer routes traffic only to the healthy instances. When the load balancer determines that an

instance is unhealthy, it stops routing traffic to that instance. The load balancer resumes routing traffic to the instance when it has been restored to a healthy state.

The load balancer checks the health of the registered instances using either the default health check configuration provided by Network Load Balancing or a health check configuration that you configure.

The health checks must reach the defined target set in the Network Load Balancing configuration for the number of successful checks before the instance is considered to be “in service” and healthy. For example, for any instance registered with Network Load Balancing - if you set the interval for health checks to 20 seconds, and you set the number of successful health checks to 10, then it will take at least 200 seconds before Network Load Balancing will route traffic to the instance.

The health check also defines a failure threshold. For example, if you set the interval to 20 seconds and you set the failure threshold at 4, then when an instance no longer responds to requests - at least 80 seconds will elapse before it is taken out of service. However, if an instance is terminated, traffic will no longer be sent to the terminated instance, but there can be a delay before the load balancer is aware that the instance was terminated. For this reason, it is important to de-register your instances before terminating them; instances are removed from service in a much shorter amount of time if they are de-registered.

3

Terminology Differences

There are some terminology differences between "normal" (non-cloud) High Availability (HA) LoadMaster units and cloud HA units. The table below outlines these differences:

Platform	Unit number	Terminology	Statuses
Non-cloud	Unit 1	HA first	Master/Standby
Non-cloud	Unit 2	HA second	Master/Standby
Cloud	Unit 1	Active	Active/Standby
Cloud	Unit 2	Standby	Active/Standby

4

Using LoadMaster HA for AWS

When using LoadMaster in High Availability on AWS, HA operates in much the same way as it does on non-cloud platforms, but with some key differences due to how HA interacts with the AWS Elastic IP feature:

- LoadMaster HA for AWS involves two LoadMasters that synchronize. Changes made to the active LoadMaster are replicated to the standby LoadMaster.
- When synchronizing the GEO settings from active to standby, any Fully Qualified Domain Name (FQDN) or cluster IP addresses that match the active's IP address are replaced with the standby's IP address. Likewise, when synchronizing from standby to active, the standby's IP address is replaced with the active's IP address.
- All user-defined settings are synchronized, with the exception of the following:
  - Default gateway (both IPv4 and IPv6)
  - IP addresses and netmasks
  - Hostname
  - Name server
  - Domain
  - Admin default gateway
  - Administrative certificate settings
  - Network interface settings: Link Status (Speed and Duplex), MTU and additional addresses
  - Virtual LAN (VLAN) configuration
  - Virtual Extensible LAN (VXLAN) configuration
  - Additional routes
- The cloud HA LoadMaster does not have a "force update" option.

- Both devices are capable of responding to Network Load Balancer health check requests.
  - The LoadMaster that is currently handling client traffic will respond with the status code **200 OK** to the AWS health check - meaning that it is healthy. Meanwhile, the standby LoadMaster will respond with the status code **503** -- meaning that it is unhealthy. In this way, all client requests are redirected by the Network Load Balancer to the healthy LoadMaster.
  - The “standby” LoadMaster (the LoadMaster which is not handling traffic) polls the “active” LoadMaster to check the availability of the service. If the probe is successful, it remains in “standby” mode, otherwise it takes over as the “active” - answering 200 OK to the AWS health check becoming capable to handle traffic.

Note: If the active unit fails, connections are directed to the standby unit. The active unit never assumes the standby role and the standby never becomes active. When the active unit becomes available again after a failure, connections are automatically directed to the active unit again. The active unit can be active or standby. The standby unit can be active or standby.You can set the **Switch to Preferred Server** option to **No Preferred Host** which allows the standby unit to maintain the active state when the active unit comes back online.For HA to work, the two LoadMasters must have different values set for the **AWS HA Mode**.We recommend always using NIC0 for HA checks on AWS.

A complete description of non-cloud LoadMaster HA can be found in the High Availability (HA), Feature Description document.

5

Configure GEO Clusters with HA

If Global Server Load Balancing (GSLB)/GEO is being used in the environment, there are unique configuration requirements for setting up clusters. This section outlines the configuration and the steps can be modified based on the environment.

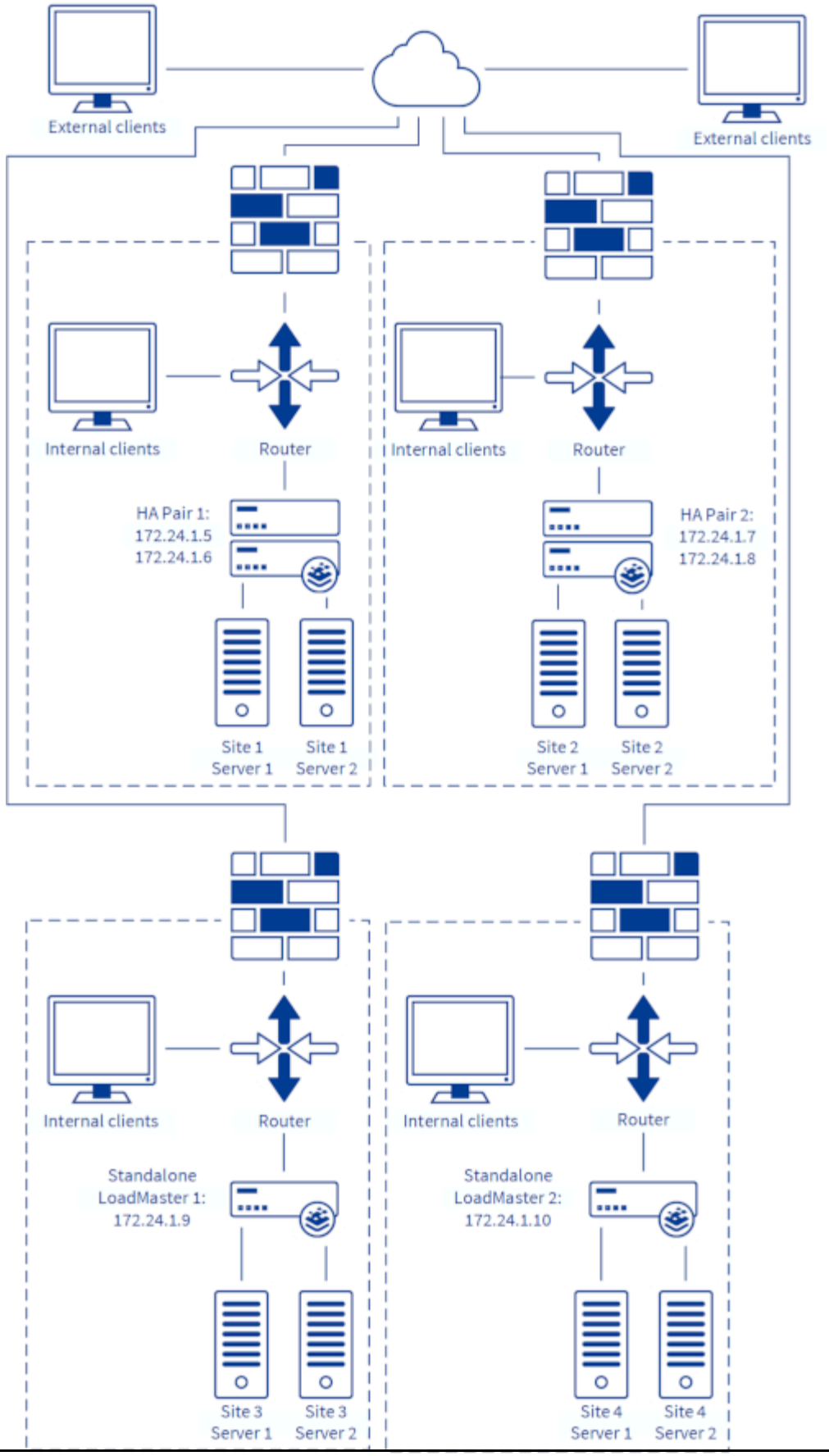
Related Links

For details, see the following topics:

- Configure GEO Clusters with HA IP Addresses

Configure GEO Clusters with HA IP Addresses

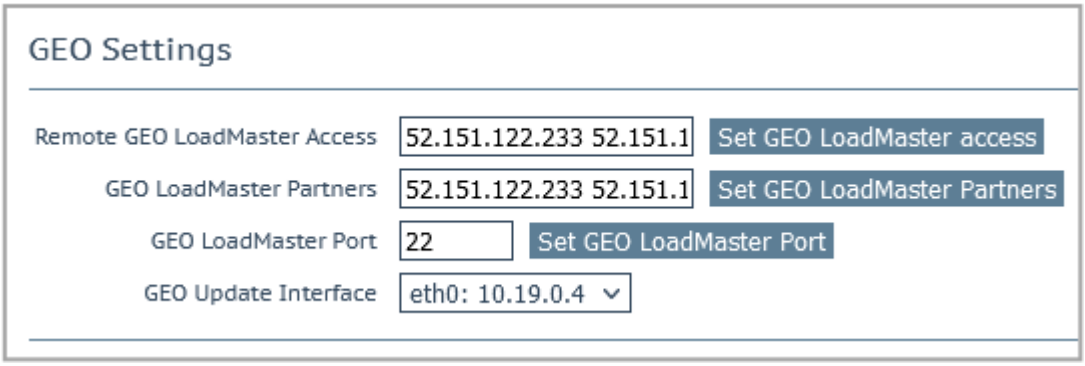
In this example, there will be two HA Pairs of LoadMasters and two standalone LoadMasters included in the GSLB environment. The steps will be the same if more or less units are in the environment.


External clients
External clients
Internal clients
Router
Internal clients
Router
HA Pair 1:
172.24.1.5
172.24.1.6
HA Pair 2:
172.24.1.7
172.24.1.8
Site 1
Server 1
Site 1
Server 2
Site 2
Server 1
Site 2
Server 2
Internal clients
Router
Internal clients
Router
Standalone
LoadMaster 1:
172.24.1.9
Standalone
LoadMaster 2:
172.24.1.10
Site 3
Server 1
Site 3
Server 2
Site 4
Server 1
Site 4
Server 2

To configure clusters of HA IP addresses between two HA pairs and two standalone systems, follow the steps below.

On HA pair 1, complete these steps:

1. Log into the Active LoadMaster in the HA pair.
2. In the main menu, go to **Certificates & Security > Remote Access**.



3. Enter the IP addresses of the ILBs in front of the HA pair of LoadMasters and the IP addresses of both standalone LoadMasters in the **Remote GEO LoadMaster Access** text box and click **Set GEO LoadMaster access**.
4. Enter the IP addresses of the ILBs in front of the HA pair of LoadMasters and the IP addresses of both standalone LoadMasters in the **GEO LoadMaster Partners** text box and click **Set GEO LoadMaster Partners**.

   Note: For the **GEO LoadMaster Partners** field - you do not need to specify the IP address (ILB IP Address) of the current LoadMaster you are connected to.

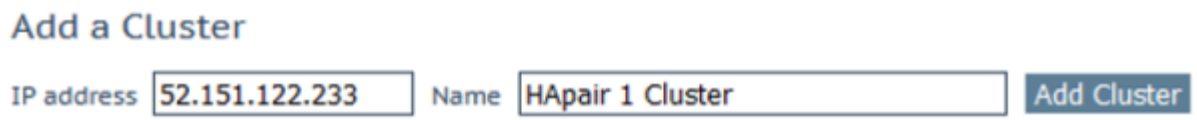
5. Reboot the Active LoadMaster.
6. Log into the new active LoadMaster.
7. In the main menu, go to **Certificates & Security > Remote Access**.
8. Re-enter the IP addresses of the ILBs in front of the HA pair of LoadMasters and the IP addresses of both standalone LoadMasters in the **Remote GEO LoadMaster Access** text box and click **Set GEO LoadMaster access**.
9. (Optional) Reboot the system again to get the original LoadMaster as active.

Repeat these steps on HA pair 2 and on the two standalone LoadMasters.

Note: Ensure the IP addresses used here are listed in the Network Security Groups for the inbound rules of TCP/22 if access is limited to specific source IP addresses.

Then, create the GEO clusters with the **Type Remote LM** using the IP addresses of the ILBs in front of the HA pairs and the IP addresses of both standalone LoadMasters. This only needs to be performed on one system as changes will be replicated to the other partner systems configured in the prior step.

1. In the main menu, go to **Global Balancing > Manage Clusters**.



2. In the **IP address** text box, enter the LoadMaster IP address.
3. Enter a **Name** for the cluster and click **Add Cluster**.
4. Click **Modify** on the relevant cluster.



5. Select **Remote LM** in the **Type** drop-down list.

Repeat these steps to add the remaining HA pair and on the two standalone LoadMasters.

Here are some example IP addresses for the scenario outlined above (involving two HA pairs and two standalone systems).

Unit	Number	IP Address
HA pair 1	LM 1	10.19.0.4
HA pair 1	LM 2	10.19.0.5
HA pair 1 ILB		52.151.122.233
HA pair 2	LM 1	10.20.0.4
HA pair 2	LM 2	10.20.0.5
HA pair 2 ILB		52.151.124.198
Standalone unit 1		20.0.54.164
Standalone unit 2		20.0.51.33

Here is an example configuration based on the scenario outlined above (involving two HA pairs and two standalone systems).

Unit	WUI Field	WUI Field Value
HA pair 1	**Remote GEO LoadMaster Access**	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33
	GEO LoadMaster Partners	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33
HA pair 2	**Remote GEO LoadMaster Access**	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33

Unit	WUI Field	WUI Field Value
	GEO LoadMaster Partners	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33
Standalone unit 1	**Remote GEO LoadMaster Access**	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33
	GEO LoadMaster Partners	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33
Standalone unit 2	**Remote GEO LoadMaster Access**	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33
	GEO LoadMaster Partners	52.151.122.233 52.151.124.198 20.0.54.164 20.0.51.33

The IP addresses colored as red above correspond to the particular unit mentioned in the **Unit** column.

Note: For the **GEO LoadMaster Partners** field - you do not need to specify the red IP address but you can include it because it is easier for management purposes.

6

Creating AWS HA Pairs

This document assumes that you already have two LoadMaster HA instances which are configured and accessible using the User Interface (UI). One should be designated as active and the other as a standby.

Note: TCP port 6973 must be allowed in the inbound rules for each LoadMaster in the HA pair to allow for synchronization traffic.

For further information and steps on how to deploy an individual LoadMaster instance, refer to the LoadMaster for AWS Installation Guide document.

There are options when choosing a Load Balancer in AWS to support the LoadMaster HA Pair. This document covers the deployment of a Network Load Balancer to allow for different persistence options on the LoadMaster. You can choose the Classic Load Balancer if persistence is not required on the LoadMaster.

Related Links

For details, see the following topics:

- Create the Network Load Balancer in AWS
- Configure the LoadMaster

Create the Network Load Balancer in AWS

Note: Before creating a Network Load Balancer in AWS, it is best practice to have an existing Elastic IP

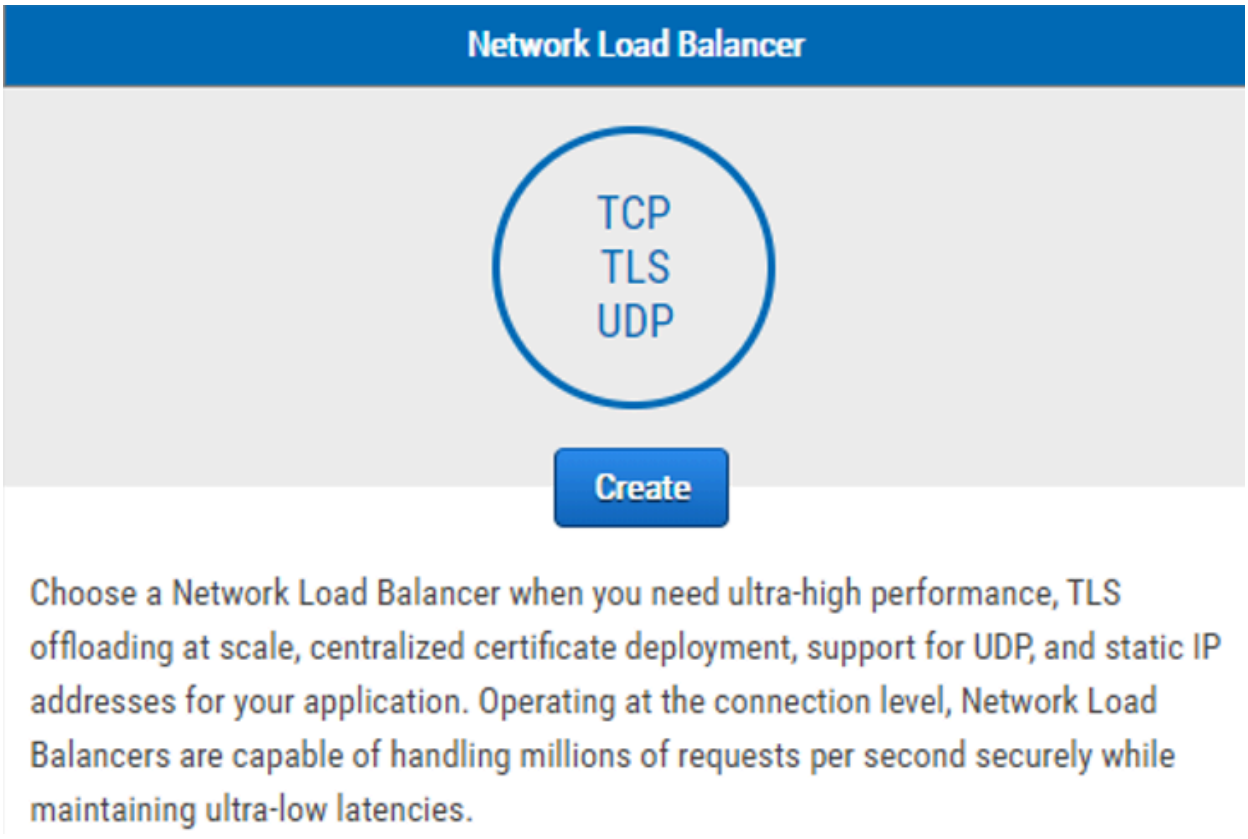
(EIP) allocated. To find out how to allocate an Elastic IP address, refer to the following AWS link: Elastic IP addresses.

To create AWS HA pairs, carry out the following steps:

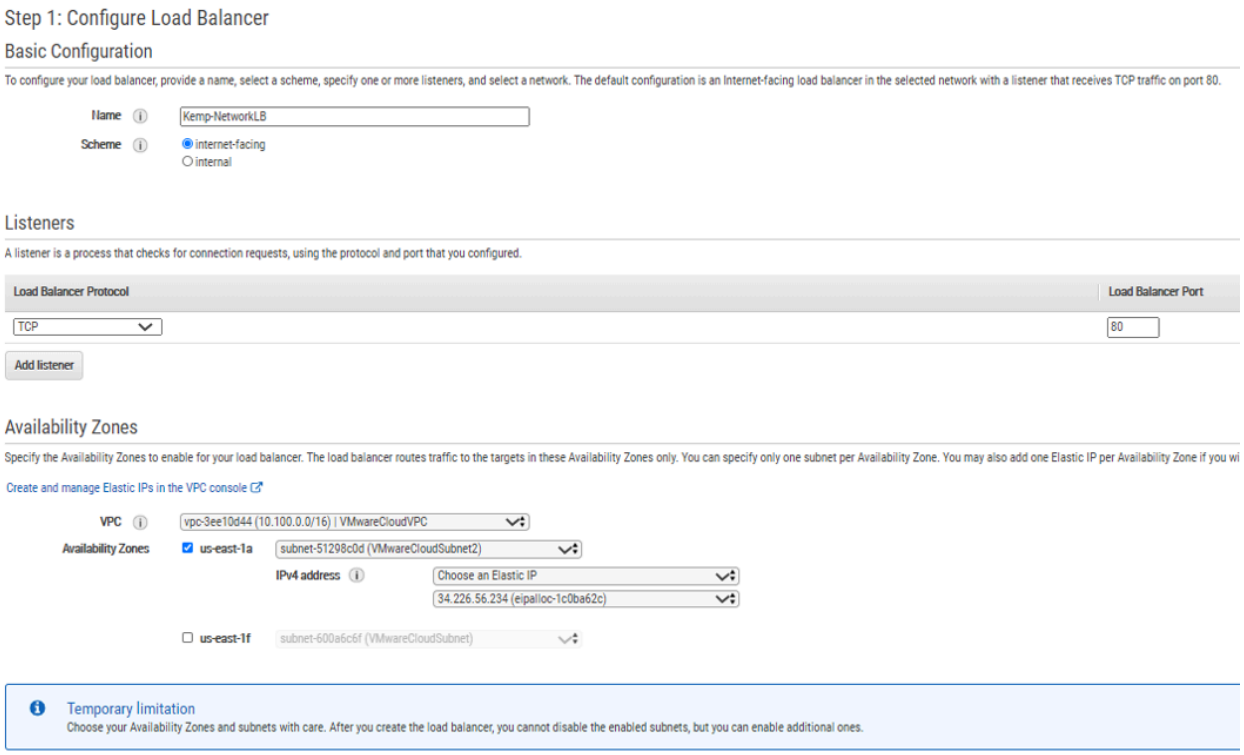
1. Open the Amazon EC2 console.
2. Navigate to **Load Balancing > Load Balancers**.



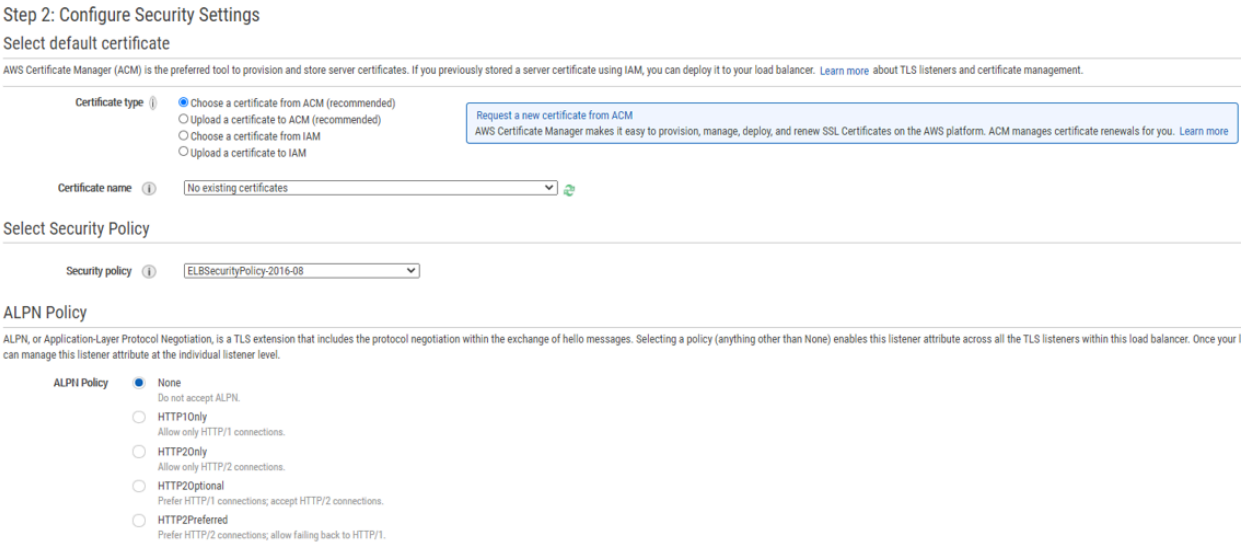
3. Click **Create Load Balancer**.



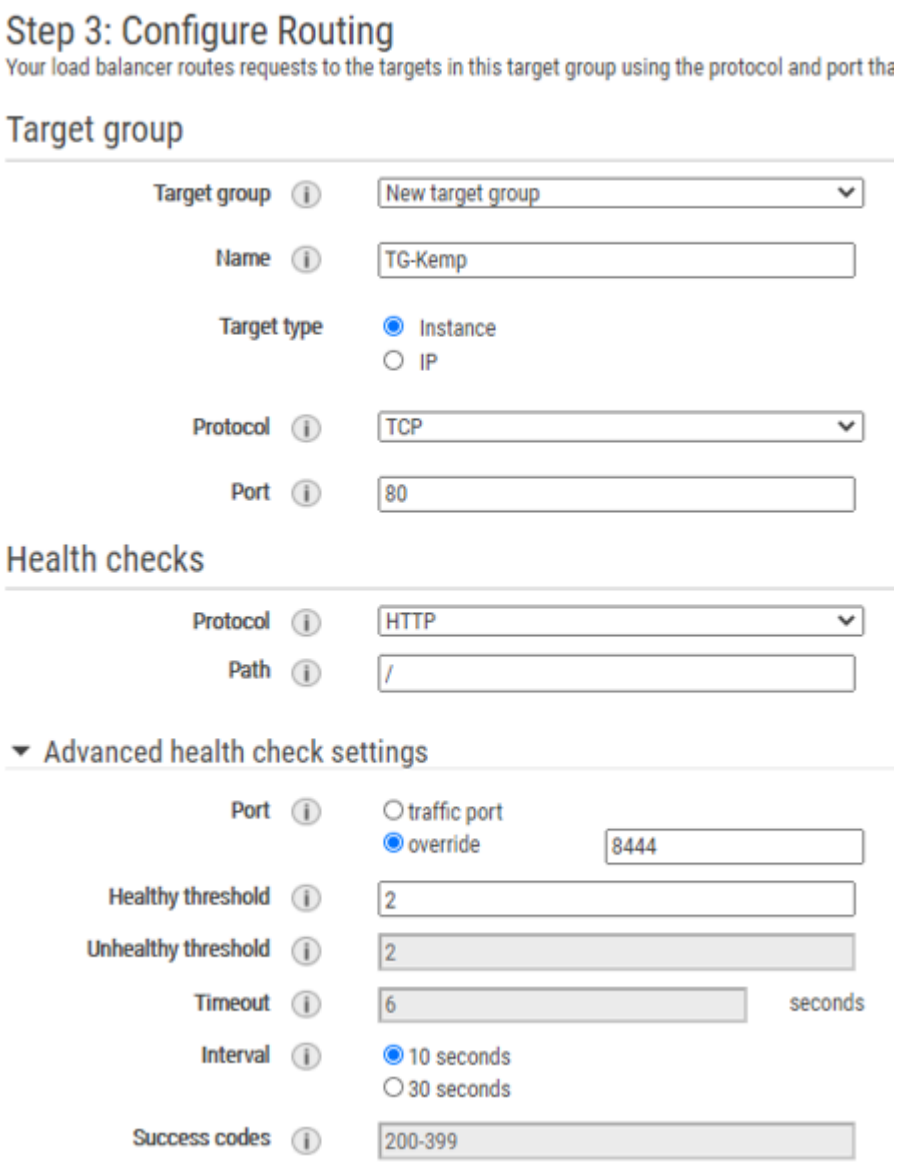
4. Click **Create** for **Network Load Balancer**.



5. Set the following options:
   1. Enter a **Name** for the Load Balancer.
   2. Select whether to provide access to **internal** only or allow for an **internet-facing** load balancer.
   3. Add one or more **Listeners** based on the application being load balanced.
   4. Select the **VPC** and **Availability Zones** to deploy the Network Load Balancer in.
   5. Select **Choose an Elastic IP** and select the EIP that was allocated earlier.
   6. Add optional **Tags**.
6. Click **Next: Configuration Security Settings**.



7. If you selected a TLS Listener, you can create or upload a certificate to encrypt the traffic.
8. Click **Next: Configure Routing**.



9. Set the following options:
   1. Select a new **Target group**.
   2. Enter a **Name** for the target group.
   3. Select **Instance** for **Target type**.
   4. Select the **Protocol** and **Port** for routing traffic to the LoadMasters.
   5. Select **HTTP** as the health check **Protocol**.
   6. Enter **/** for the health check **Path**.
   7. Under **Advanced health check settings**, select **override** for **Port**.
   8. Enter **8444** for the **override** port.
   9. Enter **2** as the **Healthy threshold**.

10. Select **10 seconds** as the **Interval**.

10. Click **Next: Register Targets**.

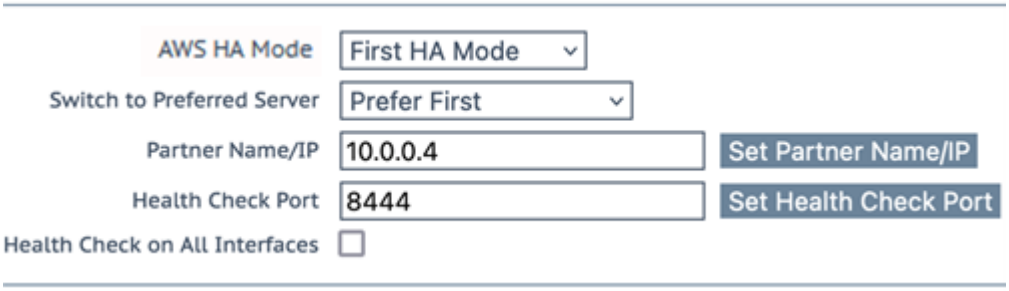


11. Select the two LoadMasters in the HA pair and click **Add to registered**.
12. Click **Next: Review**.
13. Click **Create**.

Configure the LoadMaster

Complete the following steps to configure the LoadMaster settings:

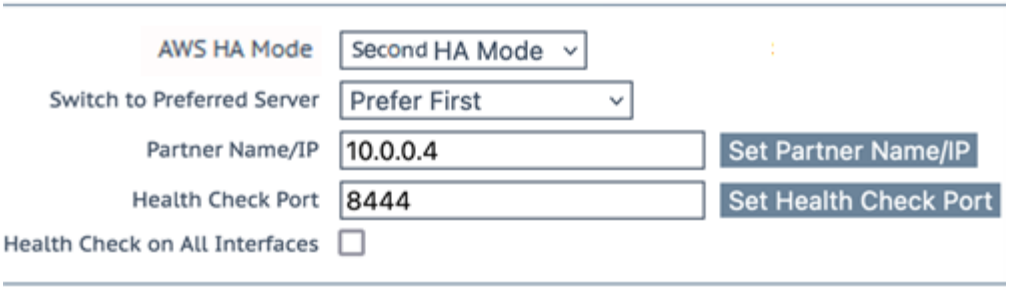
1. Log in to the UI of the active LoadMaster.
2. In the main menu, go to **System Configuration > AWS HA Parameters**.



3. Select **First HA Mode** from the **AWS HA Mode** drop-down list.
4. Select the desired option in the **Switch to Preferred Server** drop-down list:
   - **No Preferred Host:** Each unit takes over when the other unit fails. No switchover is performed when the partner is restarted.
   - **Prefer First:** The HA1 (active) unit always takes over. This is the default option.
5. Enter the IP address of the standby LoadMaster in the **Partner Name/IP** text box and click **Set Partner Name/IP**.
6. Enter the health check port selected earlier in the **Health Check Port** text box and click **Set Health Check Port**.
7. If using a multi-arm configuration, select the **Health Check on All Interfaces** check box.

Note: If this option is disabled, the health check listens on the primary eth0 address.

8. Log in to the UI of the standby LoadMaster.
9. In the main menu, go to **AWS HA Parameters**.



10. Select **Second HA Mode** from the **AWS HA Mode** drop-down list.
11. Enter the IP address of the active LoadMaster in the **Partner Name/IP** text box and click **Set Partner Name/IP**.
12. Enter the health check port selected earlier in the **Health Check Port** text box and click **Set Health Check Port**.

    Note: The **Health Check Port** must be the same on both the active and standby units in order for HA to function correctly.

13. If using a multi-arm configuration, select the **Health Check on All Interfaces** check box.

Note: If this option is disabled, the health check listens on the primary eth0 address.

In the Amazon EC2 console, go to the ELB and select the **Instances** tab. The active instance should be marked as **InService**. The standby instance should be marked as **OutOfService**.

In the LoadMaster, set up a HTTP and HTTPS Virtual Service with Real Servers. These should then be available using the ELB Domain and they should properly fail over.



If a unit is in standby mode, WUI access is restricted to **Local Administration** only. Full WUI access is available if the unit is in an active or unchecked state.

Virtual Service Restrictions

There are some situations where Virtual Service settings may prevent HA from functioning correctly. Please follow the guidelines below to avoid any issues:

- Do not set up a Virtual Service on the same port as the health check port
- Do not set up a TCP Virtual Service on port 6973 on the interface where HA sync is configured
- Do not set up a TCP Virtual Service on port 22 on a LoadMaster interface port

7

LoadMaster Firmware Upgrades/ Downgrades

Do not downgrade from firmware version 7.2.36 or higher to a version below 7.2.36. If you do this, the LoadMaster becomes inaccessible and you cannot recover it.

You should never leave two LoadMasters with different firmware versions paired as HA in a production environment. To avoid complications, follow the steps below in sequence and do not perform any other actions in between the steps. Please upgrade/downgrade during a maintenance window and expect service disruption because there are reboots.

The steps below are high-level, for detailed step-by-step instructions on how to upgrade the LoadMaster firmware, refer to the Updating the LoadMaster Software Feature Description on the documentation page: Updating the LoadMaster Software Technical Note.

Upgrade the LoadMaster Firmware

To upgrade the LoadMaster firmware with the least disruption, follow the steps below in sequence:

1. Identify the STAND-BY unit.
2. Upgrade the LoadMaster firmware on the STAND-BY unit. Once the STAND-BY unit has rebooted, it remains in the STAND-BY state and the WUI is limited to the Local Administration options.
3. Upgrade the LoadMaster firmware on the ACTIVE unit. When the ACTIVE unit is rebooting, the STAND-BY unit becomes ACTIVE.
4. Depending on Preferred Host settings in the HA configuration, the Standby unit may failback over to the Active unit.

After these steps are completed the upgrade is finished.

Downgrade the LoadMaster Firmware

To downgrade the LoadMaster firmware with the least disruption, follow the steps below in sequence:

1. Identify the STAND-BY unit.
2. Downgrade the LoadMaster firmware on the STAND-BY unit. Once the STANDY-BY unit has rebooted, it remains in the STAND-BY state and the WUI is limited to the Local Administration options.
3. Downgrade the LoadMaster firmware on the ACTIVE unit. When the ACTIVE unit is rebooting, the STAND-BY unit becomes ACTIVE.
4. Depending on Preferred Host settings in the HA configuration, the Standby unit may failback over to the Active unit.

After these steps are completed the downgrade is finished.

8

Best Practices for Backups

Hypervisor snapshots cannot be used to restore a LoadMaster to a working state. The best way to back up your LoadMaster settings is by using the native backup and restore facility in the LoadMaster WUI or API.

To back up your LoadMaster configuration, follow these steps:

1. In the main menu, go to **System Configuration > System Administration > Backup/Restore**.
2. Click **Create Backup File**.

You can create a remote host for automated backups using SCP to save backups to a remote server.

For further details on backing up and restoring the LoadMaster configuration, including certificates and cipher sets, refer to the following links:

- Backup and Restore Technical Note
- How to Create and Restore a LoadMaster Configuration or Certificate Backup

9

First/Second Unconnected

When initially setting up cloud HA, the active unit should have **First** in the top-right corner of the LoadMaster WUI.

The standby unit should show **Second**.

After setting up the load balancer (Internal Load Balancer (ILB) for Azure or Network Load Balancer for AWS) the units should switch from:

- First to First Unconnected
- Second to Second Unconnected

This means the LoadMasters have not been polled by the load balancer. Once the load balancer has the health check correctly set, the units should switch from:

- First Unconnected to First (Active)/First (Standby)
- Second (Unconnected) to Second (Active)/Second (Standby)

10

References

Unless otherwise specified, the following documents can be found at https://docs.progress.com/.

LoadMaster for AWS, Installation Guide

High Availability (HA), Feature Description