

High Availability Options

- High Availability Options, on page 1
- VPN Load Balancing, on page 2

High Availability Options

Distributed VPN Clustering, Load balancing and Failover are high-availability features that function differently and have different requirements. In some circumstances you may use multiple capabilities in your deployment. The following sections describe these features. Refer to the appropriate release of the ASA General Operations ASDM Configuration Guide for details on Distributed VPN and Failover. Load Balancing details are included here.

VPN and Clustering on the Secure Firewall eXtensible Operating System (FXOS) Chassis

An ASA FXOS Cluster supports one of two mutually exclusive modes for S2S VPN, centralized or distributed:

- Centralized VPN Mode. The default mode. In centralized mode, VPN connections are established with the control unit of the cluster only.

  VPN functionality is limited to the control unit and does not take advantage of the cluster high availability capabilities. If the control unit fails, all existing VPN connections are lost, and VPN connected users see a disruption in service. When a new control unit is elected, you must reestablish the VPN connections.

  When you connect a VPN tunnel to a Spanned interface address, connections are automatically forwarded to the control unit. VPN-related keys and certificates are replicated to all units.

- Distributed VPN Mode. In this mode, S2S IPsec IKEv2 VPN connections are distributed across members of an ASA cluster providing scalability. Distributing VPN connections across the members of a cluster allows both the capacity and throughput of the cluster to be fully utilized, significantly scaling VPN support beyond Centralized VPN capabilities.

Note Centralized VPN clustering mode supports S2S IKEv1 and S2S IKEv2.

Distributed VPN clustering mode supports S2S IKEv2 only.

Distributed VPN clustering mode is supported on the Firepower 9300 only.

Remote access VPN is not supported in centralized or distributed VPN clustering mode.

VPN Load Balancing

VPN load balancing is a mechanism for equitably distributing remote-access VPN traffic among the devices in a VPN load-balancing group. It is based on simple distribution of traffic without taking into account throughput or other factors. A VPN load-balancing group consists of two or more devices. One device is the director, and the other devices are member devices. Group devices do not need to be of the exact same type, or have identical software versions or configurations.

All active devices in a VPN load-balancing group carry session loads. VPN load balancing directs traffic to the least-loaded device in the group, distributing the load among all devices. It makes efficient use of system resources and provides increased performance and high availability.

Failover

A failover configuration requires two identical ASAs connected to each other through a dedicated failover link and, optionally, a stateful failover link. The health of the active interfaces and units is monitored to determine when specific failover conditions are met. If those conditions occur, failover occurs. Failover supports both VPN and firewall configurations.

The ASA supports two failover configurations: Active/Active failover and Active/Standby failover.

With Active/Active failover, both units can pass network traffic. This is not true load balancing, although it might appear to have the same effect. When failover occurs, the remaining active unit takes over passing the combined traffic, based on the configured parameters. Therefore, when configuring Active/Active failover, you must make sure that the combined traffic for both units is within the capacity of each unit.

With Active/Standby failover, only one unit passes traffic, while the other unit waits in a standby state and does not pass traffic. Active/Standby failover lets you use a second ASA to take over the functions of a failed unit. When the active unit fails, it changes to the standby state, while the standby unit changes to the active state. The unit that becomes active assumes the IP addresses (or, for transparent firewall, the management IP address) and MAC addresses of the failed unit and begins passing traffic. The unit that is now in standby state takes over the standby IP addresses of the active unit. If an active unit fails, the standby takes over without any interruption to the client VPN tunnel.

VPN Load Balancing

About VPN Load Balancing

If you have a remote-client configuration in which you are using two or more ASAs connected to the same network to handle remote sessions, you can configure these devices to share their session load by creating a

VPN load-balancing group. VPN Load balancing directs session traffic to the least loaded device, thus distributing the load among all devices. It makes efficient use of system resources and provides increased performance an availability.

All devices in the VPN load-balancing group carry session loads. One device in the group, the *director*, directs incoming connection requests to the other devices, called *member devices*. The director monitors all devices in the group, keeps track of how busy each is, and distributes the session load accordingly. The role of director is not tied to a physical device; it can shift among devices. For example, if the current director fails, one of the member devices in the group takes over that role and immediately becomes the new director.

The VPN load-balancing group appears to outside clients as a single, virtual IP address. This IP address is not tied to a specific physical device. It belongs to the current director. A VPN client attempting to establish a connection connects first to the virtual IP address. The director then sends back to the client the public IP address of the least-loaded available host in the group. In a second transaction (transparent to the user), the client connects directly to that host. In this way, the VPN load-balancing group director directs traffic evenly and efficiently across resources.

If an ASA in the group fails, the terminated sessions can immediately reconnect to the virtual IP address. The director then directs these connections to another active device in the group. If the director fails, a member device in the group immediately and automatically takes over as the new director. Even if several devices in the group fail, users can continue to connect to the group as long as any one device in the group is up and available.

For each VPN load balancing cluster device, you must configure the public/outside (lbpublic) and private/inside (lbprivate) interfaces.

- Public interface: The device's outside interface used for initial communication to the cluster IP address. This interface is used for the Hello handshake.
- Private interface: The device's inside interface used for messaging between the load balancing cluster members. These messages include keepalives, topology messages, and out-of-service messages related to load balancing.

VPN Load-Balancing Algorithm

The VPN load-balancing group director maintains a sorted list of group members in ascending IP address order. The load of each member is computed as an integer percentage (the number of active sessions). AnyConnect Client inactive sessions do not count towards the SSL VPN load for VPN load balancing. The director redirects the IPsec and SSL VPN tunnel to the device with the lowest load until it is 1 percent higher than the rest. When all members are 1% higher than the director, the director redirects traffic to itself.

For example, if you have one director and two members, the following cycle applies:



Note All nodes start with 0%, and all percentages are rounded half-up.

1. The director takes the connection if all members have a load at 1% higher than the director.
2. If the director does not take the connection, the session is taken by whichever member device has the lowest load percentage.
3. If all members have the same percentage load, the member with the least number of sessions gets the session.

4. If all members have the same percentage load and the same number of sessions, the member with the lowest IP address gets the session.

VPN Load-Balancing Group Configurations

A VPN load-balancing group can consist of ASAs of the same release or of mixed releases subject to the following restrictions:

- VPN load-balancing groups that consist of both same release ASAs can run VPN load balancing for a mixture of IPsec and AnyConnect Client sessions.
- VPN load-balancing groups that include mixed release ASAs or same release ASAs can support IPsec and clientless SSL sessions. In such a configuration, however, the ASAs might not reach their full IPsec capacity.

The director of the group assigns session requests to the members of the group. The ASA regards all sessions, SSL VPN or IPsec, as equal, and assigns them accordingly. You can configure the number of IPsec and SSL VPN sessions to allow, up to the maximum allowed by your configuration and license.

We have tested up to 10 nodes in a VPN load-balancing group. Larger groups might work, but we do not officially support such topologies.

VPN Load Balancing Director Election

Director Election Process

Each non-master in the virtual cluster maintains a local topology database. This database is updated by the master whenever the topology of the cluster is changed. Each non-master goes into master election state when either no Hello response is received from the master or no Keepalive response is received from the master after maximum retries.

The member performs the following functions during director election:

- Compares the priority of each load balancing unit found in the local topology database.
- If two units with the same priority are found, one with the lower IP address is elected.
- If the member itself is elected, it claims the virtual IP address.
- If one of the other members is elected, the member sends a Hello request to the elected master.
- When two member units try to claim the virtual IP address, the ARP subsystem detects the duplicate IP address condition and sends a notification to ask the member with higher MAC address to give up the director role.

Hello Handshake

Each member sends a Hello request to the virtual cluster IP address on the outside interface upon startup. If a Hello request is received, the master sends its own Hello request to the member. The non-director member returns a Hello response upon receiving of a Hello request from the director. This concludes the Hello handshake.

Once Hello handshake is completed, the connection is initiated on the inside interface if encryption is configured. If no Hello response is received by the member after maximum retries, the member goes into master election state.

Keepalive Messages

After a Hello handshake is completed between a member and the director, each member unit sends periodic Keepalive requests to the master with its load information. Keepalive requests are sent by a member unit at one second intervals during normal processing if there is no outstanding keepalive responses from the director. This means that the next keepalive request is sent the next second as long as keepalive responses from the previous request was received. If the member did not receive a keepalive response from the director for the previous keepalive request, no keepalive request are sent the next second. Instead, the member's keepalive timeout logic starts.

The keepalive timeout works as follows:

1. If a member is waiting for an outstanding keepalive response from the director, the member does not send the regular one second interval keepalive request.
2. The member waits for 3 seconds and sends a keepalive request at the 4th second.
3. The member repeats step #2 above five(5) times as long as there is no keepalive response from the director.
4. Then, the member declares the director as gone and starts a new director election cycle.

Frequently Asked Questions About VPN Load Balancing

- Multiple Context Mode
- IP Address Pool Exhaustion
- Unique IP Address Pools
- Using VPN Load Balancing and Failover on the Same Device
- VPN Load Balancing on Multiple Interfaces
- Maximum Simultaneous Sessions for VPN Load-Balancing Groups

Multiple Context Mode

Q. Is VPN load balancing supported in multiple context mode?

A. Neither VPN load balancing nor stateful failover is supported in multiple context mode.

IP Address Pool Exhaustion

Q. Does the ASA consider IP address pool exhaustion as part of its VPN load-balancing method?

A. No. If the remote access VPN session is directed to a device that has exhausted its IP address pools, the session does not establish. The load-balancing algorithm is based on load, and is computed as an integer percentage (number of active and maximum sessions) that each member supplies.

Unique IP Address Pools

Q. To implement VPN load balancing, must the IP address pools for AnyConnect Clients or IPsec clients on different ASAs be unique?

A. Yes. IP address pools must be unique for each device.

Using VPN Load Balancing and Failover on the Same Device

Q. Can a single device use both VPN load balancing and failover?

A. Yes. In this configuration, the client connects to the IP address of the group and is redirected to the least-loaded ASA in the group. If that device fails, the standby unit takes over immediately, and there is no impact to the VPN tunnel.

VPN Load Balancing on Multiple Interfaces

Q. If we enable SSL VPN on multiple interfaces, is it possible to implement VPN load balancing for both of the interfaces?

A. You can define only one interface to participate in the VPN load-balancing group as the public interface. The idea is to balance the CPU loads. Multiple interfaces converge on the same CPU, so the concept of VPN load balancing on multiple interfaces does not improve performance.

Maximum Simultaneous Sessions for VPN Load-Balancing Groups

Q. Consider a deployment of two Firepower 1150s, each with a 100-user SSL VPN license. In a VPN load-balancing group, does the maximum total number of users allow 200 simultaneous sessions, or only 100? If we add a third device later with a 100-user license, can we now support 300 simultaneous sessions?

A. With VPN load balancing, all devices are active, so the maximum number of sessions that your group can support is the total of the number of sessions for each of the devices in the group, in this case 300.

Licensing for VPN Load Balancing

VPN load balancing has the following licensing requirements:

- An active 3DES/AES license.

  ASA checks for the existence of this crypto license before enabling VPN load balancing. If it does not detect an active 3DES or AES license, the ASA prevents the enabling of VPN load balancing and also

prevents internal configuration of 3DES by the VPN load-balancing system unless the license permits this usage.

- A valid Security Plus license for this feature activated on your firewall.
- You must have enough of these Security Plus licenses in your Smart account to be compliant.

Prerequisites for VPN Load Balancing

Also refer to the Guidelines and Limitations for VPN Load Balancing, on page 7.

- VPN load balancing is disabled by default. You must explicitly enable VPN load balancing.
- You must have first configured the public (outside) and private (inside) interfaces. Subsequent references in this section use the names outside and inside.

  To do so, go to **Configuration > Device Setup > Interface Settings > Interfaces**.
- You must have previously configured the interface to which the virtual IP address refers. Establish a common virtual IP address, UDP port (if necessary), and IPsec shared secret for the group.
- All devices that participate in a group must share the same cluster-specific values: IP address, encryption settings, encryption key, and port.
- To use VPN load-balancing group encryption, first enable IKEv1 on the inside interface using the **crypto ikev1 enable** command, with the inside interface specified; otherwise you will get an error message when you try to configure VPN load-balancing group encryption.
- The Local CA feature is not supported if you use Active/Active stateful failover or VPN load-balancing. The Local CA cannot be subordinate to another CA; it can act only as the Root CA.

Guidelines and Limitations for VPN Load Balancing

Eligible Clients

VPN Load balancing is effective only on remote sessions initiated with the following clients:

- AnyConnect Secure Mobility Client (Release 3.0 and later)
- ASA 5505 (when acting as an Easy VPN client)
- Firepower 1010 (when acting as an Easy VPN client)
- IOS EZVPN Client devices supporting IKE-redirect (IOS 831/871)
- Clientless SSL VPN

Client Considerations

VPN load balancing works with IPsec clients and SSL VPN client and clientless sessions. All other VPN connection types (L2TP, PPTP, L2TP/IPsec), including LAN-to-LAN, can connect to an ASA on which VPN load balancing is enabled, but they cannot participate in VPN load balancing.

When multiple ASA nodes are grouped for load balancing, and using Group URLs is desired for AnyConnect Client connections, the individual ASA nodes must:

- Configure each remote access connection profile with a Group URL for each VPN load-balancing virtual address (IPv4 and IPv6).
- Configure a Group URL for this node's VPN load-balancing public address.

Load Balancing Group

- ASA supports 10 devices per VPN load balancing group.
- UCAPL mode does not support VPN load balancing, even when encryption is disabled. In UCAPL mode, use IKEv2 for secure tunnel establishment.

Context Mode

VPN load balancing is not supported in multiple context mode.

FIPS

Cluster encryption not supported with FIPS.

Certificate Verification

When performing certificate verification for VPN load balancing with AnyConnect Client, and the connection is redirected by an IP address, the client does all of its name checking through this IP address. Make sure the redirection IP address is listed in the certificates common name or the subject alt name. If the IP address is not present in these fields, then the certificate will be deemed untrusted.

Following the guidelines defined in RFC 2818, if a **subject alt name** is included in the certificate, we only use the **subject alt name** for name checks, and we ignore the common name. Make sure that the IP address of the server presenting the certificate is defined in the **subject alt name** of the certificate.

For a standalone ASA, the IP address is the IP of that ASA. In a VPN load-balancing group situation, it depends on the certificate configuration. If the group uses one certificate, then the certificate should have SAN extensions for the virtual IP address and group FQDN and should contain Subject Alternative Name extensions that have each ASA's IP and FQDN. If the group uses multiple certificates, then the certificate for each ASA should have SAN extensions for the virtual IP, group FQDN, and the individual ASA's IP address and FQDN.

Geographical VPN Load Balancing

In a VPN load balancing environment where the DNS resolutions are being changed at regular intervals, you must carefully consider how to set the time to live (TTL) value. For the DNS load balance configuration to work successfully with AnyConnect Client, the ASA name-to-address mapping must remain the same from the time the ASA is selected until the tunnel is fully established. If too much time passes before the credentials are entered, the lookup restarts and a different IP address may become the resolved address. If the DNS mapping changes to a different ASA before the credentials are entered, the VPN tunnel fails.

Geographical load balancing for VPN often uses a Cisco Global Site Selector (GSS). The GSS uses DNS for the load balancing, and the time to live (TTL) value for DNS resolution is defaulted to 20 seconds. You can significantly decrease the likelihood of connection failures if you increase the TTL value on the GSS. Increasing to a much higher value allows ample time for the authentication phase when the user is entering credentials and establishing the tunnel.

To increase the time for entering credentials, you may also consider disabling Connect on Start Up.

IKE/IPSec Security Associations

Cluster encryption sessions do not sync to standby in a VPN load balancer topology.

Configuring VPN Load Balancing

If you have a remote-client configuration in which you are using two or more ASAs connected to the same network to handle remote sessions, you can configure these devices to share their session load. This feature is called VPN load balancing, which directs session traffic to the least loaded device, thereby distributing the load among all devices. VPN load balancing makes efficient use of system resources and provides increased performance and system availability.

To use VPN load balancing, do the following on each device in the group:

- Configure the VPN load-balancing group by establishing common VPN load-balancing group attributes. This includes a virtual IP address, UDP port (if necessary), and IPsec shared secret for the group. All participants in the group must have an identical group configuration, except for the device priority within the group.
- Configure the participating device by enabling VPN load balancing on the device and defining device-specific properties, such as its public and private addresses. These values vary from device to device.

Configure VPN Load Balancing with the High Availability and Scalability Wizard

Procedure

Step 1 Choose **Wizards** > **High Availability and Scalability**.

Step 2 In the Configuration Type screen, click **Configure VPN Cluster Load Balancing**, and click **Next**.

Step 3 Choose the single IP address that represents the entire VPN load-balancing group. Specify an IP address that is within the public subnet address range shared by all the ASAs in the group.

Step 4 Specify the UDP port for the VPN load-balancing group in which this device is participating. The default value is 9023. If another application is using this port, enter the UDP destination port number that you want to use for VPN load balancing.

Step 5 To enable IPsec encryption and ensure that all VPN load-balancing information communicated between the devices is encrypted, check the **Enable IPsec Encryption** check box.

Step 6 Specify and verify the **IPsec shared secret**. The value that you enter appears as consecutive asterisk characters.

Step 7 Specify the priority assigned to this device within the group. The range is from 1 to 10. The priority indicates the likelihood of this device becoming the group director, either at startup or when an existing director fails. The higher the priority set (for example, 10), the more likely that this device will become the director.

Note
If the devices in the VPN load-balancing group are powered up at different times, the first device to be powered up assumes the role of director. Each device in the group checks when it is powered up to ensure that the group has a director. If none exists, that device assumes the role. Devices powered up and added to the group later become group members. If all the devices in the group are powered up simultaneously, the device with the highest priority setting becomes the director. If two or more devices in the group are powered up simultaneously, and both have the highest priority setting, the one with the lowest IP address becomes the director.

Step 8 Choose **Public Interface of This Device**.

Step 9 Choose the **Private Interface of This Device**.

Step 10 Check the **Send FQDN to client instead of an IP address when redirecting** check box to have the director send a fully qualified domain name using the host and domain name of the device instead of the outside IP address when redirecting VPN client connections to that device.

Step 11 Click **Next**. Review your configuration in the Summary screen.

Step 12 Click **Finish**.

The VPN load-balancing group configuration is sent to the ASA.

What to do next

When multiple ASA nodes are grouped for load balancing, and using Group URLs is desired for AnyConnect Client connections, the individual ASA nodes must:

- Configure each remote access connection profile with a Group URL for each VPN load-balancing virtual address (IPv4 and IPv6).
- Configure a Group URL for this node's VPN load-balancing public address.

Group URLs are configured in the **Configuration > Remote Access VPN > Network (Client) Access >** AnyConnect Client **Connection Profiles >** *connection profile name* > **Add or Edit > Advanced > Group Alias / Group URL** pane.

Configure VPN Load Balancing (Without the Wizard)

Procedure

Step 1 Select **Configuration** > **Remote Access VPN** > **Load Balancing.**

Step 2 Check **Participate in Load Balancing** to indicate that this ASA is a participant in the load-balancing cluster

You must enable load balancing in this way on every ASA participating in load balancing.

Step 3 Configure the following fields in the **VPN Cluster Configuration** area. These values must be the same for the entire virtual cluster. All servers in the cluster must have an identical cluster configuration.

- **Cluster IPv4 Address**—Specifies the single IPv4 address that represents the entire IPv4 virtual cluster. Choose an IP address that is within the public subnet address range shared by all the ASAs in the virtual cluster.
  - **UDP Port**—Specifies the UDP port for the virtual cluster in which this device is participating. The default value is 9023. If another application is using this port, enter the UDP destination port number you want to use for load balancing.
- **Cluster IPv6 Address**—Specifies the single IPv6 address that represents the entire IPv6 virtual cluster. Choose an IP address that is within the public subnet address range shared by all the ASAs in the virtual cluster. Clients with IPv6 addresses can make AnyConnect Client connections through the ASA cluster's public-facing IPv6 address or through a GSS server. Likewise, clients with IPv6 addresses can make

AnyConnect Client VPN connections through the ASA cluster's public-facing IPv4 address or through a GSS server. Either type of connection can be load-balanced within the ASA cluster.

Note
In the Cluster IPv4 Address and Cluster IPv6 Address fields, you can also specify the fully qualified domain name of the virtual cluster, provided that you have a DNS server group configured with at least one DNS server, and DNS lookup is enabled on one of the ASA's interfaces.

- **Enable IPsec Encryption**—Enables or disables IPsec encryption. If you check this box, you must also specify and verify a shared secret. The ASAs in the virtual cluster communicate via LAN-to-LAN tunnels using IPsec. To ensure that all load-balancing information communicated between the devices is encrypted, check this box.
- **IPsec Shared Secret**—Specifies the shared secret between IPsec peers when you have enabled IPsec encryption. The value you enter in the box appears as consecutive asterisk characters.
- **Verify Secret**—Re-enter the shared secret. Confirms the shared secret value entered in the IPsec Shared Secret box.

Step 4 Configure the fields in the **VPN Server Configuration** area for a specific ASA:

- **Public Interface**—Specifies the name or IP address of the public interface for this device.
- **Private Interface**—Specifies the name or IP address of the private interface for this device.
- **Priority**—Specifies the priority assigned to this device within the cluster. The range is from 1 to 10. The priority indicates the likelihood of this device becoming the virtual cluster master, either at start-up or when an existing master fails. The higher you set the priority (for example, 10), the more likely this device becomes the virtual cluster master.

  Note
  If the devices in the virtual cluster are powered up at different times, the first device to be powered up assumes the role of virtual cluster master. Because every virtual cluster requires a master, each device in the virtual cluster checks when it is powered-up to ensure that the cluster has a virtual master. If none exists, that device takes on the role. Devices powered up and added to the cluster later become backup devices. If all the devices in the virtual cluster are powered up simultaneously, the device with the highest priority setting becomes the virtual cluster master. If two or more devices in the virtual cluster are powered up simultaneously, and both have the highest priority setting, the one with the lowest IP address becomes the virtual cluster master.

- **NAT Assigned IPv4 Address**—Specifies the IP address that this device's IP address is translated to by NAT. If NAT is not being used (or if the device is not behind a firewall using NAT), leave the field blank.
- NAT Assigned IPv6 Address—Specifies the IP address that this device's IP address is translated to by NAT. If NAT is not being used (or if the device is not behind a firewall using NAT), leave the field blank.
- **Send FQDN to client**—Check this check box to cause the VPN cluster master to send a fully qualified domain name using the host and domain name of the cluster device instead of the outside IP address when redirecting VPN client connections to that cluster device.

  By default, the ASA sends only IP addresses in load-balancing redirection to a client. If certificates are in use that are based on DNS names, the certificates will be invalid when redirected to a backup device.

As a VPN cluster master, this ASA can send a fully qualified domain name (FQDN), using reverse DNS lookup, of a cluster device (another ASA in the cluster), instead of its outside IP address, when redirecting VPN client connections to that cluster device.

All of the outside and inside network interfaces on the load-balancing devices in a cluster must be on the same IP network.

Note
When using IPv6 and sending FQDNS down to client, those names must be resolvable by the ASA via DNS.

What to do next

When multiple ASA nodes are clustered for load balancing, and using Group URLs is desired for AnyConnect Client connections, the individual ASA nodes must:

- Configure each remote access connection profile with a Group URL for each load balancing virtual cluster address (IPv4 and IPv6).
- Configure a Group URL for this node's VPN Load Balancing public address.

Group URLs are configured in the **Configuration > Remote Access VPN > Network (Client) Access >** AnyConnect Client **Connection Profiles >** *connection profile name* > **Add or Edit > Advanced > Group Alias / Group URL** pane.

Enable Clientless SSL VPN Load Balancing Using FQDNs

Procedure

Step 1 Enable the use of FQDNs for VPN load balancing by checking the **Send FQDN to client instead of an IP address when redirecting** check box.

Step 2 Add an entry for each of your ASA outside interfaces into your DNS server, if such entries are not already present. Each ASA outside IP address should have a DNS entry associated with it for lookups. These DNS entries must also be enabled for Reverse Lookup.

Step 3 Enable DNS lookups on your ASA in the dialog box **Configuration** > **Device Management** > **DNS** > **DNS Client** for whichever interface has a route to your DNS server.

Step 4 Define your DNS server IP address on the ASA. To do this, click **Add** on this dialog box. This opens the Add DNS Server Group dialog box. Enter the IPv4 or IPv6 address of the DNS server you want to add; for example, 192.168.1.1 or 2001:DB8:2000::1.

Step 5 Click **OK** and **Apply**.

Feature History for VPN Load Balancing

Feature Name	Releases	Feature Information
VPN Load balancing	7.2(1)	This feature was introduced.