ASA Cluster for the Firepower 9300 Chassis

Clustering lets you group multiple Firepower 9300 chassis ASAs together as a single logical device. The Firepower 9300 chassis series includes the Firepower 9300. A cluster provides all the convenience of a single device (management, integration into a network) while achieving the increased throughput and redundancy of multiple devices.

Note The Firepower 9300 does not support a cluster across multiple chassis (inter-chassis); only intra-chassis clustering is supported.

Note Some features are not supported when using clustering. See Unsupported Features with Clustering, on page 22.

About Clustering on the Firepower 9300 Chassis

The cluster consists of multiple devices acting as a single logical unit. When you deploy a cluster on the Firepower 9300 chassis, it does the following:

- Creates a *cluster-control link* (by default, port-channel 48) for unit-to-unit communication.

  For intra-chassis clustering, this link utilizes the Firepower 9300 backplane for cluster communications.

- Creates the cluster bootstrap configuration within the application.

When you deploy the cluster, the chassis supervisor pushes a minimal bootstrap configuration to each unit that includes the cluster name, cluster control link interface, and other cluster settings. Some parts of the bootstrap configuration may be user-configurable within the application if you want to customize your clustering environment.

- Assigns data interfaces to the cluster as *Spanned* interfaces.

  For intra-chassis clustering, spanned interfaces are not limited to EtherChannels, like it is for inter-chassis clustering.The Firepower 9300 supervisor uses EtherChannel technology internally to load-balance traffic to multiple modules on a shared interface, so any data interface type works for Spanned mode.

  Note Individual interfaces are not supported, with the exception of a management interface.

- Assigns a management interface to all units in the cluster.

The following sections provide more detail about clustering concepts and implementation. See also Reference for Clustering, on page 22.

Bootstrap Configuration

When you deploy the cluster, the Firepower 9300 chassis supervisor pushes a minimal bootstrap configuration to each unit that includes the cluster name, cluster control link interface, and other cluster settings. Some parts of the bootstrap configuration are user-configurable if you want to customize your clustering environment.

Cluster Members

Cluster members work together to accomplish the sharing of the security policy and traffic flows.

One member of the cluster is the **control** unit. The control unit is determined automatically. All other members are **data** units.

You must perform all configuration on the control unit only; the configuration is then replicated to the data units.

Some features do not scale in a cluster, and the control unit handles all traffic for those features. See Centralized Features for Clustering, on page 23.

Master and Slave Unit Roles

One member of the cluster is the master unit. The master unit is determined automatically. All other members are slave units.

You must perform all configuration on the master unit only; the configuration is then replicated to the slave units.

Some features do not scale in a cluster, and the master unit handles all traffic for those features. See Centralized Features for Clustering, on page 23.

Cluster Control Link

The cluster-control link is an EtherChannel (port-channel 48) for unit-to-unit communication. For intra-chassis clustering, this link utilizes the Firepower 9300 backplane for cluster communications.

Cluster control link traffic includes both control and data traffic.

Control traffic includes:

- Control unit election.
- Configuration replication.
- Health monitoring.

Data traffic includes:

- State replication.
- Connection ownership queries and data packet forwarding.

Cluster Control Link Network

The Firepower 9300 chassis auto-generates the cluster control link interface IP address for each unit based on the chassis ID and slot ID: 127.2.*chassis_id*.*slot_id*. You cannot set this IP address manually, either in FXOS or within the application. The cluster control link network cannot include any routers between units; only Layer 2 switching is allowed.

Cluster Interfaces

For intra-chassis clustering, you can assign both physical interfaces or EtherChannels (also known as port channels) to the cluster. Interfaces assigned to the cluster are Spanned interfaces that load-balance traffic across all members of the cluster.

Individual interfaces are not supported, with the exception of a management interface.

Connecting to a VSS or vPC

We recommend connecting EtherChannels to a VSS or vPC to provide redundancy for your interfaces.

Configuration Replication

All units in the cluster share a single configuration. You can only make configuration changes on the control unit, and changes are automatically synced to all other units in the cluster.

ASA Cluster Management

One of the benefits of using ASA clustering is the ease of management. This section describes how to manage the cluster.

Management Interface

You must assign a Management type interface to the cluster. This interface is a special individual interface as opposed to a Spanned interface. The management interface lets you connect directly to each unit.

The Main cluster IP address is a fixed address for the cluster that always belongs to the current control unit. You also configure a range of addresses so that each unit, including the current control unit, can use a Local address from the range. The Main cluster IP address provides consistent management access to an address; when a control unit changes, the Main cluster IP address moves to the new control unit, so management of the cluster continues seamlessly.

For example, you can manage the cluster by connecting to the Main cluster IP address, which is always attached to the current control unit. To manage an individual member, you can connect to the Local IP address.

For outbound management traffic such as TFTP or syslog, each unit, including the control unit, uses the Local IP address to connect to the server.

Control Unit Management Vs. Data Unit Management

All management and monitoring can take place on the control unit. From the control unit, you can check runtime statistics, resource usage, or other monitoring information of all units. You can also issue a command to all units in the cluster, and replicate the console messages from data units to the control unit.

You can monitor data units directly if desired. Although also available from the control unit, you can perform file management on data units (including backing up the configuration and updating images). The following functions are not available from the control unit:

- Monitoring per-unit cluster-specific statistics.
- Syslog monitoring per unit (except for syslogs sent to the console when console replication is enabled).
- SNMP
- NetFlow

RSA Key Replication

When you create an RSA key on the control unit, the key is replicated to all data units. If you have an SSH session to the Main cluster IP address, you will be disconnected if the control unit fails. The new control unit uses the same key for SSH connections, so that you do not need to update the cached SSH host key when you reconnect to the new control unit.

ASDM Connection Certificate IP Address Mismatch

By default, a self-signed certificate is used for the ASDM connection based on the Local IP address. If you connect to the Main cluster IP address using ASDM, then a warning message about a mismatched IP address might appear because the certificate uses the Local IP address, and not the Main cluster IP address. You can ignore the message and establish the ASDM connection. However, to avoid this type of warning, you can enroll a certificate that contains the Main cluster IP address and all the Local IP addresses from the IP address pool. You can then use this certificate for each cluster member. See https://www.cisco.com/c/en/us/td/docs/security/asdm/identity-cert/cert-install.html for more information.

Requirements and Prerequisites for Clustering on the Firepower 9300 Chassis

Maximum Clustering Units Per Model

- Firepower 9300—16 modules. For example, you can use 1 module in 16 chassis, or 2 modules in 8 chassis, or any combination that provides a maximum of 16 modules.

Switch Requirements

- Be sure to complete the switch configuration and successfully connect all the EtherChannels from the chassis to the switch(es) before you configure clustering on the Firepower 9300 chassis.
- For supported switch characteristics, see Cisco FXOS Compatibility.

Licenses for Clustering on the Firepower 9300 Chassis

The clustering feature itself does not require any licenses. To use Strong Encryption and other optional licenses, you can only request licenses on the control unit; the licenses are aggregated with the data units. If you have licenses on multiple units, they combine into a single running ASA cluster license. License configuration completed on the control unit is not replicated to the data units. You can only configure separate license entitlements on data units if you disable clustering, configure the licensing, and then re-enable clustering.

Note To use ASDM and other strong encryption features, after you deploy the cluster you must enable the Strong Encryption (3DES) license on the control unit using the ASA CLI. This license is inherited by the data units; you do not need to configure this license separately on each unit. The Strong Encryption (3DES) license is not available with any type of evaluation license.

Note If the control unit fails, and does not rejoin within 30 days (the licensing grace period), then the inherited licenses disappear. You must then manually configure the missing licenses on the new control unit.

Clustering Guidelines and Limitations

- When significant topology changes occur (such as adding or removing an EtherChannel interface, enabling or disabling an interface on the Firepower 9300 chassis or the switch, adding an additional switch to form a VSS or vPC) you should disable the health check feature, and also disable interface monitoring for the disabled interfaces . When the topology change is complete, and the configuration change is synced to all units, you can re-enable the health check feature.

- When adding a unit to an existing cluster, or when reloading a unit, there will be a temporary, limited packet/connection drop; this is expected behavior. In some cases, the dropped packets can hang connections; for example, dropping a FIN/ACK packet for an FTP connection will make the FTP client hang. In this case, you need to reestablish the FTP connection.
- If you use a Windows 2003 server connected to a Spanned EtherChannel interface, when the syslog server port is down, and the server does not throttle ICMP error messages, then large numbers of ICMP messages are sent back to the cluster. These messages can result in some units of the cluster experiencing high CPU, which can affect performance. We recommend that you throttle ICMP error messages.
- We recommend connecting EtherChannels to a VSS or vPC for redundancy.
- Within a chassis, you cannot cluster some security modules and run other security modules in standalone mode; you must include all security modules in the cluster.

Defaults

- The cluster health check feature is enabled by default with the holdtime of 3 seconds. Interface health monitoring is enabled on all interfaces by default.
- Connection rebalancing is disabled by default. If you enable connection rebalancing, the default time between load information exchanges is 5 seconds.

Configure Clustering on the Firepower 9300 Chassis

You can easily deploy the cluster from the Firepower 9300 chassis supervisor. All initial configuration is automatically generated for each unit. This section describes the default bootstrap configuration and optional customization you can perform on the ASA. This section also describes how to manage cluster members from within the ASA. You can also manage cluster membership from the Firepower 9300 chassis. See the Firepower 9300 chassis documentation for more information.

Procedure

Step 1 FXOS: Add an ASA Cluster, on page 6
Step 2 ASA: Change the Firewall Mode and Context Mode, on page 11
Step 3 ASA: Configure Data Interfaces, on page 11
Step 4 ASA: Customize the Cluster Configuration, on page 13
Step 5 ASA: Manage Cluster Members, on page 17

FXOS: Add an ASA Cluster

You can add a single Firepower 9300 chassis as an intra-chassis cluster.

Create an ASA Cluster

Set the scope to the image version.

You can easily deploy the cluster from the Firepower 9300 chassis supervisor. All initial configuration is automatically generated for each unit.

In a Firepower 9300 chassis, you must enable clustering for all 3 module slots, even if you do not have a module installed. If you do not configure all 3 modules, the cluster will not come up.

For multiple context mode, you must first deploy the logical device, and then enable multiple context mode in the ASA application.

To change the ASA to transparent firewall mode, complete the initial deployment, and then change the firewall mode within the ASA CLI.

When you deploy a cluster, the Firepower 9300 chassis supervisor configures each ASA application with the following bootstrap configuration. You can later modify parts of the bootstrap configuration from the ASA, if desired (shown in **Bold** text).

```
interface Port-channel48
   description Clustering Interface
cluster group <service_type_name>
   key <secret>
   local-unit unit-<chassis#-module#>

   cluster-interface port-channel48 ip 127.2.<chassis#>.<module#> 255.255.255.0
   priority <auto>
   health-check holdtime 3
   health-check data-interface auto-rejoin 3 5 2
   health-check cluster-interface auto-rejoin unlimited 5 1
   enable

ip local pool cluster_ipv4_pool <ip_address>-<ip_address> mask <mask>

interface <management_ifc>
   management-only individual
   nameif management
   security-level 0
   ip address <ip_address> <mask> cluster-pool cluster_ipv4_pool
   no shutdown

http server enable
http 0.0.0.0 0.0.0.0 management
route management <management_host_ip> <mask> <gateway_ip> 1
```

Note The **local-unit** name can only be changed if you disable clustering.

Before you begin

- Download the application image you want to use for the logical device from Cisco.com, and then upload that image to the Firepower 9300 chassis.
- Gather the following information:
  - Management interface ID, IP address, and network mask
  - Gateway IP address

Procedure

Step 1 Configure interfaces.

a) Add at least one Data type interface or EtherChannel (also known as a port-channel) before you deploy the cluster. See Add an EtherChannel (Port Channel) or Configure a Physical Interface.

b) Add a Management type interface or EtherChannel. See Add an EtherChannel (Port Channel) or Configure a Physical Interface.

The management interface is required. Note that this management interface is not the same as the chassis management interface that is used only for chassis management (in FXOS, you might see the chassis management interface displayed as MGMT, management0, or other similar names).

Step 2 Choose **Logical Devices**.

Step 3 Click **Add Device**, and set the following parameters:

a) Provide a **Device Name**.

This name is used internally by the chassis supervisor to configure management settings and to assign interfaces; it is not the device name used in the application configuration.

b) For the **Template**, choose **Cisco Adaptive Security Appliance**.

c) Choose the **Image Version**.

d) For the **Instance Type**, only the **Native** type is supported.

e) For the **Usage**, click the **Cluster** radio button.

f) Click **OK**.

You see the Provisioning - *device name* window.

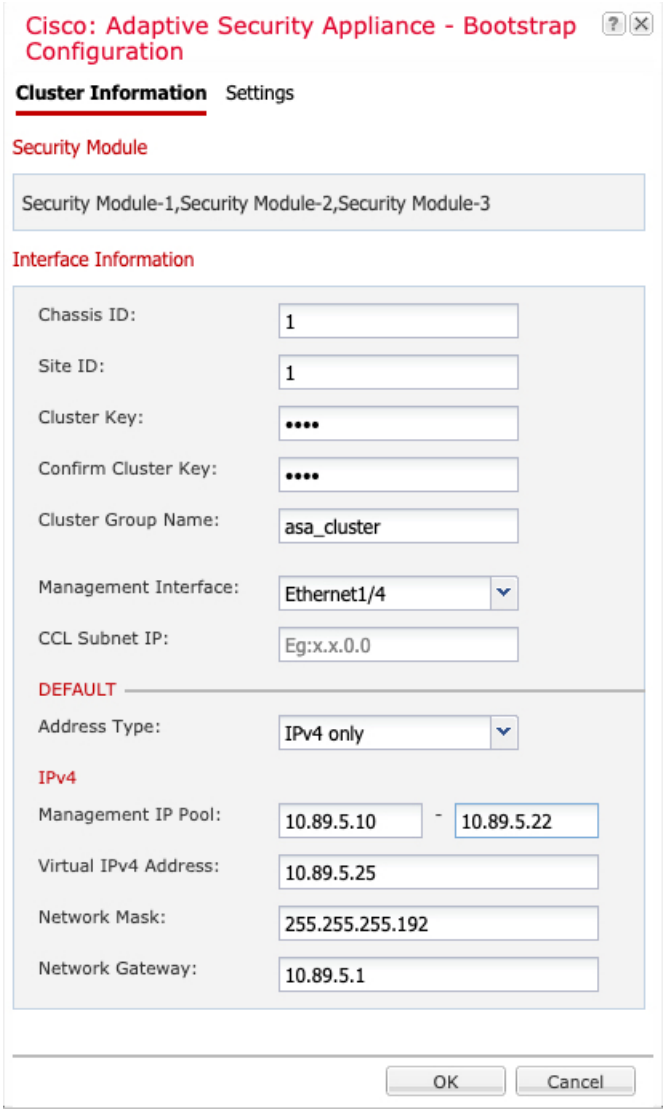
Step 4 Choose the interfaces you want to assign to this cluster.

All valid interfaces are assigned by default.

Step 5 Click the device icon in the center of the screen.

A dialog box appears where you can configure initial bootstrap settings. These settings are meant for initial deployment only, or for disaster recovery. For normal operation, you can later change most values in the application CLI configuration.

Step 6 On the **Cluster Information** page, complete the following.



a) In the **Cluster Key** field, configure an authentication key for control traffic on the cluster control link.

The shared secret is an ASCII string from 1 to 63 characters. The shared secret is used to generate the key. This option does not affect datapath traffic, including connection state update and forwarded packets, which are always sent in the clear.

b) Set the **Cluster Group Name**, which is the cluster group name in the logical device configuration.

The name must be an ASCII string from 1 to 38 characters.

c) Choose the **Management Interface**.

This interface is used to manage the logical device. This interface is separate from the chassis management port.

d) Choose the **Address Type** for the management interface.

This information is used to configure a management interface in the ASA configuration. Set the following information:

- **Management IP Pool**—Configure a pool of Local IP addresses, one of which will be assigned to each cluster unit for the interface, by entering the starting and ending addresses separated by a hyphen.

  Include at least as many addresses as there are units in the cluster. Note that for the Firepower 9300, you must include 3 addresses per chassis, even if you do not have all module slots filled. If you plan to expand the cluster, include additional addresses. The Virtual IP address (known as the Main cluster IP address) that belongs to the current control unit is *not* a part of this pool; be sure to reserve an IP address on the same network for the Main cluster IP address. You can use IPv4 and/or IPv6 addresses.

- **Network Mask** or **Prefix Length**
- **Network Gateway**
- **Virtual IP address**—Set the management IP address of the current control unit. This IP address must be on the same network as the cluster pool addresses, but not be part of the pool.

Step 7 On the **Settings** page, complte the following.

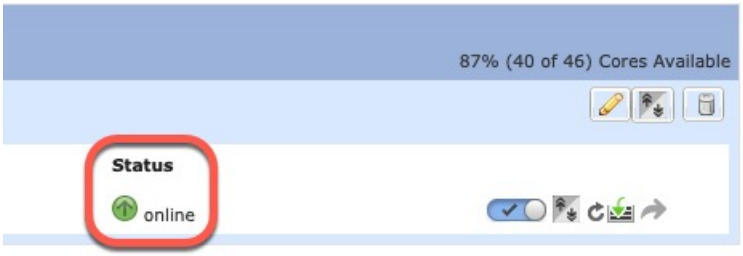
a) Enter and confirm a **Password** for the admin user.

The pre-configured ASA admin user is useful for password recovery; if you have FXOS access, you can reset the admin user password if you forget it.

Step 8 Click **OK** to close the configuration dialog box.

Step 9 Click **Save**.

The chassis deploys the logical device by downloading the specified software version and pushing the bootstrap configuration and management interface settings to the application instance. Check the **Logical Devices** page for the status of the new logical device. When the logical device shows its **Status** as **online**, you can start configuring the cluster in the application. You may see the "Security module not responding" status as part of the process; this status is normal and is temporary.



Step 10 Connect to the control unit ASA to customize your clustering configuration.

ASA: Change the Firewall Mode and Context Mode

By default, the FXOS chassis deploys a cluster in routed or transparent firewall mode, and single context mode.

- Change the firewall mode— To change the mode after you depoy, change the mode on the control unit; the mode is automatically changed on all data units to match. See Set the Firewall Mode (Single Mode). In multiple context mode, you set the firewall mode per context. See Configure a Security Context.
- Change to multiple context mode—To change to multiple context mode after you deploy, change the mode on the control unit; the mode is automatically changed on all data units to match. See Enable Multiple Context Mode.

ASA: Configure Data Interfaces

This procedure configures basic parameters for each data interface that you assigned to the cluster when you deployed it in FXOS. For inter-chassis clustering, data interfaces are always Spanned EtherChannel interfaces.

Note The management interface was pre-configured when you deployed the cluster. You can also change the management interface parameters in ASA, but this procedure focuses on data interfaces. The management interface is an individual interface, as opposed to a Spanned interface. See Management Interface, on page 4 for more information.

Before you begin

- For multiple context mode, start this procedure in the system execution space. If you are not already in the System configuration mode in the Configuration > Device List pane, double-click **System** under the active device IP address.
- For transparent mode, configure the bridge group. See Configure the Bridge Virtual Interface (BVI).
- When using Spanned EtherChannels for inter-chassis clustering, the port-channel interface will not come up until clustering is fully enabled. This requirement prevents traffic from being forwarded to a unit that is not an active unit in the cluster.

Procedure

Step 1 Depending on your context mode:

- For single mode, choose the **Configuration > Device Setup > Interface Settings > Interfaces** pane.
- For multiple mode in the System execution space, choose the **Configuration > Context Management > Interfaces** pane.

Step 2 Select the interface, and click **Edit**.

The **Edit Interface** dialog box appears.

Step 3 Set the following:

- (For EtherChannels) **MIO Port-channel ID**—Enter the same ID used in FXOS.
- **Enable Interface** (checked by default)

The rest of the fields on this screen are described later in this procedure.

Step 4 To configure the MAC address and optional parameters, click the **Advanced** tab.

- In the **MAC Address Cloning** area, set a manual global MAC address for the EtherChannel. Do not set the Standby MAC Address; it is ignored. You must configure a MAC address for a Spanned EtherChannel to avoid potential network connectivity problems. With a manually-configured MAC address, the MAC address stays with the current control unit. If you do not configure a MAC address, then if the control unit changes, the new control unit uses a new MAC address for the interface, which can cause a temporary network outage.

  In multiple context mode, if you share an interface between contexts, you should instead enable auto-generation of MAC addresses so you do not need to set the MAC address manually. Note that you must manually configure the MAC address using this command for *non-shared* interfaces.

Step 5 (Optional) Configure VLAN subinterfaces on this EtherChannel. The rest of this procedure applies to the subinterfaces.

Step 6 (Multiple context mode) Before you complete this procedure, you need to allocate interfaces to contexts.

a) Click **OK** to accept your changes.
b) Allocate interfaces.
c) Change to the context that you want to configure: in the **Device List** pane, double-click the context name under the active device IP address.
d) Choose the **Configuration > Device Setup > Interface Settings > Interfaces** pane, select the port-channel interface that you want to customize, and click **Edit**.

   The **Edit Interface** dialog box appears.

Step 7 Click the **General** tab.

Step 8 (Transparent Mode) From the **Bridge Group** drop-down list, choose the bridge group to which you want to assign this interface.

Step 9 In the **Interface Name** field, enter a name up to 48 characters in length.

Step 10 In the **Security level** field, enter a level between 0 (lowest) and 100 (highest).

Step 11 (Routed Mode) For an IPv4 address, click the **Use Static IP** radio button and enter the IP address and mask. DHCP and PPPoE are not supported. For transparent mode, you configure the IP address for the bridge group interface, not the EtherChannel interface.

Step 12 (Routed Mode) To configure an IPv6 address, click the **IPv6** tab.

For transparent mode, you configure the IP address for the bridge group interface, not the EtherChannel interface.

a) Check the **Enable IPv6** check box.
b) In the **Interface IPv6 Addresses** area, click **Add**.

   The **Add IPv6 Address for Interface** dialog box appears.

   Note The **Enable address autoconfiguration** option is not supported.

c) In the **Address/Prefix Length** field, enter the global IPv6 address and the IPv6 prefix length. For example, 2001:DB8::BA98:0:3210/64.

d) (Optional) To use the Modified EUI-64 interface ID as the host address, check the **EUI-64** check box. In this case, just enter the prefix in the **Address/Prefix Length** field.
e) Click **OK**.

Step 13 Click **OK** to return to the **Interfaces** screen.

Step 14 Click **Apply**.

ASA: Customize the Cluster Configuration

If you want to change bootstrap settings after you deploy the cluster or configure additional options, such as clustering health monitoring, TCP connection replication delay, flow mobility, and other optimizations, you can do so on the control unit.

Configure Basic ASA Cluster Parameters

You can customize cluster settings on the control unit.

Before you begin

- For multiple context mode, complete this procedure in the system execution space on the control unit. If you are not already in the System configuration mode, in the **Configuration** > **Device List** pane, double-click **System** under the active device IP address.
- The local-unit **Member Name** and several other options can only be set on the FXOS chassis, or they can only be changed on the ASA if you disable clustering, so they are not included in the following procedure.

Procedure

Step 1 Choose **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster**.

Step 2 (Optional) Configure the following optional parameters:

- **Enable connection rebalancing for TCP traffic across all the ASAs in the cluster**—Enables connection rebalancing. This parameter is disabled by default. If enabled, ASAs in a cluster exchange load information periodically, and offload new connections from more loaded devices to less loaded devices. The frequency, between 1 and 360 seconds, specifies how often the load information is exchanged. This parameter is not part of the bootstrap configuration, and is replicated from the control unit to the data units.
- **Enable health monitoring of this device within the cluster**—Enables the cluster unit health check feature, and determines the amount of time between unit keepalive status messages, between .8 and 45 seconds; The default is 3 seconds. **Note**: When you are adding new units to the cluster, and making topology changes on the ASA or the switch, you should disable this feature temporarily until the cluster is complete, and also disable interface monitoring for the disabled interfaces (**Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster** > **Cluster Interface Health Monitoring**). You can re-enable this feature after cluster and topology changes are complete. To determine unit health, the ASA cluster units send keepalive messages on the cluster control link to other units. If a unit does not receive any keepalive messages from a peer unit within the holdtime period, the peer unit is considered unresponsive or dead.

- **Replicate console output**—Enables console replication from data units to the control unit. This feature is disabled by default. The ASA may print out some messages directly to the console for certain critical events. If you enable console replication, data units send the console messages to the control unit so that you only need to monitor one console port for the cluster. This parameter is not part of the bootstrap configuration, and is replicated from the control unit to the data units.

Step 3 In the **Cluster Control Link** area, you can configue the cluster control link MTU. Other options in this area cannot be configured on the ASA.

- **MTU**—Specify the maximum transmission unit for the cluster control link interface to be at least 100 bytes higher than the highest MTU of the data interfaces. We suggest setting the MTU to the maximum of 9000; the minimum value is 1400 bytes. Because the cluster control link traffic includes data packet forwarding, the cluster control link needs to accommodate the entire size of a data packet plus cluster traffic overhead.

  For example, because the maximum MTU is 9000, then the highest data interface MTU can be 8900, while the cluster control link can be set to 9000.

Step 4 (Optional) In the **Cluster LACP** area, you can enable static port priority. The ASA uses cLACP to negotiate the EtherChannel with the neighbor switch. ASAs in a cluster collaborate in cLACP negotiation so that they appear as a single (virtual) device to the switch. Other options in this area cannot be configured on the ASA without disabling clustering.

- **Enable static port priority**—Disables dynamic port priority in LACP. Some switches do not support dynamic port priority, so this parameter improves switch compatibility. Moreover, it enables support of more than 8 active spanned EtherChannel members, up to 32 members. Without this parameter, only 8 active members and 8 standby members are supported. If you enable this parameter, then you cannot use any standby members; all members are active. This parameter is not part of the bootstrap configuration, and is replicated from the control unit to the data units.

Step 5 Click **Apply**.

Configure Interface Health Monitoring

You might want to disable health monitoring of non-essential interfaces, for example, the management interface. You can monitor any port-channel ID or single physical interface ID. Health monitoring is not performed on VLAN subinterfaces or virtual interfaces such as VNIs or BVIs. You cannot configure monitoring for the cluster control link; it is always monitored.

Procedure

Step 1 Choose **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster** > **Cluster Interface Health Monitoring**.

Step 2 Select an interface in the **Monitored Interfaces** box, and click **Add** to move it to the **Unmonitored Interfaces** box.

Interface status messages detect link failure. If all physical ports for a given logical interface fail on a particular unit, but there are active ports under the same logical interface on other units, then the unit is removed from the cluster. If a unit does not receive interface status messages within the holdtime, then the amount of time

before the ASA removes a member from the cluster depends on the type of interface and whether the unit is an established member or is joining the cluster. Health check is enabled by default for all interfaces.

You might want to disable health monitoring of non-essential interfaces, for example, the management interface. You can specify any port-channel ID or single physical interface ID. Health monitoring is not performed on VLAN subinterfaces or virtual interfaces such as VNIs or BVIs. You cannot configure monitoring for the cluster control link; it is always monitored.

When any topology changes occur (such as adding or removing a data interface, enabling or disabling an interface on the ASA, Firepower 9300 chassis, or the switch, or adding an additional switch to form a VSS or vPC) you should disable the health check feature (**Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster**) and also disable interface monitoring for the disabled interfaces. When the topology change is complete, and the configuration change is synced to all units, you can re-enable the health check feature.

Step 3 Click **Apply**.

Configure the Cluster TCP Replication Delay

Enable the cluster replication delay for TCP connections to help eliminate the "unnecessary work" related to short-lived flows by delaying the director/backup flow creation. Note that if a unit fails before the director/backup flow is created, then those flows cannot be recovered. Similarly, if traffic is rebalanced to a different unit before the flow is created, then the flow cannot be recovered. You should not enable the TCP replication delay for traffic on which you disable TCP randomization.

Procedure

Step 1 Choose **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster Replication**.

Step 2 Click **Add** and set the following values:

- **Replication delay**—Set the seconds between 1 and 15.
- **HTTP**—Set the delay for all HTTP traffic. This setting is enabled by default for 5 seconds.
- **Source Criteria**
  - **Source**—Set the source IP address.
  - Service—(Optional) Set the source port. Typically you set either the source or the destination port, but not both.
- **Destination Criteria**
  - **Source**—Set the destination IP address.
  - **Service**—(Optional) Set the destination port. Typically you set either the source or the destination port, but not both.

Step 3 Click **OK**.

Step 4 Click **Apply**.

FXOS: Remove a Cluster Unit

The following sections describe how to remove units temporarily or permanently from the cluster.

Temporary Removal

A cluster unit will be automatically removed from the cluster due to a hardware or network failure, for example. This removal is temporary until the conditions are rectified, and it can rejoin the cluster. You can also manually disable clustering.

To check whether a device is currently in the cluster, check the cluster status on the Firepower Chassis Manager **Logical Devices** page:

- Disable clustering in the application—You can disable clustering using the application CLI. Enter the **cluster remove unit** *name* command to remove any unit other than the one you are logged into. The bootstrap configuration remains intact, as well as the last configuration synced from the control unit, so you can later re-add the unit without losing your configuration. If you enter this command on a data unit to remove the control unit, a new control unit is elected.

  When a device becomes inactive, all data interfaces are shut down; only the Management interface can send and receive traffic. To resume traffic flow, re-enable clustering. The Management interface remains up using the IP address the unit received from the bootstrap configuration. However if you reload, and the unit is still inactive in the cluster (for example, you saved the configuration with clustering disabled), the Management interface is disabled.

  To reenable clustering, on the ASA enter **cluster group** *name* and then **enable**.

- Disable the application instance—In Firepower Chassis Manager on the **Logical Devices** page, click the **Slider enabled** ( ). You can later reenable it using the **Slider disabled** ( ).

- Shut down the security module/engine—In Firepower Chassis Manager on the **Security Module/Engine** page, click the **Power Off icon**.

- Shut down the chassis—In Firepower Chassis Manager on the **Overview** page, click the **Shut Down icon**.

Permanent Removal

You can permanently remove a cluster member using the following methods.

- Delete the logical device—In Firepower Chassis Manager on the **Logical Devices** page, click the **Delete** ( ). You can then deploy a standalone logical device, a new cluster, or even add a new logical device to the same cluster.
- Remove the chassis or security module from service—If you remove a device from service, you can add replacement hardware as a new member of the cluster.

ASA: Manage Cluster Members

After you deploy the cluster, you can change the configuration and manage cluster members.

Become an Inactive Member

To become an inactive member of the cluster, disable clustering on the unit while leaving the clustering configuration intact.

Note When an ASA becomes inactive (either manually or through a health check failure), all data interfaces are shut down; only the management-only interface can send and receive traffic. To resume traffic flow, re-enable clustering; or you can remove the unit altogether from the cluster. The management interface remains up using the IP address the unit received from the cluster IP pool. However if you reload, and the unit is still inactive in the cluster (for example, you saved the configuration with clustering disabled), then the management interface is disabled. You must use the console port for any further configuration.

Before you begin

- For multiple context mode, perform this procedure in the system execution space. If you are not already in the System configuration mode in the Configuration > Device List pane, double-click **System** under the active device IP address.

Procedure

Step 1 Choose **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster** > **Cluster Configuration**.

Step 2 Uncheck the **Participate in ASA cluster** check box.

Note Do not uncheck the **Configure ASA cluster settings** check box; this action clears all cluster configuration, and also shuts down all interfaces including the management interface to which ASDM is connected. To restore connectivity in this case, you need to access the CLI at the console port.

Step 3 Click **Apply**.

Deactivate a Data Unit from the Control Unit

To deactivate a data unit, perform the following steps.

Note When an ASA becomes inactive, all data interfaces are shut down; only the management-only interface can send and receive traffic. To resume traffic flow, re-enable clustering. The management interface remains up using the IP address the unit received from the cluster IP pool. However if you reload, and the unit is still inactive in the cluster (for example, if you saved the configuration with clustering disabled), the management interface is disabled. You must use the console port for any further configuration.

Before you begin

For multiple context mode, perform this procedure in the system execution space. If you are not already in the System configuration mode in the **Configuration > Device List** pane, double-click **System** under the active device IP address.

Procedure

Step 1 Choose **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster**.

Step 2 Select the data unit that you want to remove, and click **Delete**.

The data unit bootstrap configuration remains intact, so that you can later re-add the data unit without losing your configuration.

Step 3 Click **Apply**.

Rejoin the Cluster

If a unit was removed from the cluster, for example for a failed interface or if you manually deactivated a member, you must manually rejoin the cluster.

Before you begin

- You must use the console port to reenable clustering. Other interfaces are shut down. The exception is if you manually disabled clustering in ASDM, then you can reenable clustering in ASDM if you did not save the configuration and reload. After reloading, the management interface is disabled, so console access is the only method to reenable clustering.
- For multiple context mode, perform this procedure in the system execution space. If you are not already in the System configuration mode in the **Configuration > Device List** pane, double-click **System** under the active device IP address.
- Make sure the failure is resolved before you try to rejoin the cluster.

Procedure

Step 1 If you still have ASDM access, you can reenable clustering in ASDM by connecting ASDM to the unit you want to reenable.

You cannot reenable clustering for a data unit from the control unit unless you add it as a new member.

a) Choose **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster**.
b) Check the **Participate in ASA cluster** check box.
c) Click **Apply**.

Step 2 If you cannot use ASDM: At the console, enter cluster configuration mode:

cluster group *name*

Example:

```
ciscoasa(config)# cluster group pod1
```

Step 3 Enable clustering.

enable

Change the Control Unit

Caution The best method to change the control unit is to disable clustering on the control unit, wait for a new control election, and then re-enable clustering. If you must specify the exact unit you want to become the control unit, use the procedure in this section. Note, however, that for centralized features, if you force a control unit change using this procedure, then all connections are dropped, and you have to re-establish the connections on the new control unit.

To change the control unit, perform the following steps.

Before you begin

For multiple context mode, perform this procedure in the system execution space. If you are not already in the System configuration mode in the Configuration > Device List pane, double-click **System** under the active device IP address.

Procedure

Step 1 Choose **Monitoring** > **ASA Cluster** > **Cluster Summary**.

Step 2 From the drop-down list, choose a data unit to become master, and click the button to make it the control unit.

Step 3 You are prompted to confirm the control unit change. Click **Yes**.

Step 4 Quit ASDM, and reconnect using the Main cluster IP address.

Execute a Command Cluster-Wide

To send a command to all members in the cluster, or to a specific member, perform the following steps. Sending a **show** command to all members collects all output and displays it on the console of the current unit. (Note that alternatively there are show commands that you can enter on the control unit to view cluster-wide statistics.) Other commands, such as **capture** and **copy**, can also take advantage of cluster-wide execution.

Before you begin

Perform this procedure at the Command Line Interface tool: choose **Tools** > **Command Line Interface**.

Procedure

Send a command to all members, or if you specify the unit name, a specific member:

cluster exec [unit *unit_name*] *command*

Example:

```
cluster exec show xlate
```

To view member names, enter **cluster exec unit ?** (to see all names except the current unit), or enter the **show cluster info** command.

Examples

To copy the same capture file from all units in the cluster at the same time to a TFTP server, enter the following command on the control unit:

```
cluster exec copy /pcap capture: tftp://10.1.1.56/capture1.pcap
```

Multiple PCAP files, one from each unit, are copied to the TFTP server. The destination capture file name is automatically attached with the unit name, such as capture1_asa1.pcap, capture1_asa2.pcap, and so on. In this example, asa1 and asa2 are cluster unit names.

The following sample output for the **cluster exec show memory** command shows memory information for each member in the cluster:

```
cluster exec show memory
unit-1-1(LOCAL):*******************************************************
Free memory:      108724634538 bytes (92%)
Used memory:        9410087158 bytes ( 8%)
-------------     ------------------
Total memory:     118111600640 bytes (100%)


unit-1-3:*******************************************************
Free memory:      108749922170 bytes (92%)
Used memory:        9371097334 bytes ( 8%)
-------------     ------------------
Total memory:     118111600640 bytes (100%)
```

```
unit-1-2:************************************************************
Free memory:      108426753537 bytes (92%)
Used memory:        9697869087 bytes ( 8%)
-------------     ------------------
Total memory:     118111600640 bytes (100%)
```

ASA: Monitoring the ASA Cluster on the Firepower 9300 chassis

You can monitor and troubleshoot cluster status and connections.

Monitoring Cluster Status

See the following screens for monitoring cluster status:

- **Monitoring > ASA Cluster > Cluster Summary**

  This pane shows cluster information about the unit to which you are connected, as well as other units in the cluster. You can also change the primary unit from this pane.

- **Cluster Dashboard**

  On the home page on the primary unit, you can monitor the cluster using the Cluster Dashboard and the Cluster Firewall Dashboard.

Capturing Packets Cluster-Wide

See the following screen for capturing packets in a cluster:

Wizards > Packet Capture Wizard

To support cluster-wide troubleshooting, you can enable capture of cluster-specific traffic on the control unit, which is then automatically enabled on all of the data units in the cluster.

Monitoring Cluster Resources

See the following screens for monitoring cluster resources:

- **Monitoring > ASA Cluster > System Resources Graphs > CPU**

  This pane lets you create graphs or tables showing the CPU utilization across the cluster members.

- **Monitoring > ASA Cluster > System Resources Graphs > Memory**.

  This pane lets you create graphs or tables showing the Free Memory and Used Memory across the cluster members.

Monitoring Cluster Traffic

See the following screens for monitoring cluster traffic:

- **Monitoring > ASA Cluster > Traffic Graphs > Connections**.

  This pane lets you create graphs or tables showing the Connections across the cluster members.

- **Monitoring > ASA Cluster > Traffic Graphs > Throughput**.

  This pane lets you create graphs or tables showing the traffic throughput across the cluster members.

Monitoring the Cluster Control Link

See the following screen for monitoring cluster status:

Monitoring > Properties > System Resources Graphs > Cluster Control Link.

This pane lets you create graphs or tables showing the cluster control link Receival and Transmittal capacity utilization.

Configuring Logging for Clustering

See the followingscreen for configuring logging for clustering:

Configuration > Device Management > Logging > Syslog Setup

Each unit in the cluster generates syslog messages independently. You can generate syslog messages with identical or different device IDs to make messages appear to come from the same or different units in the cluster.

Reference for Clustering

This section includes more information about how clustering operates.

ASA Features and Clustering

Some ASA features are not supported with ASA clustering, and some are only supported on the control unit. Other features might have caveats for proper usage.

Unsupported Features with Clustering

These features cannot be configured with clustering enabled, and the commands will be rejected.

- Unified Communication features that rely on TLS Proxy
- The following application inspections:
  - CTIQBE
  - GTP
  - H323, H225, and RAS
  - IPsec passthrough
  - MGCP

    - MMP
    - RTSP
    - SCCP (Skinny)
    - WAAS
    - WCCP
- Botnet Traffic Filter
- Auto Update Server
- DHCP client, server, and proxy. DHCP relay is supported.
- Failover
- Dead Connection Detection (DCD)
- FIPS mode

Centralized Features for Clustering

The following features are only supported on the control unit, and are not scaled for the cluster.

Note Traffic for centralized features is forwarded from member units to the control unit over the cluster control link.

If you use the rebalancing feature, traffic for centralized features may be rebalanced to non-control units before the traffic is classified as a centralized feature; if this occurs, the traffic is then sent back to the control unit.

For centralized features, if the control unit fails, all connections are dropped, and you have to re-establish the connections on the new control unit.

- The following application inspections:
    - DCERPC
    - NetBIOS
    - PPTP
    - RADIUS
    - RSH
    - SUNRPC
    - TFTP
    - XDMCP
- Dynamic routing
- Static route tracking

- IGMP multicast control plane protocol processing (data plane forwarding is distributed across the cluster)
- PIM multicast control plane protocol processing (data plane forwarding is distributed across the cluster)
- Authentication and Authorization for network access. Accounting is decentralized.
- Filtering Services

Features Applied to Individual Units

These features are applied to each ASA unit, instead of the cluster as a whole or to the control unit.

- QoS—The QoS policy is synced across the cluster as part of configuration replication. However, the policy is enforced on each unit independently. For example, if you configure policing on output, then the conform rate and conform burst values are enforced on traffic exiting a particular ASA. In a cluster with 3 units and with traffic evenly distributed, the conform rate actually becomes 3 times the rate for the cluster.
- Threat detection—Threat detection works on each unit independently; for example, the top statistics is unit-specific. Port scanning detection, for example, does not work because scanning traffic will be load-balanced between all units, and one unit will not see all traffic.
- Resource management—Resource management in multiple context mode is enforced separately on each unit based on local usage.

AAA for Network Access and Clustering

AAA for network access consists of three components: authentication, authorization, and accounting. Authentication and authorization are implemented as centralized features on the clustering control unit with replication of the data structures to the cluster data units. If a control unit is elected, the new control unit will have all the information it needs to continue uninterrupted operation of the established authenticated users and their associated authorizations. Idle and absolute timeouts for user authentications are preserved when a control unit change occurs.

Accounting is implemented as a distributed feature in a cluster. Accounting is done on a per-flow basis, so the cluster unit owning a flow will send accounting start and stop messages to the AAA server when accounting is configured for a flow.

Connection Settings

Connection limits are enforced cluster-wide (see **Configuration** > **Firewall** > **Service Policy** page). Each unit has an estimate of the cluster-wide counter values based on broadcast messages. Due to efficiency considerations, the configured connection limit across the cluster might not be enforced exactly at the limit number. Each unit may overestimate or underestimate the cluster-wide counter value at any given time. However, the information will get updated over time in a load-balanced cluster.

FTP and Clustering

- If FTP data channel and control channel flows are owned by different cluster members, then the data channel owner will periodically send idle timeout updates to the control channel owner and update the idle timeout value. However, if the control flow owner is reloaded, and the control flow is re-hosted, the parent/child flow relationship will not longer be maintained; the control flow idle timeout will not be updated.

- If you use AAA for FTP access, then the control channel flow is centralized on the control unit.

Identity Firewall and Clustering

Only the control unit retrieves the user-group from the AD and the user-ip mapping from the AD agent. The control unit then populates the user information to data units, and data units can make a match decision for user identity based on the security policy.

Multicast Routing and Clustering

The control unit handles all multicast routing packets and data packets until fast-path forwarding is established. After the connection is established, each data unit can forward multicast data packets.

NAT and Clustering

NAT can affect the overall throughput of the cluster. Inbound and outbound NAT packets can be sent to different ASAs in the cluster, because the load balancing algorithm relies on IP addresses and ports, and NAT causes inbound and outbound packets to have different IP addresses and/or ports. When a packet arrives at the ASA that is not the NAT owner, it is forwarded over the cluster control link to the owner, causing large amounts of traffic on the cluster control link. Note that the receiving unit does not create a forwarding flow to the owner, because the NAT owner may not end up creating a connection for the packet depending on the results of security and policy checks.

If you still want to use NAT in clustering, then consider the following guidelines:

- NAT pool address distribution for dynamic PAT—The control unit evenly pre-distributes addresses across the cluster. If a member receives a connection and they have no addresses assigned, then the connection is forwarded to the control unit for PAT. If a cluster member leaves the cluster (due to failure), a backup member will get the PAT IP address, and if the backup exhausts its normal PAT IP address, it can make use of the new address. Make sure to include at least as many NAT addresses as there are units in the cluster, plus at least one extra address, to ensure that each unit receives an address, and that a failed unit can get a new address if its old address is in use by the member that took over the address. Use the **show nat pool cluster** command to see the address allocations.
- Reusing a PAT pool in multiple rules—To use the same PAT pool in multiple rules, you must be careful about the interface selection in the rules. You must either use specific interfaces in all rules, or "any" in all rules. You cannot mix specific interfaces and "any" across the rules, or the system might not be able to match return traffic to the right node in the cluster. Using unique PAT pools per rule is the most reliable option.
- No round-robin—Round-robin for a PAT pool is not supported with clustering.
- Dynamic NAT xlates managed by the control unit—The control unit maintains and replicates the xlate table to data units. When a data unit receives a connection that requires dynamic NAT, and the xlate is not in the table, it requests the xlate from the control unit. The data unit owns the connection.
- Stale xlates—The xlate idle time on the connection owner does not get updated. Thus, the idle time might exceed the idle timeout. An idle timer value higher than the configured timeout with a refcnt of 0 is an indication of a stale xlate.
- Per-session PAT feature—Although not exclusive to clustering, the per-session PAT feature improves the scalability of PAT and, for clustering, allows each data unit to own PAT connections; by contrast, multi-session PAT connections have to be forwarded to and owned by the control unit. By default, all TCP traffic and UDP DNS traffic use a per-session PAT xlate, whereas ICMP and all other UDP traffic

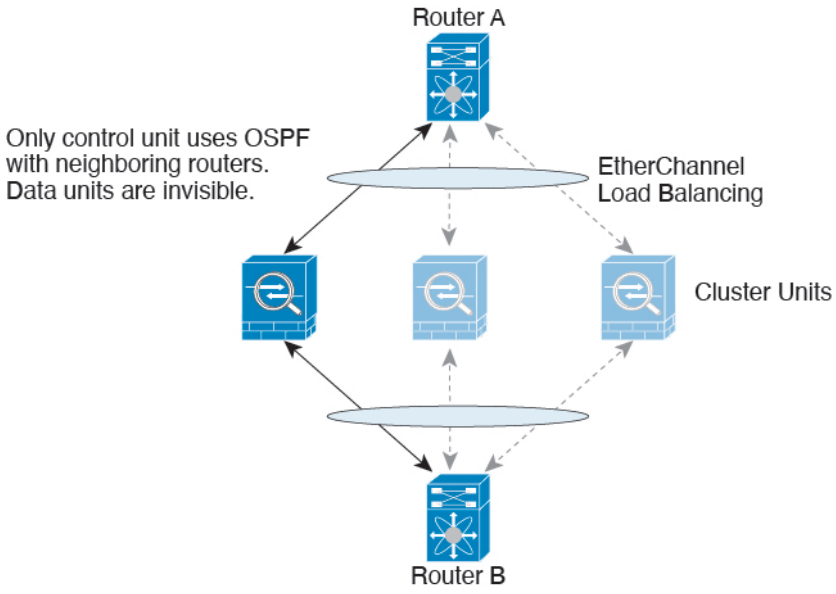
uses multi-session. You can configure per-session NAT rules to change these defaults for TCP and UDP, but you cannot configure per-session PAT for ICMP. For traffic that benefits from multi-session PAT, such as H.323, SIP, or Skinny, you can disable per-session PAT for the associated TCP ports (the UDP ports for those H.323 and SIP are already multi-session by default). For more information about per-session PAT, see the firewall configuration guide.

- No static PAT for the following inspections—
  - FTP
  - PPTP
  - RSH
  - SQLNET
  - TFTP
  - XDMCP
  - SIP
- If you have an extremely large number of NAT rules, over ten thousand, you should enable the transactional commit model using the **asp rule-engine transactional-commit nat** command in the device CLI. Otherwise, the unit might not be able to join the cluster.

Dynamic Routing and Clustering

The routing process only runs on the control unit, and routes are learned through the control unit and replicated to secondaries. If a routing packet arrives at a data unit, it is redirected to the control unit.

Figure 1: Dynamic Routing



After the data units learn the routes from the control unit, each unit makes forwarding decisions independently.

The OSPF LSA database is not synchronized from the control unit to data units. If there is a control unit switchover, the neighboring router will detect a restart; the switchover is not transparent. The OSPF process picks an IP address as its router ID. Although not required, you can assign a static router ID to ensure a consistent router ID is used across the cluster. See the OSPF Non-Stop Forwarding feature to address the interruption.

SIP Inspection and Clustering

A control flow can be created on any unit (due to load balancing); its child data flows must reside on the same unit.

TLS Proxy configuration is not supported.

SNMP and Clustering

An SNMP agent polls each individual ASA by its Local IP address. You cannot poll consolidated data for the cluster.

You should always use the Local address, and not the Main cluster IP address for SNMP polling. If the SNMP agent polls the Main cluster IP address, if a new control unit is elected, the poll to the new control unit will fail.

When using SNMPv3 with clustering, if you add a new cluster unit after the initial cluster formation, then SNMPv3 users are not replicated to the new unit.You must re-add them on the control unit to force the users to replicate to the new unit, or directly on the data unit.

Syslog and NetFlow and Clustering

- Syslog—Each unit in the cluster generates its own syslog messages. You can configure logging so that each unit uses either the same or a different device ID in the syslog message header field. For example, the hostname configuration is replicated and shared by all units in the cluster. If you configure logging to use the hostname as the device ID, syslog messages generated by all units look as if they come from a single unit. If you configure logging to use the local-unit name that is assigned in the cluster bootstrap configuration as the device ID, syslog messages look as if they come from different units.
- NetFlow—Each unit in the cluster generates its own NetFlow stream. The NetFlow collector can only treat each ASA as a separate NetFlow exporter.

Cisco TrustSec and Clustering

Only the control unit learns security group tag (SGT) information. The control unit then populates the SGT to data units, and data units can make a match decision for SGT based on the security policy.

Performance Scaling Factor

When you combine multiple units into a cluster, you can expect the total cluster performance to be approximately:

- 80% of the combined TCP or CPS throughput
- 90% of the combined UDP throughput
- 60% of the combined Ethernet MIX (EMIX) throughput, depending on the traffic mix.

For example, for TCP throughput, the Firepower 9300 with 3 SM-44 modules can handle approximately 135 Gbps of real world firewall traffic when running alone. For 2 chassis, the maximum combined throughput will be approximately 80% of 270 Gbps (2 chassis x 135 Gbps): 216 Gbps.

Control Unit Election

Members of the cluster communicate over the cluster control link to elect a control unit as follows:

1. When you deploy the cluster, each unit broadcasts an election request every 3 seconds.
2. Any other units with a higher priority respond to the election request; the priority is set when you deploy the cluster and is not configurable.
3. If after 45 seconds, a unit does not receive a response from another unit with a higher priority, then it becomes the control unit.

Note If multiple units tie for the highest priority, the cluster unit name and then the serial number is used to determine the control unit.

4. If a unit later joins the cluster with a higher priority, it does not automatically become the control unit; the existing control unit always remains as the control unit unless it stops responding, at which point a new control unit is elected.
5. In a "split brain" scenario when there are temporarily multiple control units, then the unit with highest priority retains the role while the other units return to data unit roles.

Note You can manually force a unit to become the control unit. For centralized features, if you force a control unit change, then all connections are dropped, and you have to re-establish the connections on the new control unit.

High Availability Within the Cluster

Clustering provides high availability by monitoring chassis, unit, and interface health and by replicating connection states between units.

Chassis-Application Monitoring

Chassis-application health monitoring is always enabled. The Firepower 9300 chassis supervisor checks the ASA application periodically (every second). If the ASA is up and cannot communicate with the Firepower 9300 chassis supervisor for 3 seconds, the ASA generates a syslog message and leaves the cluster.

If the Firepower 9300 chassis supervisor cannot communicate with the application after 45 seconds, it reloads the ASA. If the ASA cannot communicate with the supervisor, it removes itself from the cluster.

Unit Health Monitoring

Each unit periodically sends a broadcast keepalivekeepalive packet over the cluster control link. If the control unit does not receive any keepalivekeepalive packets or other packets from a data unit within the configurable

timeout period, then the control unit removes the data unit from the cluster. If the data units do not receive packets from the control unit, then a new control unit is elected from the remaining members.

If units cannot reach each other over the cluster control link because of a network failure and not because a unit has actually failed, then the cluster may go into a "split brain" scenario where isolated data units will elect their own control units. For example, if a router fails between two cluster locations, then the original control unit at location 1 will remove the location 2 data units from the cluster. Meanwhile, the units at location 2 will elect their own control unit and form their own cluster. Note that asymmetric traffic may fail in this scenario. After the cluster control link is restored, then the control unit that has the higher priority will keep the control unit's role. See Control Unit Election, on page 28 for more information.

Interface Monitoring

Each unit monitors the link status of all hardware interfaces in use, and reports status changes to the control unit. When you enable health monitoring, all physical interfaces are monitored by default (including the main EtherChannel for EtherChannel interfaces). Only named interfaces that are in an Up state can be monitored. For example, all member ports of an EtherChannel must fail before a *named* EtherChannel is removed from the cluster (depending on your minimum port bundling setting). You can optionally disable monitoring per interface.

If a monitored interface fails on a particular unit, but it is active on other units, then the unit is removed from the cluster. The amount of time before the ASA removes a member from the cluster depends on whether the unit is an established member or is joining the cluster. The ASA does not monitor interfaces for the first 90 seconds that a unit joins the cluster. Interface status changes during this time will not cause the ASA to be removed from the cluster. For an established member, the unit is removed after 500 ms.

Status After Failure

When a unit in the cluster fails, the connections hosted by that unit are seamlessly transferred to other units; state information for traffic flows is shared over the control unit's cluster control link.

If the control unit fails, then another member of the cluster with the highest priority (lowest number) becomes the control unit.

The ASA automatically tries to rejoin the cluster, depending on the failure event.

Note When the ASA becomes inactive and fails to automatically rejoin the cluster, all data interfaces are shut down; only the management-only interface can send and receive traffic. The management interface remains up using the IP address the unit received from the cluster IP pool. However if you reload, and the unit is still inactive in the cluster, the management interface is disabled. You must use the console port for any further configuration.

Rejoining the Cluster

After a cluster member is removed from the cluster, how it can rejoin the cluster depends on why it was removed:

- Failed cluster control link—After you resolve the problem with the cluster control link, you must manually rejoin the cluster by re-enabling clustering.
- Failed data interface—The ASA automatically tries to rejoin at 5 minutes, then at 10 minutes, and finally at 20 minutes. If the join is not successful after 20 minutes, then the ASA disables clustering. After you resolve the problem with the data interface, you have to manually enable clustering.

- Failed unit—If the unit was removed from the cluster because of a unit health check failure, then rejoining the cluster depends on the source of the failure. For example, a temporary power failure means the unit will rejoin the cluster when it starts up again as long as the cluster control link is up. The unit attempts to rejoin the cluster every 5 seconds.
- Failed Chassis-Application Communication—When the ASA detects that the chassis-application health has recovered, the ASA tries to rejoin the cluster automatically.
- Internal error—Internal failures include: application sync timeout; inconsistent application statuses; and so on. After you resolve the problem, you must manually rejoin the cluster by re-enabling clustering.

Data Path Connection State Replication

Every connection has one owner and at least one backup owner in the cluster. The backup owner does not take over the connection in the event of a failure; instead, it stores TCP/UDP state information, so that the connection can be seamlessly transferred to a new owner in case of a failure. The backup owner is usually also the director.

Some traffic requires state information above the TCP or UDP layer. See the following table for clustering support or lack of support for this kind of traffic.

Table 1: Features Replicated Across the Cluster

Traffic	State Support	Notes
Up time	Yes	Keeps track of the system up time.
ARP Table	Yes	Transparent mode only.
MAC address table	Yes	Transparent mode only.
User Identity	Yes	Includes AAA rules (uauth) and identity firewall.
IPv6 Neighbor database	Yes	—
Dynamic routing	Yes	—
SNMP Engine ID	**No**	—
Centralized VPN (Site-to-Site)	**No**	VPN sessions will be disconnected if the control unit fails.

How the Cluster Manages Connections

Connections can be load-balanced to multiple members of the cluster. Connection roles determine how connections are handled in both normal operation and in a high availability situation.

Connection Roles

See the following roles defined for each connection:

- Owner—Usually, the unit that initially receives the connection. The owner maintains the TCP state and processes packets. A connection has only one owner. If the original owner fails, then when new units receive packets from the connection, the director chooses a new owner from those units.
- Backup owner—The unit that stores TCP/UDP state information received from the owner, so that the connection can be seamlessly transferred to a new owner in case of a failure. The backup owner does not take over the connection in the event of a failure. If the owner becomes unavailable, then the first unit to receive packets from the connection (based on load balancing) contacts the backup owner for the relevant state information so it can become the new owner.

  As long as the director (see below) is not the same unit as the owner, then the director is also the backup owner. If the owner chooses itself as the director, then a separate backup owner is chosen.
- Director—The unit that handles owner lookup requests from forwarders. When the owner receives a new connection, it chooses a director based on a hash of the source/destination IP address and ports, and sends a message to the director to register the new connection. If packets arrive at any unit other than the owner, the unit queries the director about which unit is the owner so it can forward the packets. A connection has only one director. If a director fails, the owner chooses a new director.

  As long as the director is not the same unit as the owner, then the director is also the backup owner (see above). If the owner chooses itself as the director, then a separate backup owner is chosen.
- Forwarder—A unit that forwards packets to the owner. If a forwarder receives a packet for a connection it does not own, it queries the director for the owner, and then establishes a flow to the owner for any other packets it receives for this connection. The director can also be a forwarder. Note that if a forwarder receives the SYN-ACK packet, it can derive the owner directly from a SYN cookie in the packet, so it does not need to query the director. (If you disable TCP sequence randomization, the SYN cookie is not used; a query to the director is required.) For short-lived flows such as DNS and ICMP, instead of querying, the forwarder immediately sends the packet to the director, which then sends them to the owner. A connection can have multiple forwarders; the most efficient throughput is achieved by a good load-balancing method where there are no forwarders and all packets of a connection are received by the owner.

  Note We do not recommend disabling TCP sequence randomization when using clustering. There is a small chance that some TCP sessions won't be established, because the SYN/ACK packet might be dropped.
- Fragment Owner—For fragmented packets, cluster units that receive a fragment determine a fragment owner using a hash of the fragment source IP address, destination IP address, and the packet ID. All fragments are then forwarded to the fragment owner over the cluster control link. Fragments may be load-balanced to different cluster units, because only the first fragment includes the 5-tuple used in the switch load balance hash. Other fragments do not contain the source and destination ports and may be load-balanced to other cluster units. The fragment owner temporarily reassembles the packet so it can determine the director based on a hash of the source/destination IP address and ports. If it is a new connection, the fragment owner will register to be the connection owner. If it is an existing connection, the fragment owner forwards all fragments to the provided connection owner over the cluster control link. The connection owner will then reassemble all fragments.

When a connection uses Port Address Translation (PAT), then the PAT type (per-session or multi-session) influences which member of the cluster becomes the owner of a new connection:

- Per-session PAT—The owner is the unit that receives the initial packet in the connection.

By default, TCP and DNS UDP traffic use per-session PAT.

- Multi-session PAT—The owner is always the control unit. If a multi-session PAT connection is initially received by a data unit, then the data unit forwards the connection to the control unit.

  By default, UDP (except for DNS UDP) and ICMP traffic use multi-session PAT, so these connections are always owned by the control unit.

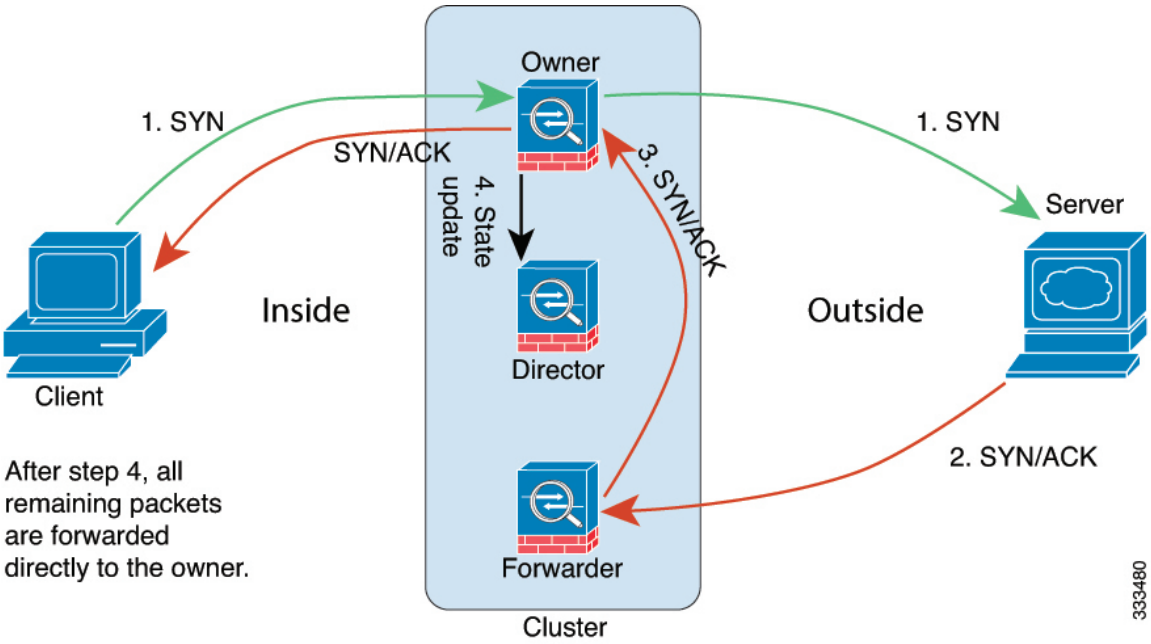
You can change the per-session PAT defaults for TCP and UDP so connections for these protocols are handled per-session or multi-session depending on the configuration. For ICMP, you cannot change from the default multi-session PAT. For more information about per-session PAT, see the firewall configuration guide.

New Connection Ownership

When a new connection is directed to a member of the cluster via load balancing, that unit owns both directions of the connection. If any connection packets arrive at a different unit, they are forwarded to the owner unit over the cluster control link. If a reverse flow arrives at a different unit, it is redirected back to the original unit.

Sample Data Flow

The following example shows the establishment of a new connection.



1. The SYN packet originates from the client and is delivered to one ASA (based on the load balancing method), which becomes the owner. The owner creates a flow, encodes owner information into a SYN cookie, and forwards the packet to the server.
2. The SYN-ACK packet originates from the server and is delivered to a different ASA (based on the load balancing method). This ASA is the forwarder.
3. Because the forwarder does not own the connection, it decodes owner information from the SYN cookie, creates a forwarding flow to the owner, and forwards the SYN-ACK to the owner.

4. The owner sends a state update to the director, and forwards the SYN-ACK to the client.
5. The director receives the state update from the owner, creates a flow to the owner, and records the TCP state information as well as the owner. The director acts as the backup owner for the connection.
6. Any subsequent packets delivered to the forwarder will be forwarded to the owner.
7. If packets are delivered to any additional units, it will query the director for the owner and establish a flow.
8. Any state change for the flow results in a state update from the owner to the director.

History for ASA Clustering on the Firepower 4100/9300

Feature Name	Version	Feature Information
Intra-chassis ASA Clustering for the Firepower 9300	9.4(1.150)	You can cluster up to 3 security modules within the Firepower 9300 chassis. All modules in the chassis must belong to the cluster. We introduced the following screen: **Configuration** > **Device Management** > **High Availability and Scalability** > **ASA Cluster Replication**