Strengthen Network Defenses through ASA Firewall Clustering

What You Will Learn

The Cisco® ASA 5585-X Adaptive Security Appliance supports clustering of multiple identical firewall nodes into one logical firewall. This capability:

- Achieves a higher throughput, connection rate, and number of concurrent connections
- Provides a predictive scalable solution
- Allows customers to buy according to their current needs and add more nodes as their traffic increases

The higher performance is especially relevant in data centers, where a large amount of traffic is processed in one place.

The Benefits of Clustering

Clustering of firewall nodes extends redundancy beyond the traditional active/standby redundancy provided on a device basis. Redundancy is now provided on a per-flow basis with up to 16 nodes in a single cluster. This shift greatly increases the availability of firewall protection and improves the stability of the network in general.

There is only one configuration for all the nodes in a cluster. Any change to the configuration is made only on the master node and is then propagated to all the slave nodes. This is one reason that all nodes in the cluster have to be identical. When the master node fails, a different node assumes the role of master with no manual intervention.

The node that receives the first packet in a flow is called the owner of the flow. Then another node is picked to be the director of the flow, based on a hash of the IP addresses and port numbers of the source and destination. All state information for the flow is then passed on to the director using a dedicated cluster control link (CCL). Thus, at all times, the state information of every flow is available at two nodes. The CCL is used solely for the nodes to share information among them about the different flows in the cluster and for data flows in cases of asymmetric traffic flows. The cluster control links are put in their own VLAN separate from the data links.

If the cluster control link of one node goes down, it is removed from the cluster and no traffic is sent to that node moving forward. So Cisco recommends providing redundancy for the cluster control links by putting them in an EtherChannel to guard against interface failures.

If the owner node fails for whatever reason, the packets for the flow are sent to a different node in the cluster. This node, if not the director of the flow, is able to find the director of the flow by using the same hash of the IP addresses and port numbers of the source and destination. The node then queries the director for state information of the flow. Then this node becomes the owner of that particular traffic flow. The director of flow is updated with the new owner information.

For asymmetric flows, where the returning packet is sent to a node different from the owner, the node that receives the packet is called the forwarder. The forwarder contacts the director of the flow to find its owner. Once the owner is known, the flow is forwarded to the owner over the cluster control link for the duration of the flow. In a network
with large number of asymmetric flows, it is recommended to assign an equal number of interfaces for the data link and the cluster control link. It is also another reason why all the nodes in the cluster must be identical, because different models have a different number of interfaces and capabilities.

A maximum of 16 nodes is supported in a single cluster. All these nodes are managed as a single logical firewall. However the ability to look at the statistics of the individual nodes in a cluster is also available. The cluster upgrade is achieved without any disruption in traffic because the nodes in the cluster are allowed to be at different minor versions during the upgrade.

In clustering, load-balancing of traffic across multiple ASA nodes is done in two ways. One is by configuring the data interfaces as spanned EtherChannel interfaces, and the other by configuring the data interfaces as individual interfaces.

**Spanned EtherChannel Interfaces**

In this type of deployment the EtherChannel does the job of load-balancing data traffic across the nodes of the cluster (see Figure 1).

**Figure 1.** Load Balancing by Spanned EtherChannel Interfaces

An EtherChannel aggregates multiple links between two devices to increase throughput. An added benefit is the high availability provided by redistributing traffic between the two devices if one interface were to fail. Link Aggregation Control Protocol (LACP) allows dynamic negotiation and the establishment of an EtherChannel between two devices. Cluster LACP (cLACP) implemented on the ASA makes multiple ASA nodes in the cluster appear as one logical firewall to the switch they are connected to. This is achieved by bundling multiple interfaces on different nodes into a single big EtherChannel on the Cisco ASA side.

A virtual port channel (vPC) allows links that are physically connected to two Cisco Nexus® 7000 Series Switches to appear as a single EtherChannel to a Cisco ASA. A vPC also allows all links to actively forward traffic, resulting in maximum use of the hardware.
The Virtual Switching System (VSS) puts to use all available Layer 2 bandwidth across redundant Cisco Catalyst® 6500 Series Switches with even load balancing. VSS is also supported with ASA clustering.

Dynamic port priority allows devices to pick which ports are put in standby state when all the available ports cannot be put into active state in an EtherChannel. Most switches support up to 8 active and 8 standby ports in an EtherChannel. By disabling dynamic port priority on the Cisco ASA cluster, the number of active ports in an EtherChannel can be increased to 16 by putting even the standby ports in the active state.

The vPC and VSS bring device redundancy on the switching infrastructure by doubling the active links that can be used for data traffic, to a maximum of 32. This increase is accomplished by using the Cisco Nexus 7000 F-Series line cards when the vPC is in operation.

Only one logical IP and MAC address is shared by all the data interfaces of the cluster. Here, the firewall is deployed in either transparent mode or routed mode. In transparent mode, the IP address is assigned to the bridge group. In routed mode, the IP address is assigned to the EtherChannel, configured as a routed interface.

Cisco recommends the spanned EtherChannel interface type because there is no impact to traffic when ASA nodes are added or removed from the cluster. The ability to dynamically add and remove nodes from the cluster lends itself to meet the changing business needs of customers while providing high availability at all times.

It is recommended that the switches on the inside and outside interfaces of the firewall have the same hashing algorithm to avoid asymmetric routing, thereby helping to achieve better performance.

**Individual Interfaces**

When individual interfaces are deployed, the adjacent routers take the task of load balancing data traffic to different ASA nodes (see Figure 2). The load balancing is accomplished using policy-based routing or equal-cost multipathing.

![Figure 2. Load Balancing by Individual Interfaces](image-url)
With policy-based routing, the load balancing of traffic among ASA nodes is static. The load balancing is accomplished in many ways; it is based on the flow or the class of traffic, for example. It is recommended to implement Cisco IOS IP Service Level Agreements (IP SLAs) with object tracking to learn which nodes have joined or left the cluster.

With equal-cost multipathing (ECMP), the load balancing is done with the help of routing protocols, which are dynamic. ECMP is the preferred method of traffic load balancing on a per-flow basis. The individual ASA nodes run the dynamic routing protocols and establish independent adjacencies with the routers and maintain their own routing tables. In cases of node failure, we rely on the dynamic routing protocols to update the router of such a failure, and the traffic is blackholed until that update comes through. For this reason, we recommend lowering the hello and dead timers of the routing protocols to help ensure faster convergence.

Every ASA node in the cluster has its own IP address for the data interfaces and for use in load-balancing traffic. Because of this, the firewall is supported only in routed mode.

EtherChannel supports bundling multiple data ports on individual ASA nodes. You would have one local EtherChannel for every node in the cluster.

**Conclusion**

As digitization spreads far and wide across the globe, data traffic keeps growing with more people and devices get connected. This brings with it many opportunities as well as challenges. Enterprises and service providers have to maintain highly available networks with low latency while ensuring data security at all times. Also many of today’s applications are bursty in nature. Firewall clustering with Cisco ASA provides a highly available solution, which scales throughput and more importantly connection rate to suit business needs. Cisco highly recommends deploying clustering where load balancing is done by spanned EtherChannel Interfaces for its simplicity in configuration and reliability.

**For More Information**

Cisco ASA 5500-X Configuration Guide