How does the Cluster Control Protocol function in working and failure scenarios for gateway clusters?

The Cluster Control Protocol (CCP) is a proprietary Check Point protocol. It is the basis of Check Point High Availability (CPHA) and new synchronization functionality. This protocol uses UDP port 8116 for internal communications between CPHA cluster members or VPN-1 Pro Security Gateways, using third-party High Availability solutions.

Cluster Control Protocol Modes

ClusterXL Cluster Control Protocol (CCP) packets used to be sent in "old-broadcast" (i.e. destination MAC address ff:ff:ff:ff:ff:ff). Since NG FP3 HF2, Cluster Control Protocol packets are sent to all interfaces in multicast by default (i.e., destination MAC address 01:00:5e:x:y:z). The layer 2 multicast MAC destination address is designed by default, whether in New Mode High Availability, or Unicast and Multicast Load Sharing Modes. If the connecting switch is incapable of forwarding the multicast MAC address, Cluster Control Protocol (CCP) can be changed to use broadcast instead of multicast.

To toggle between these two modes, run the following commands on the cluster gateway members.

Switching to broadcast mode:

```
cphaconf set_ccp broadcast
```

Switching to multicast mode:

```
cphaconf set_ccp multicast
```

Cluster Control Protocol (CCP) packets will be sent out in broadcast (destination MAC address ff:ff:ff:ff:ff:ff) after the command `cphaconf set_ccp broadcast` is run. No reboot or `cpstop` and `cpstart` are needed. The change from multicast to broadcast will survive a reboot. The change from the above commands is reflected in the file `$FWDIR/boot/ha_boot.conf` (UNIX) or `%FWDIR%\boot\ha_boot.conf` (Windows).

Default settings in the `ha_boot.conf` file are:

```
ha_installed 1
ccp_mode multicast
```

After the command `cphaconf set_ccp broadcast` is run, the `ha_boot.conf` file will look like the following:

```
ha_installed 1
ccp_mode broadcast
```

State Synchronization

When using ClusterXL, UDP port 8116 traffic is sent on all interfaces of the gateway cluster members (except those in the $FWDIR/conf/discntd.if file for UNIX or %FWDIR%\conf\discntd.if file for Windows). UDP port 8116 traffic cannot be turned off unless ClusterXL is turned off in the `cpconfig` utility, or the `cphastop` command runs on the VPN-1 Pro Security Gateway. When using third-party cluster solutions, UDP port 8116 traffic is only sent out through the synchronization interfaces.

UDP port 8116 traffic is necessary for cluster status health checks when a Check Point ClusterXL clustering solution is implemented. Third-party OPSEC clustering solutions conduct cluster status health checks, notwithstanding State Synchronization interfaces.

It is important to note that the Check Point High Availability (CPHA) is located between the VPN-1/FireWall-1 kernel and the network interface card (NIC). The Security Policy cannot block the synchronization data (UDP port 8116). This is the reason it is not necessary to create explicit rules accepting UDP port 8116 traffic in the SmartDashboard Rule Base for the Security Policy to be installed on the gateway cluster. This is also the reason CCP packets should be captured by the snoop, tcpdump, or ethereal utilities rather than the fw monitor packet capture utility. To stop the Check Point High Availability module from functioning and to completely prevent UDP 8116 data from being sent out, disable the Check Point High Availability (CPHA) in the `cpconfig` menu on the VPN-1/FireWall-1 gateway.

In VPN-1/FireWall-1 NG FP3, OPSEC cluster solutions like Nokia VRRP sent UDP port 8116 Cluster Control Protocol (CCP) packets on all interfaces besides the secured synchronization interface. This occurred when VPN-1/FireWall-1 was first brought up and before the Security Policy was installed. At the same time, VPN-1/FireWall-1 did not recognize which interface was the secure interface, so sent traffic to all interfaces. After the Security Policy was installed, the cluster stopped sending packets through the other interfaces, and only sent them through secured interfaces. This behavior has changed since NG with Application Intelligence. Cluster Control Protocol (CCP) packets are only sent through synchronization interfaces when using OPSEC clustering solutions.

State Synchronization is the mechanism which synchronizes the connections tables between the cluster gateway members. State Synchronization in NG with Application Intelligence uses the following two new synchronization modes:

Full Synchronization

Full Synchronization synchronizes the whole connections table when a member joins a cluster, or when a member comes up from reboot, or `cpstop` and `cpstart.` Full Synchronization uses TCP port 256

Delta Synchronization

Delta Synchronization is a constant update of the new connections among cluster members. Delta synchronization uses UDP port 8116.

Role of the Cluster Control Protocol (CPP)

The Cluster Control Protocol (CCP) serves an integral role in the operation of ClusterXL. Specifically, Cluster Control Protocol (CCP) is responsible for the following:

- Health-status Reports
- Cluster-member Probing
- State-change Commands
- Querying for Cluster Membership
- State-table Synchronization

Health-Status Reports

Cluster Control Protocol (CCP) reports the status of a cluster member roughly three times a second, per interface. These reports contain the state of the transmitting cluster member, as well as the presumed state of the other cluster members.

Cluster-member Probing

If a cluster member fails to receive status for another member on a given segment, Cluster Control Protocol (CCP) will probe that segment in an attempt to elicit a response. The purpose of such probes is to detect the nature of possible interface failures, and to determine which module has the problem. The outcome of this probe will determine the action taken next.

State-change Commands

If a cluster member wishes to change state, the command to do so takes place on the defined secured interface.

Querying-cluster Membership

When a cluster member comes online, it will send a series of Cluster Control Protocol (CCP) query/response messages, to gain knowledge of cluster membership.

State-table Synchronization

When State Synchronization is enabled, connection information is updated between cluster members on the defined secured interface.

How are IP addresses assigned within the gateway cluster?

Load Sharing Multicast, Load Sharing Unicast, and High Availability New modes use unique, real IP addresses for the cluster members interfaces. The cluster members physical IP addresses do not have to be routable on the Internet. The cluster itself uses virtual IP addresses for the internal and external cluster interfaces. Only the external cluster virtual IP address must be routable.

High Availability Legacy mode uses shared IP addresses on the internal and external cluster member interfaces. The SmartCenter Server must not be connected to these shared interfaces.

How ClusterXL responds to several types of failures in New Mode High Availability

Interface Failure

The scenario assumes two cluster members where the external interface of the primary has failed:

1. At the point of failure, the primary member will recognize that no Cluster Control Protocol (CCP) messages have been heard on the failed interface. As such, it announces via FWHA_MY_STATE on all other segments, that there may be an issue in the inbound direction with one of the interfaces.

2. The primary will also note that no Cluster Control Protocol (CCP) responses have been received on the failed interface. This causes the primary to announce on all other segments via FWHA_MY_STATE that the outbound direction for one of the interfaces is in question.

3. At the same time as the above events, the secondary member will recognize that no CCP packets have been received, and will begin sending FWHA_PROBE_REQ messages on the affected segment. In addition, the secondary will attempt ARP requests to hosts belonging to the affected segment, and will begin pings to those hosts that respond. This is done in an attempt to diagnose which member has the problem. The pings will continue as long as CCP packets cannot identify by other means that the interface is alive. This will happen when there are N cluster members, and N-1 of them are down. When more than two members are present, such pings will only be issued if all other cluster members do not respond to Cluster Control (CCP) probing.

4. Since no FWHA_PROBE_RPLY message is received as a response, but the ping requests are being answered, the secondary member concludes that its own interfaces are up and working, and that the interface of the primary has failed. Therefore, it announces via FWHA_MY_STATE that all of its own interfaces are operational.

5. With this report from the secondary, the primary concludes the issue is with its own interface, and moves its Check Point High Availability (CPHA) status to Down state.

6. The secondary issues gratuitous ARPs for both the physical and cluster address per IP segment, and moves itself to the Active/Active-Attention state.

Reboot

The scenario assumes two cluster members, in which a reboot has occurred:

1. As the primary goes down, it changes its state to Down/Dead, and announces this as part of FWHA_MY_STATE.

2. This triggers the secondary to prepare itself to become the active member. It does so by sending a series of gratuitous ARPs for both its physical IP and cluster IP for each clustered segment. This will update all necessary hosts/routers on each segment with the relevant updated MAC address information.

3. The secondary now moves to the Active/Active-Attention state, and assumes responsibility for processing all connections.

4. Though the primary is now considered Down/Dead, it will still be able to send/receive Cluster Control Protocol (CCP) packets until its interfaces are brought down. Once this occurs, the secondary, which is now in the Active/Active-Attention state, will make notice of the fact that no Cluster Control Protocol (CCP) packets are being received.

5. The secondary will do several things in an effort to ascertain why no Cluster Control Protocol (CCP) packets are being received. First, it sends FWHA_IF_PROB_Req packets on all segments in which no CCP packets have been heard. This is to elicit a response from any member capable of responding. The secondary will ARP on each segment for IPs belonging to that segment, and ping those hosts that respond. The pings will continue as long as CCP packets cannot identify that the interface is alive. This will happen when there are N cluster members, and N-1 of them are down. When more than two members are present, such pings will only be issued if all other cluster members do not respond to CCP probing.

6. Once the primary has rebooted, but before the policy is installed, it will begin sending FWHA_IFCONF_REPLY packets regularly. The primary node does so without knowing which cluster it belongs to, so a random ID is used.

7. After the primary learns the cluster ID, it begins sending FWHA_MY_STATE. The primary at this stage announces itself as Down/Dead.

8. The primary fetches the policy from another cluster member if possible, otherwise from the SmartCenter Server.

9. The primary now initiates full synchronization on the secured interface via the FW1 protocol.