ClusterXL

NG with Application Intelligence (R55)

IMPORTANT
Check Point recommends that customers stay up-to-date with the latest service packs and versions of security products, as they contain security enhancements and protection against new and changing attacks.

For additional technical information about Check Point products, consult Check Point's SecureKnowledge at:

http://support.checkpoint.com/kb/

See the latest version of this document in the User Center at:

http://www.checkpoint.com/support/technical/documents/docs_r55.html
productive use of gd. If you have questions, ask. "Derived works" includes all programs that utilize the library. Credit must be given in user-accessible documentation. This software is provided "AS IS." The copyright holders disclaim all warranties, either express or implied, including but not limited to implied warranties of merchantability and fitness for a particular purpose, with respect to this code and accompanying documentation. Although their code does not appear in gd 2.0.4, the authors wish to thank David Koblas, David Rowley, and Hutchison Avenue Software Corporation for their prior contributions.

Check Point Software Technologies Ltd.

U.S. Headquarters: 800 Bridge Parkway, Redwood City, CA 94065. Tel: (650) 628-2000 Fax: (650) 654-4233, info@CheckPoint.com
International Headquarters: 3A Jabotinsky Street, Ramat Gan, 52520, Israel, Tel: 972-3-753 4555 Fax: 972-3-575 9256, http://www.checkpoint.com
# Table Of Contents

## Chapter 1 Introduction to ClusterXL
- Summary of Contents 9
- The Need for Gateway Clusters 10
  - Reliability through High Availability 10
  - Enhanced Reliability and Performance through Load Sharing 10
- Check Point ClusterXL Gateway Clustering Solution 10
- The Cluster Control Protocol 11
- Installation, Licensing and Platform Support 11
- Clustering Definitions and Terms 12

## Chapter 2 Synchronizing Connection Information Across the Cluster
- The Need to Synchronize Cluster Information 15
- The Check Point State Synchronization Solution 16
  - Introduction to State Synchronization 16
  - The Synchronization Network 16
  - How State Synchronization Works 17
  - Non-Synchronized Services 18
  - Choosing Services That Don’t Require Synchronization 18
  - Duration Limited Synchronization 19
  - Non-Sticky connections 19
  - Example of a Non-Sticky Connection: The TCP 3-Way Handshake 20
  - How the Synchronization Mechanism Handles Non-Sticky Connections 21
  - Synchronizing Clusters over a Wide Area Network 22
  - Synchronized Cluster Restrictions 22
- Configuring State Synchronization 23
  - Configuring State Synchronization 23
  - Configuring a Service so that it is not Synchronized 24
  - How to create a Synchronized and an Unsynchronized Version of the Same Service 24
  - Configuring Duration Limited Synchronization 25

## Chapter 3 High Availability and Load Sharing in ClusterXL
- Introduction to High Availability and Load Sharing 27
  - Load Sharing 28
  - High Availability 28
- Example ClusterXL Topology 29
  - Defining the Cluster Virtual IP Addresses 30
  - Defining the Cluster Member IP Addresses 30
  - The Synchronization Network 31
  - Associating the Cluster Virtual IP addresses with the Cluster Member IP addresses 31
- ClusterXL Modes 32
  - Introduction to ClusterXL Modes 32
Chapter 6  ClusterXL Advanced Configuration

Upgrading ClusterXL Clusters 110
Working with VPN-1 Pro and Clusters 110
How to Configure VPN and Clusters 110
How to Define a Cluster Object for a VPN Peer with a Separate Manager 111
Working with NAT and Clusters 111
Cluster Fold and Cluster Hide 111
Configuring NAT on the Gateway Cluster 112
Configuring NAT on a Cluster Member 112
Working with VLANs and Clusters 113
VLAN Support in ClusterXL 113
Connecting Several Clusters on the Same VLAN 113
Advanced Cluster Configuration using Module Configuration Parameters 117
How to Configure Module Configuration Parameters 117
How to Configure Module Configuration Parameters to Survive a Boot 117
Controlling the Clustering and Synchronization Timers 119
Blocking New Connections Under Load 119
Working with SmartView Tracker Active Mode 120
Reducing the Number of Pending Packets 121
Configuring Full Synchronization Advanced Options 121
Defining Disconnected Interfaces 122
Defining a Disconnected Interface on Unix 123
Defining a Disconnected Interface on Windows 123
Enhanced Enforcement of the TCP 3-Way Handshake  123
Configuring Cluster Addresses on Different Subnets  124
  Introduction 124
  Configuration 124
  Example 126
  Limitations 127
Moving from High Availability Legacy to High Availability New Mode or Load Sharing with
  Minimal Effort 129
  On the Modules 130
  On the Management 130
Moving from High Availability Legacy to High Availability New Mode or Load Sharing with
  Minimal Downtime 131
Moving from a Single Gateway to a ClusterXL Cluster  133
  On the Single Gateway Machine 133
  On Machine 'B' 133
  On the Management 133
  On Machine 'A' 133
  On the Management 134
Adding Another Member to an Existing Cluster  134
Configuring ISP Redundancy on a Cluster  135

Appendix A  High Availability Legacy Mode
  Introduction to High Availability Legacy Mode  137
  Example of High Availability HA Legacy Mode Topology 138
    Shared Interfaces IP and MAC Address Configuration 138
    The Synchronization Interface 139
  Implementation Planning Considerations for HA Legacy Mode 139
    IP Address Migration 139
    SmartCenter Server Location 140
    Routing Configuration 140
    Switch (Layer 2 Forwarding) Considerations 140
  Configuring High Availability Legacy Mode 141
    Routing Configuration 141
    SmartDashboard configuration 142

Appendix B  Example cphaprob Script
  More information 143
  The clusterXL_monitor_process script 143

Appendix C  ClusterXL Command Line Interface

Index  149
CHAPTER 1

Introduction to ClusterXL

In This Chapter

Summary of Contents  page 9
The Need for Gateway Clusters  page 10
Check Point ClusterXL Gateway Clustering Solution  page 10
The Cluster Control Protocol  page 11
Installation, Licensing and Platform Support  page 11
Clustering Definitions and Terms  page 12

Summary of Contents

- Chapter 1, “Introduction to ClusterXL briefly describes the need for Gateway Clusters, introduces ClusterXL and the Cluster Control Protocol, specifies installation and licensing requirements, and lists some clustering definitions and terms.
- Chapter 2, “Synchronizing Connection Information Across the Cluster” describes State Synchronization, what not to synchronize, and how to configure State Synchronization.
- Chapter 3, “High Availability and Load Sharing in ClusterXL describes the ClusterXL Load Sharing and High Availability modes, talks about failover and the compatibility with other Check Point software and hardware.
- Chapter 4, “Working with OPSEC Certified Clustering Products describes the special considerations for working with OPSEC clustering products.
The Need for Gateway Clusters

Reliability through High Availability

Firewalls and VPN connections are business critical devices for an organization. A failure of the firewall or the VPN connection results in immediate loss of active connections in and out of the organization. Many of these connections, such as financial transactions, may be mission critical, and losing them will result in loss of critical data.

Firewalls and VPN connections must therefore be highly available. The gateway between the organization and the world must remain open, under all conceivable circumstances.

Enhanced Reliability and Performance through Load Sharing

In a Load Sharing Gateway Cluster, all the machines in the Cluster are active at all times. This makes the cluster more reliable, because if one machine fails or is brought down for maintenance, the remaining machines are already active and working. They do not have to be “woken up”.

Load Sharing also brings significant performance advantages. Putting to work multiple Gateways instead of a single Gateway provides linear performance increases for CPU intensive applications, such as VPNs, Security Servers, Policy Servers, and SmartDirectory (LDAP).

Check Point ClusterXL Gateway Clustering Solution

ClusterXL is a software-based Load Sharing and High Availability solution that distributes network traffic between clusters of redundant VPN-1 Pro Gateways, and provides transparent failover between machines in a cluster.

A VPN-1 Pro Gateway cluster is a group of identical VPN-1 Pro Gateways that are connected in such a way that if one fails, another immediately take its place (FIGURE 1-1).
ClusterXL uses unique physical IP and MAC addresses for the cluster member, and virtual IP addresses to represent the cluster itself. Virtual addresses (in all configurations other than High Availability Legacy mode) do not belong to any real machine interface.

ClusterXL supplies an infrastructure that ensures that no data is lost in case of a failure, by making sure each gateway cluster is aware of the connections going through the other members. Passing information about connections and other VPN-1 Pro states between the cluster members is called State Synchronization.

VPN-1 Pro Gateway Clusters can also be built using OPSEC certified High Availability and Load Sharing products. OPSEC Certified Clustering products use the same State Synchronization infrastructure as ClusterXL.

**The Cluster Control Protocol**

The Cluster Control Protocol (CCP) is the glue that links together the machines in the Check Point Gateway cluster. CCP traffic is distinct from ordinary network traffic, and can be seen using any network sniffer.

CCP runs on UDP port 8116, and has the following roles:

- Allows cluster members to report their own states and learn about the states of other members, by sending keep-alive packets (applies only to ClusterXL clusters).
- State synchronization.

Check Point's Cluster Control Protocol is used by each of the four ClusterXL modes, as well as by OPSEC clusters. However, the tasks performed by this protocol, and the manner in which they are implemented, may differ between the modes.

**Installation, Licensing and Platform Support**

ClusterXL must only be installed in a distributed configuration, in which the SmartCenter Server and the Cluster members are on different machines. ClusterXL is part of the standard VPN-1 Pro installation.

To install a policy on a gateway cluster:
You must have a license for VPN-1 Pro (with SKU: CPMP-VPG) installed on at least one of the cluster members. For Check Point Express you must have the matching Express license (with SKU: CPXP-VPX) installed on at least one of the cluster members.

On the other member(s) it is possible to install a secondary module license with SKU: CPMP-HVPG and for a Check Point Express with SKU: CPXP-HVPX.

For each ClusterXL Load Sharing cluster there are two supplemental management add-on licenses. At least one of the products with the following SKUs are required on the SmartCenter Server: CPMP-CXLS-U-NG or CPMP-CXLS-500-NG or CPMP-CXL-HA-1-NG.

After upgrading to NG with Application Intelligence (R.55), a previous version license for ClusterXL automatically counts as a legitimate Load Sharing license eliminating the requirement in point 3.

ClusterXL is currently available for Windows NT, Windows 2000, Windows 2003 Server, Solaris, Linux and SecurePlatform. See the platform support matrix in the Release Notes at:

http://www.checkpoint.com/techsupport/installation/ng/release_notes.html

### Clustering Definitions and Terms

Different vendors give different meanings to terms that relate to gateway clusters, High Availability and Load Sharing. Check Point uses the following definitions in discussing clustering:

**Cluster**
A group of machines that work together to provide Load Sharing and/or High Availability.

**Failure**
A hardware or software problem that causes a machine to be unable to filter packets. A failure of an Active machine leads to a Failover.

**Failover**
A machine taking over packet filtering in place of another machine in the cluster that suffered a Failure.
High Availability
The ability to maintain a connection when there is a Failure by having another machine
in the cluster take over the connection, without any loss of connectivity. Only the
Active machine filters packets, and the others do not. One of the machines in the
cluster is configured as the Active machine. If a Failover occurs on the Active machine,
one of the other machines in the cluster assumes its responsibilities.

Active Up
When the High Availability machine that was Active and suffered a Failure becomes
available again, it returns to the cluster, not as the Active machine but as one of the
standby machines in the cluster.

Primary Up
When the High Availability machine that was Active and suffered a Failure becomes
available again, it resumes its responsibilities as the Primary machine.

Hot Standby
Also known as Active/Standby. Means the same as High Availability.

Load Sharing
Also known as Active/Active. All machines in the cluster filter packets. Load Sharing
provides High Availability, and gives transparent Failover to any of the other machines
in the cluster when a Failure occurs.

Multicast Load Sharing
In Load Sharing Multicast mode of ClusterXL, every member of the cluster receives all
the packets sent to the cluster IP address. A router or Layer 3 switch forwards packets to
all cluster members using multicast. A ClusterXL decision algorithm on all cluster
members decides which cluster member should perform enforcement processing on the
packet.

Unicast Load Sharing
In Load Sharing Unicast mode of ClusterXL, one machine (the “Pivot”) receives all
traffic from a router with a unicast configuration, and redistributes the packets to the
other machines in the cluster. The Pivot machine is chosen automatically by ClusterXL.

Critical Device
A device which the administrator has defined to be critical to the operation of the
cluster member. A critical device is also known as a Problem Notification (pnote).
Critical devices are constantly monitored. If a critical device stops functioning, this is
declared as a Failure. A device can be hardware, or a process. The fwd and cphad
processes are predefined by default as critical devices. The Security Policy is also predefined as a critical device. The administrator can add to the list of critical devices using the `cphaprob` command.

**State Synchronization**

The technology that maintains connections after Failover. State Synchronization is used by both ClusterXL and third-party clustering solutions. It works by replicating VPN-1 Pro kernel tables.

**Secured interface**

An interface on a secure network. The synchronization network should be secured because of the sensitivity of the data that passes across it. One way of securing a network is to ensure that all interfaces connected to it are in a single locked room. Connecting the synchronization interfaces via a cross cable is another way of securing an interface.
Synchronizing Connection Information Across the Cluster

In This Chapter

The Need to Synchronize Cluster Information ... page 15
The Check Point State Synchronization Solution ... page 16
Choosing Services That Don’t Require Synchronization ... page 18
Configuring State Synchronization ... page 23

The Need to Synchronize Cluster Information

Firewalls are critical devices for an organization. A failure of the firewall will result in immediate loss of active connections in and out of the organization. Many of these connections, such as financial transactions, may be mission critical, and losing them will result in loss of critical data. ClusterXL supplies an infrastructure that ensures that no data is lost in case of a failure, by making sure each gateway cluster is aware of the connections going through the other members. Passing information about connections and other VPN-1 Pro states between the cluster members is called State Synchronization.
The Check Point State Synchronization Solution

In This Section

- Introduction to State Synchronization  page 16
- The Synchronization Network  page 16
- How State Synchronization Works  page 17
- Non-Synchronized Services  page 18
- Choosing Services That Don’t Require Synchronization  page 18
- Duration Limited Synchronization  page 19
- Non-Sticky connections  page 19
- Example of a Non-Sticky Connection: The TCP 3-Way Handshake  page 20
- How the Synchronization Mechanism Handles Non-Sticky Connections  page 21
- Synchronizing Clusters over a Wide Area Network  page 22
- Synchronized Cluster Restrictions  page 22

Introduction to State Synchronization

State Synchronization enables all machines in the cluster to be aware of the connections passing through each of the other machines. It ensures that if there is a failure in a cluster member, connections that were handled by the failed machine will be maintained by the other machines.

Every IP based service (including TCP and UDP) recognized by VPN-1 Pro is synchronized.

State Synchronization is used both by ClusterXL and by third-party OPSEC-certified clustering products.

Machines in a ClusterXL Load Sharing configuration must be synchronized. Machines in a ClusterXL High Availability configuration do not have to be synchronized, though if they are not, connections will be lost upon failover.

The Synchronization Network

The Synchronization Network is used to transfer synchronization information about connections and other VPN-1 Pro states between cluster members.
Because the synchronization network carries the most sensitive Security Policy information in the organization, it is important to make sure that it is secured against both malicious and unintentional interference. It is therefore recommended to secure the synchronization interfaces by:

- using a dedicated synchronization network, and
- connecting the physical network interfaces of the cluster members directly using a cross-cable. In a cluster with three or more members, use a dedicated hub or switch.

**Note** - Starting NG with Application Intelligence (R55), it is possible to run synchronization across a WAN. For details, see “Synchronizing Clusters over a Wide Area Network” on page 22.

Following these recommendations guarantees the safety of the synchronization network because no other networks carry synchronization information.

It is possible to define more than one synchronization network for backup purposes. It is recommended that the backup be a dedicated network.

As of NG with Application Intelligence and later version clusters, it is possible, though less secure, to use one of the existing internal or DMZ cluster networks for the backup synchronization network. There are two restrictions to the synchronization network: First, VLANs cannot be used in the synchronization network in any version. Second, in older versions, the interface used for the synchronization network must be a real interface with a real IP address (as opposed to a cluster IP or a virtual IP).

### How State Synchronization Works

Synchronization works in two modes:

- **Full sync.** transfers all VPN-1 Pro kernel table information from one cluster member to another. It is handled by the `fwd` daemon using an encrypted TCP connection.
- **Delta sync.** transfers changes in the kernel tables between cluster members. Delta sync is handled by the VPN-1 Pro-1 kernel using UDP multicast on port 8116.

Full sync. is used for initial transfers of state information, for many thousands of connections. If a cluster member is brought up after being down, it will perform full sync. Once all members are synchronized, only updates are transferred via delta sync. Delta sync is much quicker than full sync.
State Synchronization traffic typically makes up around 90% of all Cluster Control Protocol (CCP) traffic. State Synchronization packets are distinguished from the rest of CCP traffic via an opcode in the UDP data header.

**Non-Synchronized Services**

In a gateway cluster, all connections on all cluster members are normally synchronized across the cluster. However, not all services that cross a gateway cluster need necessarily be synchronized.

- It is possible to decide not to synchronize TCP, UDP and Other types of service. By default, all these services are synchronized.
- The VRRP and IP Clustering control protocols, as well as the IGMP protocol, are not synchronized by default (although you can choose to turn on synchronization for these protocols). Protocols that run solely between cluster members need not be synchronized. Although it is possible to synchronize them, no benefit will be gained if the cluster is configured to do so. The synchronization information is not relevant for this case because it will not help in case of a failover. Therefore the following protocols are not synchronized by default: IGMP, VRRP, IP clustering and some other OPSEC cluster control protocols.
- Broadcasts and multicasts are not synchronized, and cannot be synchronized.

It is possible to have both a synchronized service and a non-synchronized definition of a service, and to use them selectively in the Rule Base.

**Choosing Services That Don't Require Synchronization**

Synchronization has some performance cost. You can decide not to synchronize a service if all the following conditions are true:

1. A significant proportion of the traffic crossing the cluster uses a particular service. Not synchronizing the service reduces the amount of synchronization traffic, thereby enhancing cluster performance.
2. The service usually opens short connections, whose loss may not be noticed. DNS (over UDP) and HTTP are typically responsible for most connections, and on the other hand frequently have very short life and inherent recoverability in the application level. Services which typically open long connections, such as FTP, should always be synchronized.
3 Configurations that ensure bi-directional stickiness for all connections do not require synchronization to operate (only to maintain High Availability). Such configurations include:

- Any cluster in High Availability mode (for example, ClusterXL New HA or Nokia VRRP)
- ClusterXL in a Load Sharing mode with clear connections (no VPN or static NAT)
- OPSEC clusters that guarantee full stickiness (refer to the OPSEC cluster's documentation)

VPN and Static NAT connections passing through a ClusterXL cluster in a Load Sharing mode (either multicast or unicast) may not maintain bi-directional stickiness; hence, State Synchronization must be turned on for such environments.

To configure a service so that it will not be synchronized, edit the Service object. See “Configuring a Service so that it is not Synchronized” on page 24.

**Duration Limited Synchronization**

Some TCP services (HTTP for example) are characterized by connections with a very short duration. There is no point in synchronizing these connections because every synchronized connection consumes gateway resources, and the connection is likely to have finished by the time a failover occurs.

For all TCP services whose Protocol Type (that is defined in the GUI) is HTTP or None, you can use this option to delay telling VPN-1 Pro about a connection, so that the connection will only be synchronized if it still exists x seconds after the connection is initiated. This feature requires a SecureXL device that supports “Delayed Notifications” and the current cluster configuration (such as Performance Pack with ClusterXL LS Multicast).

This capability is only available if a SecureXL-enabled device is installed on the VPN-1 Pro Gateway through which the connection passes.

The setting is ignored if connection templates are not offloaded from the ClusterXL-enabled device. See the SecureXL documentation for additional information.

**Non-Sticky connections**

A connection is called sticky if all packets of this connection are handled by a single cluster member. In a non-sticky connection, a reply packet returns through a different gateway than the original packet.
The synchronization mechanism knows how to properly handle non-sticky connections. In a non-sticky connection, a cluster member gateway can receive an out-of-state packet, which VPN-1 Pro normally drops because it poses a security risk.

In Load Sharing configurations, all cluster members are active, and in Static NAT and encrypted connections, the source and destination IP addresses change. Therefore, Static NAT and encrypted connections through a Load Sharing cluster may be non-sticky. Non-stickiness may also occur with Hide NAT, but ClusterXL has a mechanism to make it sticky.

In High Availability configurations, all packets reach the Active machine, so non-sticky connections do not occur. If failover occurs during connection establishment, the connection is lost, but synchronization can be performed later.

If the other members do not know about a non-sticky connection, the packet will be out-of-state, and the connection will be dropped for security reasons. However, the Synchronization mechanism knows how to inform other members of the connection. The Synchronization mechanism thereby prevent out-of-state packets in valid, but non-sticky connections, so that these non-sticky connections are allowed.

Non-sticky connections will also occur if the network administrator has configured asymmetric routing, where a reply packet returns through a different gateway than the original packet.

Normally, TCP streaming connections are not allowed to be non-sticky (meaning, be handled by a different cluster member on each direction) to avoid excessive synchronization.

By default, a TCP streaming that is determined to be non-sticky is reset by the detecting Firewall module. However, you can override this behavior, and allow non-sticky TCP streaming connections, by setting the global flag `fwtcpstr_reject_synced` to 0. To restore the default behavior, this flag should be set to 1.

**Example of a Non-Sticky Connection: The TCP 3-Way Handshake**

The 3-way handshake that initiates all TCP connections can very commonly lead to a non-sticky (often called asymmetric routing) connection. The following situation may arise:

Client A initiates a connection by sending a SYN packet to server B (see FIGURE 2-1). The SYN passes through Gateway C, but the SYN/ACK reply returns through Gateway D. This is a non-sticky connection, because the reply packet returns through a different gateway than the original packet.
Gateway D is notified of the SYN packet via the synchronization network. If gateway D is updated before the SYN/ACK packet sent by server B reaches this machine, the connection is handled normally. If, however, synchronization is delayed, and the SYN/ACK packet is received on gateway D before the SYN flag has been updated, then the gateway will treat the SYN/ACK packet as out-of-state, and will drop the connection.

See “Enhanced Enforcement of the TCP 3-Way Handshake” on page 123 for additional information.

**FIGURE 2-1** A Non-sticky (asymmetrically routed) connection

---

**How the Synchronization Mechanism Handles Non-Sticky Connections**

The synchronization mechanism prevents out-of-state packets in valid, but non-sticky connections. The way it does this is best illustrated with reference to the 3-way handshake that initiates all TCP data connections. The 3-way handshake proceeds as follows:

1. SYN (client to server)
2. SYN/ACK (server to client)
3. ACK (client to server)
4. Data (client to server)

To prevent out-of-state packets, the following sequence (called “Flush and Ack”) occurs (The step numbers correspond to the numbers in FIGURE 2-1):
Cluster member receives first packet (SYN) of a connection.
2 Suspects that it is non-sticky.
3 Hold the SYN packet.
4 Send the pending synchronization updates to all cluster members (including all changes relating to this packet).
5 Wait for all the other cluster members to acknowledge the information in the sync packet.
6 Release held SYN packet.
7 All cluster members are ready for the SYN-ACK.

**Synchronizing Clusters over a Wide Area Network**

Organizations are sometimes faced with the need to locate cluster members in geographical locations that are distant from each other. A typical example is a replicated data center whose locations are widely separated for disaster recovery purposes. In such a configuration it is clearly impractical to use a cross cable as the synchronization network (as described in “The Synchronization Network” on page 16).

Starting from NG with Application Intelligence (R55), the synchronization network can be spread over remote sites, which makes it easier to deploy geographically distributed clustering. There are two limitations to this capability:

1 The synchronization network must guarantee no more than 100ms latency and no more than 5% packet loss.
2 The synchronization network may only include switches and hubs. No routers are allowed on the synchronization network, because routers drop Cluster Control Protocol packets.

To monitor and troubleshoot geographically distributed clusters, a command line is available. See “Troubleshooting Synchronization (cphaprob [-reset] syncstat)” on page 88.

**Synchronized Cluster Restrictions**

The following restrictions apply to synchronizing cluster members:

1 Only cluster members running on the same platform can be synchronized.

For example, it is not possible to synchronize a Windows 2000 cluster member with a Solaris 8 cluster member.
2 The cluster members must be the same software version.
   For example, it is not possible to synchronize a Version NG FP3 cluster member with a version NG with application Intelligence cluster member.

3 A user-authenticated connection through a cluster member will be lost if the cluster member goes down. Other synchronized cluster members will be unable to resume the connection.
   However, a client-authenticated connection or session-authenticated connection will not be lost.
   The reason for these restrictions is that user authentication state is maintained on Security Servers, which are processes, and thus cannot be synchronized on different machines in the way that data can be synchronized. However, the state of session authentication and client authentication is stored in kernel tables, and thus can be synchronized.

4 The state of connections using resources is maintained in a Security Server, so these connections cannot be synchronized for the same reason that user-authenticated connections cannot be synchronized.

5 Accounting information is accumulated in each cluster member and reported separately to the SmartCenter Server, where the information is aggregated. In case of a failover, accounting information that was accumulated on the failed member but not yet reported to the SmartCenter Server is lost. To minimize the problem it is possible to reduce the period in which accounting information is “flushed”. To do this, in the cluster object’s Logs and Masters > Additional Logging page, configure the attribute Update Account Log every:.

Configuring State Synchronization

In This Section

Configuring State Synchronization  page 23
Configuring a Service so that it is not Synchronized  page 24

Configuring State Synchronization

Configure State synchronization as part of the process of configuring ClusterXL and OPSEC certified clustering products. Configuring State synchronization involves

• Setting up a synchronization network for the gateway cluster
Configuring State Synchronization

- Installing VPN-1 Pro and turning on the synchronization capability during the configuration phase.
- In SmartDashboard, ensuring State Synchronization is selected in Synchronization page of the cluster object.

For configuration details, see
- “Configuring Routing for the Client Machines” on page 44
- “Configuring OPSEC-Certified Clustering Products” on page 58

**Configuring a Service so that it is not Synchronized**

For background information about configuring services so that they are not synchronized, see “Non-Synchronized Services” on page 18.

1. In the Services branch of the objects tree, double click the TCP, UDP or Other type service that you do not wish to synchronize.
2. In the **Service Properties** window, click **Advanced** to display the **Advanced Services Properties** window.
3. Deselect **Synchronize connections on the cluster**.

**How to create a Synchronized and an Unsynchronized Version of the Same Service**

It is possible to have both a synchronized and a non-synchronized definition of the service, and to use them selectively in the Security Rule Base.

1. Define a new TCP, UDP and Other type service. Give it a name that distinguishes it from the existing service.
2. Copy all the definitions from the existing service into the **Service Properties** window of the new service.
3. In the new service, click **Advanced** to display the **Advanced Services Properties** window.
4. Copy all the definitions from the existing service into the **Advanced Service Properties** window of the new service.
5. Set **Synchronize connections on the cluster** in the new service, so that it is different from the setting in the existing service.
Configuring Duration Limited Synchronization

For background information about the synchronization of services that have limited duration, see “Duration Limited Synchronization” on page 19.

1. In the Services branch of the objects tree, double click the TCP, UDP or Other type service that you wish to synchronize.

2. In the Service Properties window, click Advanced to display the Advanced Services Properties window.

3. Select Start synchronizing x seconds after connection initiation.

   **Note**: As this feature is limited to HTTP-based services, the Start synchronizing - seconds after connection initiation checkbox is not displayed for other services.

4. In the seconds field, enter the number of seconds or select the number of seconds from the dropdown list, for which you want synchronization to be delayed after connection initiation.
High Availability and Load Sharing in ClusterXL

In This Chapter

- Introduction to High Availability and Load Sharing
- Example ClusterXL Topology
- ClusterXL Modes
- Failover
- Implementation Planning Considerations
- ClusterXL Hardware Requirements
- Configuring ClusterXL
- Check Point Software and Hardware Compatibility

Introduction to High Availability and Load Sharing

ClusterXL is a software-based Load Sharing and High Availability solution that distributes network traffic between clusters of redundant VPN-1 Pro gateways.

ClusterXL provides:
- Transparent failover in case of machine failures
- Zero downtime for mission-critical environments (when using State Synchronization)
- Enhanced throughput (in Load Sharing modes)
- Transparent upgrades
All machines in the cluster are aware of the connections passing through each of the other machines. The cluster members synchronize their connection and status information across a secure synchronization network.

The glue that binds the machines in a ClusterXL cluster is the Cluster Control Protocol (CCP), which is used to pass synchronization and other information between the cluster members.

**Load Sharing**

ClusterXL Load Sharing distributes traffic within a cluster of gateways so that the total throughput of multiple machines is increased.

In Load Sharing configurations, all functioning machines in the cluster are active, and handle network traffic (Active/Active operation).

If any individual Check Point gateway in the cluster becomes unreachable, transparent failover occurs to the remaining operational machines in the cluster, thus providing High Availability. All connections are shared between the remaining gateways without interruption.

**High Availability**

High Availability allows organizations to maintain a connection when there is a failure in a cluster member, without Load Sharing between cluster members. In a High Availability cluster, only one machine is active (Active/Standby operation). In the event that the active cluster member becomes unreachable, all connections are re-directed to a designated backup without interruption. In a synchronized cluster, the backup cluster members are updated with the state of the connections of the active cluster member.

In a High Availability cluster, each machine is given a priority. The highest priority machine serves as the gateway in normal circumstances. If this machine fails, control is passed to the next highest priority machine. If that machine fails, control is passed to the next machine, and so on.

Upon gateway recovery, it is possible to maintain the current active gateway (Active Up), or to switch to the highest priority gateway (Primary Up). Note that in Active Up configuration, changing and installing the Security Policy may restart the ClusterXL configuration handshake on the members, which may lead to another member being chosen as the Active machine.
Example ClusterXL Topology

In This Section

Defining the Cluster Virtual IP Addresses  page 30
Defining the Cluster Member IP Addresses  page 30
The Synchronization Network  page 31
Associating the Cluster Virtual IP addresses with the Cluster Member IP addresses  page 31

ClusterXL uses unique physical IP and MAC addresses for the cluster member, and virtual IP addresses to represent the cluster itself. Cluster interface addresses do not belong to any real machine interface.

FIGURE 3-1 shows a two-member ClusterXL cluster, and contrasts the virtual IP addresses of the cluster, and the physical IP addresses of the cluster members. Only one routable IP address is required in a ClusterXL cluster, for the virtual cluster interface that faces the Internet. All cluster member physical IP addresses can be non-routable, thus saving routable addresses.

Each cluster member has three interfaces: one external interface, one internal interface, and one for synchronization. Cluster member interfaces facing in each direction are connected via a switch, router, or VLAN switch.

All cluster members must be in the same network. For example, there must not be a router in between cluster members.

Refer to the sections following FIGURE 3-1 for a description of the ClusterXL configuration concepts shown in the example.

Note ·

1. High Availability Legacy Mode uses a different Topology, and is discussed in the Appendix: “High Availability Legacy Mode” on page 137.

2. In the examples in this and subsequent sections, addresses in the range 192.168.0.0 to 192.168.255.255 which are RFC 1918 private addresses are used to represent routable (public) IP addresses.
Defining the Cluster Virtual IP Addresses

In FIGURE 3-1, the IP address of the cluster is 192.168.10.100, and this is the only routable IP address in the cluster.

The cluster has an internal and an external virtual IP addresses. The external IP address is 192.168.10.100, and the internal IP address is 10.10.0.100.

The SmartCenter Management Server can be located anywhere, and should be routable to either the internal or external cluster addresses.

Defining the Cluster Member IP Addresses

The guidelines for configuring each cluster member machine are as follows:

- All machines within the cluster must have at least three interfaces:
  - an interface facing the external cluster interface, which in turn faces the internet
  - an interface facing the internal cluster interface, which in turn faces the internal network
• an interface to use for synchronization.

**Note** - NG with Application Intelligence presents an option to use only two interfaces per member, one external and one internal and to run sync over the internal interface. However, this configuration is not recommended and should be used for backup only. For more information see Chapter 2, “Synchronizing Connection Information Across the Cluster”.

All interfaces pointing in a certain direction must be on the same network.

For example, in the configuration in FIGURE 3-1, there are two cluster members, Member_A and Member_B. Each has an interface with an IP address facing the Internet through a hub or a switch. This is the External interface with IP address 192.168.10.1 on Member_A and 192.168.10.2 on Member_B, and is the interface that the cluster external interface sees.

### The Synchronization Network

State Synchronization between cluster members ensures that if there is a failover, connections that were handled by the failed machine will be maintained. The synchronization network is used to pass connection synchronization and other state information between cluster members. This network therefore carries all the most sensitive security policy information in the organization, and so it is important to make sure the network is secure. It is possible to define more than one synchronization network for backup purposes.

To secure the synchronization interfaces, they should be directly connected by a cross cable, or in the case of a three or more member cluster, by means of a dedicated hub, switch.

Machines in a Load Sharing cluster must be synchronized because SYNC is used in normal traffic flow. Machines in a High Availability cluster do not have to be synchronized, though if they are not, connections may be lost upon failover.

FIGURE 3-1 shows a SYNC interface with a unique IP address on each machine. 10.0.10.1 on Member_A and 10.0.10.2 on Member_B.

### Associating the Cluster Virtual IP addresses with the Cluster Member IP addresses

ClusterXL clusters require a unique IP address for every member. However, in some organizations, only one routable (real) IP address is available.

To allow organizations with only one routable address to use ClusterXL, it is possible to configure different subnets for the cluster IP addresses and the member addresses. For example, the cluster IP address could be 192.168.10.100 (a real routable address) and the
members IP addresses can be 172.16.0.1 and 172.16.0.2 (private, non-routable addresses). All member interfaces pointing in a certain direction must be on the same network.

Each virtual cluster IP address is mapped to the member interface addresses in the same direction. To do this, the concept of a member network is introduced.

If a member's interface IP address is on a different subnet than the cluster virtual IP address, the mapping of virtual cluster IP address to the member interface addresses is done by specifying in the cluster object the network that the members reside on, that is, the member network.

A member network can be configured on all cluster interfaces. In FIGURE 3-1, the internal cluster virtual IP address is in the same network as the member interfaces, so that the member network does not need to be explicitly defined.

A member network can be used in place of NAT for the cluster members themselves. This is because the real IP addresses of the cluster members are effectively hidden behind the cluster IP address. However, NAT must still be defined for nodes or networks behind the cluster.

**ClusterXL Modes**

In This Section

- *Introduction to ClusterXL Modes*  page 32
- *Load Sharing Multicast Mode*  page 33
- *Load Sharing Unicast Mode*  page 34
- *High Availability New Mode*  page 36
- *Mode Comparison Table*  page 38

**Introduction to ClusterXL Modes**

ClusterXL has four working modes. This section briefly describes each mode and its relative advantages and disadvantages.

- Load Sharing Multicast Mode
- Load Sharing Unicast Mode
- High Availability New Mode
- High Availability Legacy Mode
High Availability Legacy Mode is discussed in the Appendix chapter: “High Availability Legacy Mode” on page 137. It is recommended that you use High Availability New Mode to avoid problems with backward compatibility.

**Note** • All examples in the section refer to the ClusterXL configuration shown in FIGURE 3-1 on page 30.

### Load Sharing Multicast Mode

Load Sharing enables you to distribute network traffic between cluster members. In contrast to High Availability, where only a single member is active at any given time, all cluster members in a Load Sharing solution are active, and the cluster is responsible for assigning a portion of the traffic to each member. This assignment is the task of a decision function, which examines each packet going through the cluster, and determines which members should handle it. Thus, a Load Sharing cluster utilizes all cluster members, which usually leads to an increase in its total throughput. See Figure 3-1 on page 30 for an example of a typical ClusterXL configuration.

It is important to understand that ClusterXL Load Sharing, when combined with State Synchronization, provides a full High Availability solution as well. When all cluster members are active, traffic is evenly distributed between the machines. In case of a failover event, caused by a problem in one of the members, the processing of all connections handled by the faulty machine is immediately taken over by the other members.

ClusterXL offers two separate Load Sharing solutions: Multicast and Unicast. The two modes differ in the way members receive the packets sent to the cluster. This section describes the Multicast mode. For a description of Unicast mode see “Load Sharing Unicast Mode” on page 34.

The Multicast mechanism, which is provided by the Ethernet network layer, allows several interfaces to be associated with a single physical (MAC) address. Unlike Broadcast, which binds all interfaces in the same subnet to a single address, Multicast enables grouping within networks. This means that it is possible to select the interfaces within a single subnet that will receive packets sent to a given MAC address.

ClusterXL uses the Multicast mechanism to associate the virtual cluster IP addresses with all cluster members. By binding these IP addresses to a Multicast MAC address, it ensures that all packets sent to the cluster, acting as a gateway, will reach all members in the cluster. Each member then decides whether it should process the packets or not. This decision is the core of the Load Sharing mechanism: it has to assure that at least one member will process each packet (so that traffic is not blocked), and that no two members will handle the same packets (so that traffic is not duplicated).
An additional requirement of the decision function is to route each connection through a single gateway, to ensure that packets that belong to a single connection will be processed by the same member. Unfortunately, this requirement cannot always be enforced, and in some cases, packets of the same connection will be handled by different members. ClusterXL handles these situations using its State Synchronization mechanism, which mirrors connections on all cluster members.

**Example**

This scenario describes a user logging from the Internet to a web server behind the Firewall cluster that is configured in Load Sharing Multicast mode.

1. The user requests a connection from 192.168.10.78 (his computer) to 10.10.0.34 (the web server).
2. A router on the 192.168.10.x network recognizes 192.168.10.100 (the cluster's virtual IP address) as the gateway to the 10.10.0.x network.
3. The router issues an ARP request to 192.168.10.100.
4. One of the active members intercepts the ARP request, and responds with the Multicast MAC assigned to the cluster IP address of 192.168.10.100.
5. When the web server responds to the user requests, it recognizes 10.10.0.100 as its gateway to the Internet.
6. The web server issues an ARP request to 10.10.0.100.
7. One of the active members intercepts the ARP request, and responds with the Multicast MAC address assigned to the cluster IP address of 10.10.0.100.
8. All packets sent between the user and the web server reach every cluster member, which decides whether to handle or drop each packet.
9. When a failover occurs, one of the cluster members goes down. However, traffic still reaches all of the active cluster members, and hence there is no need to make changes in the network's ARP routing. All that changes is the cluster's decision function, which takes into account the new state of the members.

**Load Sharing Unicast Mode**

Load Sharing Unicast mode provides a Load Sharing solution adapted to environments where Multicast Ethernet cannot operate. In this mode a single cluster member, referred to as *Pivot*, is associated with the cluster's virtual IP addresses, and is thus the only member to receive packets sent to the cluster. The pivot is then responsible for propagating the packets to other cluster members, creating a Load Sharing mechanism. Distribution is performed by applying a decision function on each packet, the same way
it is done in Load Sharing Multicast mode. The difference is that only one member performs this selection: any non-pivot member that receives a forwarded packet will handle it, without applying the decision function. Note that non-pivot members are still considered as “active”, since they perform routing and Firewall tasks on a share of the traffic (although they do not perform decisions.).

Even though the pivot member is responsible for the decision process, it still acts as a Firewall module that processes packets (for example, the decision it makes can be to handle a packet on the local machine). However, since its additional tasks can be time consuming, it is usually assigned a smaller share of the total load.

When a failover event occurs in a non-pivot member, its handled connections are redistributed between active cluster members, providing the same High Availability capabilities of New High Availability and Load Sharing Multicast. When the pivot member encounters a problem, a regular failover event occurs, and, in addition, another member assumes the role of the new pivot. The pivot member is always the active member with the highest priority. This means that when a former pivot recuperates, it will retain its previous role.

See Figure 3-1 on page 30 for an example of a typical ClusterXL configuration.

**Example**

In this scenario, we use a Load Sharing Unicast cluster as the gateway between the user's computer and the web server.

1. The user requests a connection from 192.168.10.78 (his computer) to 10.10.0.34 (the web server).
2. A router on the 192.168.10.x network recognizes 192.168.10.100 (the cluster's virtual IP address) as the gateway to the 10.10.0.x network.
3. The router issues an ARP request to 192.168.10.100.
4. The pivot member intercepts the ARP request, and responds with the MAC address that corresponds to its own unique IP address of 192.168.10.1.
5. When the web server responds to the user requests, it recognizes 10.10.0.100 as its gateway to the Internet.
6. The web server issues an ARP request to 10.10.0.100.
7. The pivot member intercepts the ARP request, and responds with the MAC address that corresponds to its own unique IP address of 10.10.0.1.
8. The user's request packet reaches the pivot member on interface 192.168.10.1.
9 The pivot decides that the second member should handle this packet, and forwards it to 192.168.10.2.

10 The second member recognizes the packet as a forwarded one, and processes it.

11 Further packets are processed by either the pivot member, or forwarded and processed by the non-pivot member.

12 When a failover occurs on the pivot, the second member assumes the role of pivot.

13 The new pivot member sends gratuitous ARP requests to both the 192.168.10.x and the 10.10.0.x networks. These requests associate the virtual IP address of 192.168.10.100 with the MAC address that correspond to the unique IP address of 192.168.10.2, and the virtual IP address of 10.10.0.100 with the MAC address that correspond to the unique IP address of 10.10.0.2.

14 Traffic sent to the cluster is now received by the new pivot, and processed by the local machine (as it is currently the only active machine in the cluster).

15 When the first machine recovers, it re-assumes the role of pivot, by associating the cluster IP addresses with its own unique MAC addresses.

**High Availability New Mode**

The High Availability New Mode provides basic High-Availability capabilities in a cluster environment. This means that the cluster can provide Firewall services even when it encounters a problem, which on a stand-alone module would have resulted in a complete loss of connectivity. When combined with Check Point's State Synchronization, ClusterXL High Availability can maintain connections through failover events, in a user-transparent manner, allowing a flawless connectivity experience. Thus, High-Availability provides a backup mechanism, which organizations can use to reduce the risk of unexpected downtime, especially in a mission-critical environment (such as one involving money transactions over the Internet.)

To achieve this purpose, ClusterXL's New High Availability mode designates one of the cluster members as the active machine, while the rest of the members are kept in a stand-by mode. The cluster's virtual IP addresses are associated with the physical network interfaces of the active machine (by matching the virtual IP address with the unique MAC address of the appropriate interface). Thus, all traffic directed at the cluster is actually routed (and filtered) by the active member. The role of each cluster member is chosen according to its priority, with the active member being the one with the highest ranking. Member priorities correspond to the order in which they appear in the **Cluster Members** page of the **Gateway Cluster Properties** window. The top-most member has the highest priority. You can modify this ranking at any time.
In addition to its role as a Firewall gateway, the active member is also responsible for informing the stand-by members of any changes to its connection and state tables, keeping these members up-to-date with the current traffic passing through the cluster. Whenever the cluster detects a problem in the active member that is severe enough to cause a failover event, it passes the role of the active member to one of the standby machines (the member with the currently highest priority). If State Synchronization is applied, any open connections are recognized by the new active machine, and are handled according to their last known state. Upon the recovery of a member with a higher priority, the role of the active machine may or may not be switched back to that member, depending on the user's configuration.

It is important to note that the cluster may encounter problems in standby machines as well. In this case, these machines are not considered for the role of active members, in the event of a failover.

See Figure 3-1, “Example ClusterXL Topology,” on page 30 for an example of a typical ClusterXL configuration.

**Example**

This scenario describes a user logging from the Internet to a web server behind the Firewall cluster.

1. The user requests a connection from 192.168.10.78 (his computer) to 10.10.0.34 (the web server).
2. A router on the 192.168.10.x network recognizes 192.168.10.100 (the cluster's virtual IP address) as the gateway to the 10.10.0.x network.
3. The router issues an ARP request to 192.168.10.100.
4. The active member intercepts the ARP request, and responds with the MAC address that corresponds to its own unique IP address of 192.168.10.1.
5. When the web server responds to the user requests, it recognizes 10.10.0.100 as its gateway to the Internet.
6. The web server issues an ARP request to 10.10.0.100.
7. The active member intercepts the ARP request, and responds with the MAC address that corresponds to its own unique IP address of 10.10.0.1.
8. All traffic between the user and the web server is now routed through the active member.
9. When a failover occurs, the standby member concludes that it should now replace the faulty active member.
10 The stand-by member sends gratuitous ARP requests to both the 192.168.10.x and the 10.10.0.x networks. These requests associate the virtual IP address of 192.168.10.100 with the MAC address that correspond to the unique IP address of 192.168.10.2, and the virtual IP address of 10.10.0.100 with the MAC address that correspond to the unique IP address of 10.10.0.2.

11 The stand-by member has now switched to the role of the active member, and all traffic directed through the cluster is routed through this machine.

12 The former active member is now considered to be “down”, waiting to recover from whatever problem that had caused the failover event.

**Mode Comparison Table**

TABLE 3-1 summarizes the similarities and differences between the ClusterXL modes.

<table>
<thead>
<tr>
<th></th>
<th>Legacy High Availability</th>
<th>New High Availability</th>
<th>Load Sharing Multicast</th>
<th>Load Sharing Unicast</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Availability</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Load Sharing</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Performance</td>
<td>Good</td>
<td>Good</td>
<td>Excellent</td>
<td>Very Good</td>
</tr>
<tr>
<td>Hardware Support</td>
<td>All</td>
<td>All</td>
<td>Not all routers are supported</td>
<td>All</td>
</tr>
<tr>
<td>SecureXL Support</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, with Performance Pack or SecureXL Turbocard.</td>
<td>Yes</td>
</tr>
<tr>
<td>State Synchronization</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VLAN Tagging Support</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Failover

In This Section

- What is a Failover?  page 39
- When Does a Failover Occur?  page 40
- What Happens When a Gateway Recovers?  page 40
- How a Recovered Cluster Member Obtains the Security Policy  page 40

What is a Failover?

A failover occurs when a Gateway is no longer able to perform its designated functions. When this happens another Gateway in the cluster assumes the failed Gateway’s responsibilities.

In a Load Sharing configuration, if one VPN-1 Pro Gateway in a cluster of Gateways goes down, its connections are distributed among the remaining Gateways. All gateways in a Load Sharing configuration are synchronized, so no connections are interrupted.

In a High Availability configuration, if one Gateway in a synchronized cluster goes down, another Gateway becomes active and “takes over” the connections of the failed Gateway. If you do not use State Synchronization, existing connections are closed when failover occurs, although new connections can be opened.

To tell each cluster member that the other gateways are alive and functioning, the ClusterXL Cluster Control Protocol maintains a heartbeat between cluster members. If a certain predetermined time has elapsed and no message is received from a cluster member, it is assumed that the cluster member is down and a failover occurs. At this point another cluster member automatically assumes the responsibilities of the failed cluster member.

It should be noted that a cluster machine may still be operational but if any of the above checks fail in the cluster, then the faulty member initiates the failover because it has determined that it can no longer function as a cluster member.

Note that more than one cluster member may encounter a problem that will result in a failover event. In cases where all cluster members encounter such problems, ClusterXL will try to choose a single member to continue operating. The state of the chosen member will be reported as Active Attention. This situation lasts until another member fully recovers. For example, if a cross cable connecting the cluster members malfunctions, both members will detect an interface problem. One of them will change to the Down state, and the other to Active Attention.
When Does a Failover Occur?

A failover takes place when one of the following occurs on the active cluster member:

- Any critical device (such as fw) fails. A critical device is a process running on a cluster member that enables the member to notify other cluster members that it can no longer function as a member. The device reports to the ClusterXL mechanism regarding its current state or it may fail to report, in which case ClusterXL decides that a failover has occurred and another cluster member takes over.
- An interface or cable fails.
- The machine crashes.
- The Security Policy is uninstalled. When the Security Policy is uninstalled the Gateway can no longer function as a firewall. If it cannot function as a firewall, it can no longer function as a cluster member and a failover occurs. Normally a policy is not uninstalled by itself but would be initiated by a user.

What Happens When a Gateway Recovers?

In a Load Sharing configuration, when the failed Gateway in a cluster recovers, all connections are redistributed among all active members.

In a High Availability configuration, when the failed Gateway in a cluster recovers, the recovery method depends on the configured cluster setting. The options are:

- **Maintain Current Active Gateway** — If one machine passes on control to a lower priority machine, control will be returned to the higher priority machine only if the lower priority machine fails. This mode is recommended if all members are equally capable of processing traffic, in order to minimize the number of failover events.
- **Switch to Higher Priority Gateway** — If the lower priority machine has control and the higher priority machine is restored, then control will be returned to the higher priority machine. This mode is recommended if one member is better equipped for handling connections, so it will be the default gateway.

How a Recovered Cluster Member Obtains the Security Policy

The administrator installs the security policy on the cluster rather than separately on individual cluster members. The policy is automatically installed on all cluster members. The policy is sent to the IP address defined in the **General Properties** page of the cluster member object.

When a failed cluster member recovers, it will first try to take a policy from one of the other cluster members. The assumption is that the other cluster members have a more up-to-date policy. If this does not succeed, it compares its own local policy to the
policy on the SmartCenter Server. If the policy on the SmartCenter Server is more up-to-date than the one on the cluster member, the policy on the SmartCenter Server will be retrieved. If the cluster member does not have a local policy, it retrieves one from the SmartCenter Server. This ensures that all cluster members use the same policy at any given moment.

**Implementation Planning Considerations**

In This Section

- High Availability or Load Sharing
- Choosing the Load Sharing Mode
- IP Address Migration

**High Availability or Load Sharing**

Whether to choose a Load Sharing (Active/Active) or a High Availability (Active/Standby) configuration depends on the need and requirements of the organization. A High Availability gateway cluster ensures fail-safe connectivity for the organization. Load Sharing provides the additional benefit of increasing performance. For a list of supported hardware devices and their configuration instructions see “Hardware Compatibility for Load Sharing Multicast Mode” on page 53.

**Choosing the Load Sharing Mode**

Load Sharing Multicast mode is the most efficient Load Sharing Mode because load is distributed optimally between all cluster members. However, not all routers can be used for Load Sharing Multicast mode. Load Sharing Multicast mode associates a multicast MAC with each unicast cluster IP address. This ensures that traffic destined for the cluster is received by all members. The ARP replies sent by a cluster member will therefore indicate that the cluster IP address is reachable via a multicast MAC address.

Some routing devices will not accept such ARP replies. For some routers, adding a static ARP entry for the cluster IP on the routing device will solve the issue. Other routers will not accept this type of static ARP entry. For those cases, use Load Sharing Unicast mode, which does not require the use of multicast for the cluster addresses.

For a list of supported hardware devices, and for their configuration instructions, see SecureKnowledge solution sk10621 (Login at [http://support.checkpoint.com/kb/](http://support.checkpoint.com/kb/) and enter the solution ID in the SecureKnowledge Search by Keyword(s) window).
**IP Address Migration**

If you wish to provide High Availability or Load Sharing to an existing single gateway configuration, it is recommended to take the existing IP addresses from the current gateway, and make these the cluster addresses (cluster virtual addresses), when feasible. Doing so will avoid altering current IPSec endpoint identities, as well keep Hide NAT configurations the same in many cases.

**ClusterXL Hardware Requirements**

In This Section

- **Router Configuration**
- **Switch Configuration**
- **Example Configuration of Cisco Catalyst 6xxx**

The FireWall-1 Gateway is usually located in an environment having other networking devices such as switches and routers. These devices and the Firewall must interact to assure network connectivity. This section outlines the requirements imposed by ClusterXL on surrounding networking equipment.

**Router Configuration**

**Unicast MAC**

When working in High Availability Legacy mode, High Availability New mode and Load Sharing Unicast mode, the Cluster IP is mapped to a regular MAC address, which is the MAC address of the active member. The router learns this MAC through regular ARP messages.

ClusterXL requires this behavior from the router although this is not a specific requirement because most routers now perform this task as a standard feature. If you have a router that does not learn MAC mapping dynamically, you'll have to configure static MAC entries.

**Multicast MAC**

When working in Load Sharing Multicast mode, the router must support sending Unicast IP packets with Multicast MAC addresses. This is required so that all cluster members will receive the data packets. In order to verify that your router supports this configuration simply try adding a MAC entry that follows the format:

```plaintext
some_unicast_mac_address 01:00:5e:xx:xx:xx
```
Some routers can learn the mapping automatically using the ARP mechanism (as for Unicast MAC). You should consult your router documentation for details. See “Hardware Compatibility for Load Sharing Multicast Mode” on page 53 for a list of supported routers and switches.

**Disabling Packets to the Router**

In Multicast Load Sharing configuration packets are transmitted using a Multicast MAC address. Some routers will send this traffic to the router itself. This may result in packet storm through the network. In order to disable it, you should consult your router's documentation.

**Switch Configuration**

**Multicast/Broadcast**

The Cluster Control Protocol (CCP) makes use of Layer 2 multicast, and so it is preferable to use a switch that supports multicast. Most switches support multicast by default. Please check your switch documentation for details.

Using multicast is efficient because a Layer 2 switch connected to non-secured interfaces forwards multicast packets only to the switch ports that connect to cluster members (and that are within a VLAN, if it is a VLAN switch).

ClusterXL uses multicast by default for CCP (except in High Availability Legacy mode, which uses broadcast).

If the connecting switch is incapable of forwarding multicast, it is possible, though less efficient, for the switch to use broadcast to forward traffic, and to configure the cluster members to use broadcast (described in “Choosing the CCP Transport Mode on the Cluster Members” on page 46).

The traffic-storm protection settings on the switch may need to be adjusted. See “Preventing Broadcast Storms on Switches” on page 106.

**IGMP**

ClusterXL does not support IGMP registration. You should disable this feature, also known as IGMP Snooping, in Switches that rely on IGMP packets to configure their ports.

**Port Mirroring**

ClusterXL does not support the use of Unicast MAC addresses with Port Mirroring for Multicast Load Sharing solutions.
Example Configuration of Cisco Catalyst 6xxx

In this example use the following IP address: 10.1.1.100 and the following MAC address: 01:00:5E:0F:11:11.

1 Disable IGMP snooping by running: `no ip igmp snooping` (under `config` for global settings or `config-if` for per-VLAN settings).

2 Configure the Multicast MAC address: `arp 10.1.1.100 01:00:5E:0F:11:11 arpa`. Disable Multicast packets from reaching the Router: `set cam static 01:00:5E:0F:11:11 module/port`. In some routers this is called the port number of the VLAN/Subnet.

Configuring ClusterXL

In This Section

- Configuring Routing for the Client Machines page 44
- Preparing the Cluster Member Machines page 45
- Choosing the CCP Transport Mode on the Cluster Members page 46
- SmartDashboard Configuration page 46

This procedure describes how to configure the Load Sharing Multicast, Load Sharing Unicast, and High Availability New Modes modes from scratch. Their configuration is identical, apart from the mode selection in SmartDashboard Gateway Cluster object. FIGURE 3-2 is used to illustrate the configuration steps.

**Note** - To configure High Availability Legacy Mode, see "High Availability Legacy Mode" on page 137

Configuring Routing for the Client Machines

3 Configure routing so that communication with the networks on the internal side of the cluster is via the cluster IP address on the external side of the cluster. For example, in FIGURE 3-2 on page 47, on the external router, configure a static route such that network 10.10.0.255 is reached via 192.168.10.100.

4 Configure routing so that communication with the networks on the external side of the cluster is via the cluster IP address on the internal side of the cluster. For example, in FIGURE 3-2 on page 47, define 10.10.0.100 as the default gateway on each machine on the internal side of the router.
Preparing the Cluster Member Machines

5 Obtain and install a Central license for ClusterXL on the SmartCenter Server.

6 Define IP addresses for each interfaces on all cluster members. For example, in FIGURE 3-2 on page 47,
   - on Member_A configure the Int Interface with address 10.10.0.1, the Ext interface with address 192.168.10.1, and the SYNC interface with address 10.0.10.1
   - on Member_B configure the Int Interface with address 10.10.0.2, the Ext interface with address 192.168.10.2, and the SYNC interface with address 10.0.10.2

7 For a VPN cluster to properly function, the cluster member clocks must be accurately synchronized to within a second of each other. On cluster members that are constantly up and running it is usually enough to set the time once. More reliable synchronization can be achieved using NTP or some other time synchronization services supplied by the operating system. The cluster member clocks are not relevant for any other (non VPN) cluster capability.

8 Connect the cluster network machines, via the switches. For the Synchronization interfaces, use a cross cable, or a dedicated switch. Make sure that each network (internal, external, Synchronization, DMZ, and so on) is configured on a separate VLAN, switch or hub.

Note - Starting NG with Application Intelligence (R55), it is possible to run synchronization across a WAN. For details, see “Synchronizing Clusters over a Wide Area Network” on page 22.

9 Install VPN-1 Pro on all cluster members.

10 During the configuration phase, enable ClusterXL and State Synchronization by selecting Enable cluster membership for this gateway on Unix machines, or This Gateway is part of a cluster on Windows.

   If you do not make this selection during installation, you can use the Check Point Configuration Tool at any time. Run the cpeconfig utility from the command line, and select the option to turn on cluster capabilities on the module. Note that on some platforms you may be asked to reboot.
Choosing the CCP Transport Mode on the Cluster Members

11 If the connecting switch is incapable of forwarding multicast, it is possible, though less efficient, for the switch to use broadcast to forward traffic.

The ClusterXL Control Protocol (CCP) on the cluster members uses multicast by default, because it is more efficient than broadcast. To toggle the CCP mode between broadcast and multicast, use the following command on each cluster member:

cphaconf set_ccp broadcast/multicast

SmartDashboard Configuration

When configuring a ClusterXL cluster in SmartDashboard, you need to separately configure the topology for both cluster and cluster member. Cluster members are defined within the Gateway Cluster object.

FIGURE 3-2 relates the physical cluster topology to the required SmartDashboard configuration. When configuring the Gateway Cluster object, configure the Virtual IP addresses of the cluster in the Topology page, and the IP addresses of the cluster members in the Cluster Members page, in the Topology tab of every cluster member. Cluster members exist solely inside the Gateway Cluster object. One (or more) interfaces of each cluster member will be the synchronization interface. Synchronization interfaces must be in the Synchronization network that is defined in the Synchronization page of the Gateway Cluster object.
1. Using SmartDashboard, define the Gateway Cluster object. In the **General** tab of the Gateway Cluster object, check ClusterXL as a product installed on the cluster.

2. Define the general IP address of the cluster. Define it to be the same as the IP address of one of the virtual cluster interfaces.

3. In the **Cluster Members** page, click **Add > New Cluster Member** to add cluster members to the cluster. Cluster members exist solely inside the Gateway Cluster object. For each cluster member:
   - In the **Cluster Members Properties > General** tab, define a name a **Name** and **IP Address**. Choose an IP address that is routable from the SmartCenter Server so that the Security Policy installation will be successful. This can be an internal or an external address, or a dedicated management interface.
   - Initialize Secure Internal Communication (SIC).
   - Define the interfaces in the **Topology** tab. To add an interface, click **Add**. In the **Interface Properties** window, check **Cluster interface**. High Availability and Load Sharing are only performed for cluster interfaces. Topology settings for cluster interfaces of cluster members are taken from the topology definitions of the virtual cluster interfaces (specified in the **Topology** tab).
page of the cluster object, in the **Interface Properties** window, **Topology** tab. The *cluster* topology definitions include anti-spoof checking, and whether the interface is internal or external.

- Define the **NAT** and **VPN** tabs, as required.

You can also add an existing gateway as a cluster member by selecting **Add > Add Gateway to Cluster** in the **Cluster Members** page and selecting the gateway from the list in the **Add Gateway to Cluster** window.

If you want to remove a gateway from the cluster, click **Remove** in the **Cluster Members** page and select **Detach Member from Cluster** or right-click on the cluster member in the **Network Objects** tree and select **Detach from Cluster**.

4 In the **ClusterXL** page (FIGURE 3-3), select either

- **Load Sharing**, The Load Sharing mode (**Multicast Mode** or **Unicast Mode**) will depend on the capabilities of the router, OR

- **High Availability New Mode**, and specify the action **Upon Gateway Recovery**. See “What Happens When a Gateway Recovers?” on page 40 for additional information.

**FIGURE 3-3** **ClusterXL** page

5 In the **Synchronization** page, define the synchronization network(s).

- Load Sharing configurations require synchronization between cluster members, and **Use State Synchronization** is checked, and grayed out.
For High Availability New mode, **Use State Synchronization** is checked by default. If you uncheck this, the cluster members will not be synchronized, and existing connections on the failed gateway will be closed when failover occurs.

6. The **Topology** page is used to define the virtual cluster IP addresses, as opposed to the cluster member addresses. In ClusterXL, the cluster topology definitions are used to manipulate ARP packets going to the cluster IP addresses.

   In the **Topology page**, define the cluster topology on each virtual cluster interface. To create a virtual cluster interface, click **Add**.

   - In the **Topology** tab, define the topology for this interface and set up anti-spoofing.
   - In the **Member Networks** tab, define the member network and its netmask if necessary (FIGURE 3-4). Member networks associate the virtual cluster IP addresses in a given direction with the IP addresses of the cluster members in the same direction. You need only define member networks if the subnets on which the physical interfaces of the **cluster members** reside are different than the subnets of the **cluster** interfaces. By default, the member networks are the subnets on which the physical interface of the cluster members reside.

   ![FIGURE 3-4 Topology page and Interface Properties Member Networks tab](image)

7. Define the other pages in the cluster object as required (NAT, VPN, Remote Access, and so on).

Check Point Software and Hardware Compatibility

In This Section

Check Point Software Compatibility (excluding SmartDefense and Web Intelligence) page 50
ClusterXL Compatibility with SmartDefense and Web Intelligence page 52
Operating System Compatibility page 53
Hardware Compatibility for Load Sharing Multicast Mode page 53
Forwarding Layer page 54

Check Point Software Compatibility (excluding SmartDefense and Web Intelligence)

TABLE 3-2 lists the products and features that are either not supported (marked as No), or are only partially supported with ClusterXL (marked as Yes, with a note). It does not apply to their use with OPSEC-certified clustering products.

TABLE 3-2 Products and features that are not fully supported with ClusterXL

<table>
<thead>
<tr>
<th>Product</th>
<th>Feature</th>
<th>Load Sharing</th>
<th>High Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>SmartCenter</td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>Authentication/Security Servers</td>
<td>Yes (1)</td>
<td>Yes (1) (10)</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>ACE servers and SecurID</td>
<td>Yes (8)</td>
<td>Yes (8)</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>Application Intelligence protocol inspection (2)</td>
<td>Yes (3)</td>
<td>Yes</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>Sequence Verifier</td>
<td>Yes (4)</td>
<td>Yes (1)</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>UDP encapsulation</td>
<td>Yes (7)</td>
<td>Yes</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>SAM</td>
<td>Yes (9)</td>
<td>Yes (9)</td>
</tr>
<tr>
<td>VPN-1 Pro</td>
<td>Remote Access VPN with Microsoft IPsec (L2TP)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SecuRemote/SecureClient</td>
<td>Software Distribution Server (SDS)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SecuRemote/SecureClient</td>
<td>IP per user in Office Mode</td>
<td>Yes (11)</td>
<td>Yes (11)</td>
</tr>
<tr>
<td>Performance Pack</td>
<td></td>
<td>Yes (12)</td>
<td>Yes (12)</td>
</tr>
<tr>
<td>FloodGate-1</td>
<td></td>
<td>Yes (4)</td>
<td>Yes (5)</td>
</tr>
</tbody>
</table>
1 Since it requires per-packet state tracking, this feature cannot be guaranteed when a session starts on one cluster member and fails over to another.

2 Application Intelligence protocol inspection includes the general HTTP worm catcher, configuration of Optimized Protocol Enforcement, and Microsoft networks inspection.

3 Application Intelligence protocol inspection is supported when connections maintain unidirectional stickiness. Unidirectional stickiness means that packets in the client-to-server direction are handled by one cluster member, while packets in the server-to-client direction are handled by another cluster member. OPSEC cluster solutions must maintain at least unidirectional stickiness for all connections in order to qualify as OPSEC clusters. Failover can break unidirectional stickiness for certain connections, and in that case, Firewall-1 will proactively reset these connections.

4 Supported when connections maintain bidirectional stickiness. Bidirectional stickiness is the situation where all packets of a connection, regardless of whether they are in the client-to-server direction or the server-to-client direction, are processed by a single cluster member.

5 Supported with bandwidth limits and guarantees that are manually divided between the members. With a 1.5 Mbps connection, and a three-member cluster, each member would have a bandwidth of 500 Kbps, and limits of 1/3 of the total. If a cluster member fails, the total bandwidth will not be automatically re-allocated among the remaining members.

6 Using OPSEC partners platform.

7 Use SecureClient NG FP3 and above.

8 Configuration instructions for ACE server in Cluster environment:

**High Availability**: To support failover scenarios, manually copy the secured file, which is created after the first authentication with the ACE server, from the initiating member to all other members.
Load Sharing:
- Every cluster member should be defined separately on the server with its unique IP address.
- Add the following entry to the `tables.def` file on the SmartCenter Server:
  
  ```
  no_hide_services_ports = {..., <5500, 17> };
  ```
  
  This forces the connection from the cluster members to the ACE server to go out with the member's IP address and not the Cluster address. Make sure the IP addresses of the cluster members are routable from the ACE server box, and then install the Security Policy.
- In some cases the agent libraries (client side) will use the wrong interface IP address in the decryption, and the authentication will fail. To overcome this problem, place a new text file `sdopts.rec` in the same directory as the `dconf.rec` file, with the following line:
  ```
  CLIENT_IP=x.x.x.x
  ```
  where `x.x.x.x` is the primary IP address, as defined on the server. This is the IP address of the interface to which the server is routed.

9 Works as two single gateways. SAM commands executed while a cluster is down are not enforced on this member.

10 In a High Availability configuration, client authentication Wait mode is not reliable. Use other client authentication modes instead.

11 The `ipassignment.conf` file must be copied manually.

12 Performance Pack on Solaris with VLANs is not supported.

ClusterXL Compatibility with SmartDefense and Web Intelligence

The SmartDefense and Web Intelligence features listed in TABLE 3-3 are supported by ClusterXL, with the limitations listed in the notes.

**TABLE 3-3** ClusterXL Compatibility with SmartDefense and Web Intelligence

<table>
<thead>
<tr>
<th>Feature</th>
<th>Load Sharing</th>
<th>High Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment Sanity Check</td>
<td>Yes (1, 3)</td>
<td>Yes (1)</td>
</tr>
<tr>
<td>Pattern Matching</td>
<td>Yes (2, 3)</td>
<td>Yes (2)</td>
</tr>
<tr>
<td>Sequence Verifier</td>
<td>Yes (2, 4)</td>
<td>Yes (2)</td>
</tr>
<tr>
<td>FTP, HTTP and SMTP Security Servers</td>
<td>Yes (2, 5)</td>
<td>Yes (2)</td>
</tr>
<tr>
<td>Web Intelligence</td>
<td>Yes (2)</td>
<td>Yes (2)</td>
</tr>
</tbody>
</table>
Notes:

1. If there is a failover when fragments are being received, the packet will be lost.
2. Does not survive failover.
3. Requires unidirectional stickiness. This means that the same member must receive all external packets, and the same member must receive all internal packets, but the same member does not have to receive both internal and external packets.
4. Requires bidirectional connection stickiness.
5. Uses the forwarding layer, described in “Forwarding Layer” on page 54.

Operating System Compatibility

The operating systems listed in TABLE 3-3 are supported by ClusterXL, with the limitations listed in the notes below. See the Release Notes for the supported versions of these operating systems.

<table>
<thead>
<tr>
<th>Operating System</th>
<th>Load Sharing</th>
<th>High Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows (1)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Solaris (2)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Linux/SecurePlatform (3)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>AIX/HP-UX (4)</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes:

1. VLANs are not supported.
2. VLANs are supported with Sun GigaSwift Ethernet Adapters. VLANs are not supported with Performance Pack.
3. VLANs are supported on all interfaces.
4. Synchronization (used for OPSEC certified clustering products) is not supported.

Hardware Compatibility for Load Sharing Multicast Mode

The following routers and switches are known to be compatible for Load Sharing in Multicast Mode:

Routers
- Cisco 7200 Series
Check Point Software and Hardware Compatibility

- Cisco 1600, 2600, 3600 Series

**Routing Switch**
- Extreme Networks Blackdiamond (Disable IGMP snooping)
- Extreme Networks Alpine 3800 Series (Disable IGMP snooping)
- Foundry Network Bigiron 4000 Series
- Nortel Networks Passport 8600 Series
- Cisco Catalyst 6500 Series (Disable IGMP snooping, Configure Multicast MAC manually)

**Switches**
- Cisco Catalyst 2900, 3500 Series
- Nortel BayStack 450
- Alteon 180e
- Dell PowerConnect 3248 and PowerConnect 5224

**Forwarding Layer**

The Forwarding Layer is a ClusterXL mechanism that allows a cluster member to pass packets to other members, after they have been locally inspected by the firewall. This feature allows connections to be opened from a cluster member to an external host.

Packets originated by cluster members are hidden behind the cluster's virtual IP. Thus, a reply from an external host is sent to the cluster, and not directly to the source member. This can pose problems in the following situations:

- The cluster is working in New High Availability mode, and the connection is opened from the stand-by machine. All packets from the external host are handled by the active machine, instead.
- The cluster is working in a Load Sharing mode, and the decision function has selected another member to handle this connection. This can happen since packets directed at a cluster IP are distributed among cluster members as with any other connection.

If a member decides, upon the completion of the Firewall inspection process, that a packet is intended for another cluster member, it can use the Forwarding Layer to hand the packet over to that destination. This is done by sending the packet over a secured network (any subnet designated as a Synchronization network) directly to that member. It is important to use secured networks only, as encrypted packets are decrypted during the inspection process, and are forwarded as clear-text (unencrypted) data.
Packets sent on the Forwarding Layer use a special source MAC address to inform the receiving member that they have already been inspected by another Firewall module. Thus, the receiving member can safely hand over these packets to the local Operating System, without further inspection. This process is secure, as synchronization networks should always be isolated from any other network (using a dedicated network).
CHAPTER 4

Working with OPSEC Certified Clustering Products

In This Chapter

Introduction to OPSEC Certified Clustering Products page 57
Configuring OPSEC-Certified Clustering Products page 58
Preparing the Cluster Member Machines page 58
SmartDashboard Configuration page 59

Introduction to OPSEC Certified Clustering Products

There are a number of OPSEC certified High Availability (sometimes called as Hot Standby) and Load Sharing (sometimes called Load Balancing) products. These products are used to build highly available VPN-1 Pro Gateway clusters and to distribute traffic evenly among the clustered gateways.

Each OPSEC certified clustering application has its particular strengths and capabilities, whether it be monitoring, management, or performance. The role of these clustering applications is to:

1. Decide which cluster member will deal with each connection.

2. Perform health checks. This involves checking the status of a cluster member (for example, Active, Standby, or Down), and checking the status of the member interfaces.

3. Perform failover.
OPSEC certified clustering products use the VPN-1 Pro state synchronization mechanism (described in Chapter 2, “Synchronizing Connection Information Across the Cluster) to exchange and update connection information and other states between cluster members.

This guide provides general guidelines for working with OPSEC-certified clustering products. Configuration details vary for each clustering product. You are therefore urged to follow the instructions supplied with the OPSEC product.

**Configuring OPSEC-Certified Clustering Products**

This procedure describes how to configure an OPSEC-certified VPN-1 Pro gateway clustering solution.

- Follow the instructions in your clustering product documentation for
  - Preparing the switches.
  - Configuring routing.

**Preparing the Cluster Member Machines**

1. Define IP addresses for all interfaces on all the cluster members.
2. Connect the cluster network machines, via the switches. For the Synchronization interfaces, a cross cable or a dedicated switch is recommended.
3. For Nokia clusters, configure VRRP or IP Clustering before installing VPN-1 Pro. For other OPSEC certified clusters, follow the vendor recommendations. Make sure that after the installation is finished, the option **Enable VPN-1/FW-1 monitoring** is turned to **Enable** in the Nokia configuration manager. This assures that IPSO will monitor changes in FireWall-1 status. For IP Clustering in IPSO 3.7 and above, FireWall-1 state is reported to the Nokia cluster for failover purposes. For VRRP this is done only from IPSO 3.8.

   When configuring VRRP on IPSO 3.7 be sure to configure a cold start delay in the VRRP configuration. The cold start delay is the delay in seconds before a rebooted member can become active after the boot. This delay enables it to synchronize its state with the master before it becomes active (as a member or as a master). Synchronization problems can occur if the value is too small. Consult the Nokia documentation for a recommended value. From IPSO 3.8 this process is done automatically without the need to manually configure a cold start.
Install VPN-1 Pro on all cluster members. During the configuration phase (or later, using the Configuration Tool):

- Install a license for VPN-1 Pro on each cluster member. No special license is required to allow the OPSEC-certified product to work with VPN-1 Pro.
- During the configuration phase, enable State Synchronization by selecting **Enable cluster membership for this gateway** on Unix machines, or **This Gateway is part of a cluster** on Windows.

## SmartDashboard Configuration

### SmartDashboard Configuration for Nokia Clusters

To configure other OPSEC-certified clusters, see “SmartDashboard Configuration for Other OPSEC Clusters” on page 62.

1. Using SmartDashboard, create the Gateway Cluster object.

2. In the **General Properties** page of the Gateway Cluster object, give the cluster a general IP address. For Nokia VRRP, use the external VRRP address. For Nokia IP Clustering, use the external IP Clustering address.

   In the list of **Check Point Products**, ensure ClusterXL is **not** selected.

3. In the **Cluster Members** page, in the **Topology** tab of each cluster member, define the interfaces for the individual members only. In Nokia VRRP and IP clustering, cluster IPs should not be added to the individual member’s **Topology** tab.

   In the **Cluster Members** page, click **Add** to add a cluster members to the cluster.

   Cluster members exist solely inside the Gateway Cluster object. For each cluster member:

   - In the **Cluster Members Properties > General** tab, define a name a **Name** and **IP Address**. Choose an IP address that is routable from the SmartCenter Server so that the Security Policy installation will be successful. This can be an internal or an external address, or a dedicated management interface.

   - Initialize Secure Internal Communication (SIC).

   - Define the interfaces for the individual members in the **Topology** tab. To add an interface, click **Add**. In the **Interface Properties** window, check **Cluster interface**.

   - Define the **NAT** and **VPN** tabs, as required.

4. In the **3rd Party Configuration** page, specify the cluster operating mode, and the OPSEC-certified solution.

   - IP clustering is the Nokia Load Sharing solution.
   - VRRP is the Nokia High Availability solution.
For Nokia IP Clustering check **Support non-sticky connections**. For Nokia VRRP, leave this option unchecked.

Many gateway clusters have a virtual cluster IP address that is defined in **Topology** page of the cluster object, in addition to physical cluster member interface addresses. The use of virtual cluster IP addresses affects the settings in the **3rd Party Configuration** page.

When a cluster member establishes an *outgoing* connection towards the Internet, the source address in the outgoing packets will usually be the physical IP address of the cluster member interface. If virtual cluster IP addresses are used, the clustering product usually changes the source IP address (using NAT) to that of the external virtual IP address of the cluster.

This corresponds to the default setting of **Hide Cluster Members’ outgoing traffic behind the Cluster’s IP address** being checked.

When a client establishes an *incoming* connection to the external virtual address of the cluster, the clustering product changes the destination IP address (using NAT) to that of the physical external address of one of the cluster members.

This corresponds to the default setting of **Forward Cluster’s incoming traffic to Cluster Members’ IP addresses** being checked.

It is recommended that for Nokia IP clustering and VRRP, both these checkboxes should be checked.

The recommended settings for Nokia VRRP and Nokia IP Clustering are shown in FIGURE 4-1.
6 In the Synchronization page, check Use State Synchronization and define the synchronization network(s).

In Nokia clusters, it is possible to “get” the synchronization network from the Nokia cluster configuration if it is already defined. In an already configured Nokia IP Clustering configuration on at least one of the members, clicking the Get... button fetches the primary and secondary synchronization network as defined in the Nokia cluster. In Nokia VRRP it is not possible to “get” the synchronization network.

Note - The Get button works only for synchronization interfaces defined in Voyager. To use Check Point State Synchronization and Nokia synchronization on other interfaces, define them manually.
The **Topology** page is used to define the virtual cluster IP addresses, as opposed to the cluster member addresses.

For Nokia IP clustering and VRRP, configure a virtual IP address in the **Topology** page, just like ClusterXL, as follows:

Define the cluster topology on each virtual cluster interface. To create a virtual cluster interface, click **Add**.

- In the **General** tab, define the cluster virtual IP addresses.
- In the **Topology** tab, define the topology for this interface.
  For Nokia IP clustering, pressing **Get** returns all the cluster IP addresses, which includes the cluster IP of the Nokia synchronization network, which is usually a dummy IP that is not used for any purpose. This IP can usually be safely deleted from the **Topology** tab. For Nokia VRRP, the **Get** option is also available.

Define the other pages in the cluster object as required (**NAT**, **VPN**, **Remote Access**, and so on).

9 Install the Security Policy on the cluster.

### SmartDashboard Configuration for Other OPSEC Clusters

1 Using SmartDashboard, create the Gateway Cluster object.

2 In the **General Properties** page of the Gateway Cluster object, give the cluster a general IP address. In general, make it the external virtual IP address of the cluster. In the list of **Check Point Products**, ensure ClusterXL is **not** selected.

3 In the **Cluster Members** page, in the **Topology** tab of each cluster member, define the interfaces for the individual members. In most OPSEC solutions, cluster IPs should not be added to the individual member’s **Topology** tab. Refer to your clustering product documentation for additional information.

  In the **Cluster Members** page, click **Add** to add a cluster member to the cluster. Cluster members exist solely inside the Gateway Cluster object. For each cluster member:
  - In the **Cluster Members Properties > General** tab, define a name a **Name** and **IP Address**. Choose an IP address that is routable from the SmartCenter Server so that the Security Policy installation will be successful. This can be an internal or an external address, or a dedicated management interface.
  - Initialize Secure Internal Communication (SIC).
  - Define the interfaces for the individual members in the **Topology** tab. To add an interface, click **Add**.
  - Define the **NAT** and **VPN** tabs, as required.
4 In the **3rd Party Configuration** page, specify the cluster operating mode, and for the **3rd Party Solution**, select **OPSEC**.

5 A non-sticky connection is one in which packets from client to server and from server to client pass through different cluster members. Non-sticky connections are a problem because they can lead to out-of-state packets being received by the cluster member. VPN-1 Pro will reject out-of-state packets, even if they belong to a valid connection.

Either the synchronization mechanism, or the OPSEC-certified clustering product need to be able identify valid non-sticky connections, so that VPN-1 Pro will allow those connections through the cluster.

Find out whether or not the OPSEC-certified clustering product can identify valid non-sticky connections.

- If the clustering product cannot identify valid non-sticky connections, the synchronization mechanism can do so instead. In that case, check **Support non-sticky connections**.
- If the clustering product can identify valid non-sticky connections, the synchronization mechanism does not have to take care of this. In that case, uncheck **Support non-sticky connections**. Usually it is safe to uncheck this option in High Availability solutions (not in Load Sharing). Unchecking this option will lead to a slight improvement in the connection establishment rate.

If the **Hide Cluster Members’ outgoing traffic behind the Clusters IP Address** option is checked, **Support non-sticky connections** should also be checked to support outgoing connections from a standby machine (unless specifically directed by OPSEC-certified clustering product guide).
Many gateway clusters have a virtual cluster IP address that is defined in the **Topology** page of the cluster object, in addition to physical cluster member interface addresses. The use of virtual cluster IP addresses affects the settings in the **3rd Party Configuration** page.

When a client behind the cluster establishes an **outgoing** connection towards the Internet, the source address in the outgoing packets, is usually the physical IP address of the cluster member interface. If virtual cluster IP addresses are used, the clustering product usually changes the source IP address (using NAT) to that of the external virtual IP address of the cluster.

This corresponds to the default setting of **Hide Cluster Members’ outgoing traffic behind the Cluster’s IP address** being checked.

When a client establishes an **incoming** connection to the external virtual address of the cluster, the clustering product changes the destination IP address (using NAT) to that of the physical external address of one of the cluster members.

This corresponds to the default setting of **Forward Cluster’s incoming traffic to Cluster Members’ IP addresses** being checked.

In the **Synchronization** page, check **Use State Synchronization** and define the synchronization network(s).

Depending on the OPSEC implementation, it might be possible to “get” the synchronization network from the OPSEC configuration if it is already defined. Refer to the OPSEC documentation to find out if this feature is implemented for a specific OPSEC.

The **Topology** page is used to define the virtual cluster IP addresses, as opposed to the cluster member addresses.

For OPSEC-certified products, the configuration is mandatory in several products, while in others it is forbidden. Refer to your cluster product documentation for details.

Define the other pages in the cluster object as required (**NAT, VPN, Remote Access**, and so on).

Install the Security Policy on the cluster.

**CPHA Command Line Behavior in OPSEC Clusters**

In This Section

- The **cphastart and cphastop Commands in OPSEC Clusters**
- The **cphaprob Command in OPSEC Clusters**

---

64
This section describes the behavior of specific command lines in OPSEC clusters.

**Note** - For details of the cpha command lines see “Monitoring and Troubleshooting Gateway Clusters” on page 67.

### The cphastart and cphastop Commands in OPSEC Clusters

The behavior of the `cphastart` and `cphastop` commands on ClusterXL clusters are described in “The cphastart and cphastop Commands” on page 83.

On OPSEC clusters, the `cphastart` command may not cause the cluster member to start working. On Nokia clusters the behavior is the same as with ClusterXL clusters.

The `cphastop` command may not cause failover on OPSEC clusters. On Nokia IP Clustering clusters (but not on VRRP clusters), the behavior is the same as with ClusterXL clusters.

As with ClusterXL clusters, these commands should only be run by VPN-1 Pro, and not directly by the user.

### The cphaprob Command in OPSEC Clusters

Use the `cphaprob` command to verify that the cluster and the cluster members are working properly. This command is relevant only for Nokia IP clustering and Nokia VRRP. For Nokia VRRP running on IPSO 3.8, these commands have the same effect as in ClusterXL. When VRRP is running on IPSO 3.7, this command will not cause failover from one member to another.

In non-Nokia OPSEC clusters the command output is either empty or the command does not have any effect.

To produce a usage printout for `cphaprob` that shows all the available commands, type `cphaprob` at the command line and press Enter. The meaning of each of these commands is explained in the following sections.

```
cphaprob -d <device> -t <timeout(sec)> -s <ok|init|problem> [-p] register
ncphaprob -f <file> register
cphaprob -d <device> [-p] unregister
ncphaprob -a unregister
ncphaprob -d <device> -s <ok|init|problem> report
cphaprob [-i[a]] [-e] list
ncphaprob state
ncphaprob [-a] if
```
cphaprob state: When running this command the machine state is only Check Point status and is not really a machine status. The command only monitors full sync success, and if a policy was successfully installed. For IP clustering, the state is accurate and also includes the status of the Nokia Cluster. For VRRP, the status is accurate for a firewall, but it does not correctly reflect the status of the Nokia machine (for example, it does not detect interface failure).

cphaprob [-a] if: Shows only the relevant information - interface name, if it is a sync interface or not. “Multicast”/“Broadcast” refers to the cluster control protocol and is relevant only for the sync interface. Note that the status of the interface is not printed since it is not monitored. (This also applies in the Nokia machine.)
In This Chapter

How to Verify the Cluster is Working Properly (cphaprob) page 68
Monitoring Cluster Status using SmartConsole Clients page 76
ClusterXL Configuration Commands (cphaconf, cphastart, cphastop) page 83
How to Initiate Failover page 83
Monitoring Synchronization (fw ctl pstat) page 84
Troubleshooting Synchronization (cphaprob [-reset] synstat) page 88
ClusterXL Error Messages page 99
Solaris Platform Specific Issues: VLAN Switch Port Flapping page 105
Preventing Broadcast Storms on Switches page 106
Member Fails to Start After Reboot page 106
How to Verify the Cluster is Working Properly (cphaprob)

In This Section

The cphaprob Command page 68
Monitoring Cluster Status (cphaprob state) page 69
Monitoring Cluster Interfaces (cphaprob [-a] if) page 71
Monitoring Critical Devices (cphaprob list) page 73
Registering a Critical Device (cphaprob -d ... register) page 74
Registering Critical Devices Listed in a File (cphaprob -f <file> register) page 74
Unregistering a Critical Device (cphaprob -d ... unregister) page 75
Reporting Critical Device Status to ClusterXL (cphaprob -d ... report) page 75
Example cphaprob Script page 75

The cphaprob Command

Use the cphaprob command to verify that the cluster and the cluster members are working properly, and to define critical devices. A critical device is a process running on a cluster member that enables the member to notify other cluster members that it can no longer function as a member. The device reports to the ClusterXL mechanism regarding its current state or it may fail to report, in which case ClusterXL decides that a failover has occurred and another cluster member takes over. When a critical device (also known as a Problem Notification, or pnote) fails, the cluster member is considered to have failed.

There are a number of built-in critical devices, and the administrator can define additional critical devices. The default critical devices are:

- The cluster interfaces on the cluster members.
- Synchronization—full synchronization completed successfully.
- Filter—the Security Policy, and whether it is loaded.
- cphad—which follows the ClusterXL process called cphamcset.
- fwd—the VPN-1 Pro daemon.

These commands can be run automatically by including them in scripts.
To produce a usage printout for `cphaprob` that shows all the available commands, type `cphaprob` at the command line and press Enter. The meaning of each of these commands is explained in the following sections.

```
cphaprob -d <device> -t <timeout(secs)> -s <ok|init|problem> [-p] register
       cphaprob -d <device> -s <ok|init|problem> [-p] unregister
       cphaprob -a unregister
       cphaprob -d <device> [-p] unregister
       cphaprob -a unregister
       cphaprob [-i[a]] [-e] list
       cphaprob state
       cphaprob [-a] if
```

**Monitoring Cluster Status (cphaprob state)**

To see the status of a cluster member, and of all the other members of the cluster, run the following command on the cluster member:

```
cphaprob state
```

Do this after setting up the cluster, and whenever you want to monitor the cluster status. The following is an example of the output of `cphaprob state`:

```
cphaprob state
Cluster mode: Load sharing (Multicast)
Number Unique Address State
1 (local) 30.0.0.1 active
2 30.0.0.2 active
```

- **Cluster mode** can be
  - Load Sharing (Multicast).
  - Load Sharing (Unicast).
  - High Availability New Mode (Primary Up or Active Up).
  - High Availability Legacy Mode (Primary Up or Active Up).
  - For third-party clustering products: “Service”.
    Refer to “Clustering Definitions and Terms” on page 12, for further information.
- The number of the member indicates the member ID for Load Sharing, and the Priority for High Availability.
In Load sharing configuration, all machines in a fully functioning cluster should be Active. In High Availability configurations, only one machine in a properly functioning cluster must be Active, and the others must be in the Standby state. Third-party clustering products show Active/Active even if one of the members is in standby state. This is because this command only reports the status of the full synchronization process. For Nokia VRKP, this command shows the exact state of the Firewall, but not the cluster member (for example, the member may not be working properly but the state of the Firewall is active).

When examining the state of the cluster member, you need to consider whether it is forwarding packets, and whether it has a problem that is stopping it forwarding packets. Each state reflects the result of a test on critical devices. TABLE 5-1 lists and explains the possible cluster states, and whether or not they represent a problem.
## Monitoring Cluster Interfaces (cphaprob [-a] if)

TABLE 5-1  Cluster States

<table>
<thead>
<tr>
<th>State</th>
<th>Meaning</th>
<th>Forwarding packets?</th>
<th>Is this state a Problem?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Everything is OK.</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Active attention</td>
<td>A problem has been detected, but the cluster member is still forwarding packets because it is the only machine in the cluster or there is no other active machines in the cluster. In any other situation the state of the machine would be <em>down</em>.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Down</td>
<td>One of the critical devices is down.</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ready</td>
<td>Can occur in following scenarios:</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>1 When a cluster is upgraded from one version of VPN-1 Pro to another, and the cluster members have different versions of VPN-1 Pro, the members with a new version have the <em>ready</em> state and the members with the previous version have the <em>active</em> state.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 Before a cluster member becomes <em>active</em>, it sends a message to the rest of the cluster, and then expects to receive confirmations from the other cluster members agreeing that it will become active. In the period of time before it receives the confirmations, the machine is in the <em>ready</em> state.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standby</td>
<td>Applies only to a High Availability configuration, and means the member is waiting for an active machine to fail in order to start packet forwarding.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Initializing</td>
<td>An initial and transient state of the cluster member. The cluster member is booting up, and ClusterXL product is already running, but VPN-1 Pro is not yet ready.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Local machine cannot hear anything coming from this cluster member.</td>
<td>Don’t know</td>
<td>Yes</td>
</tr>
</tbody>
</table>

To see the state of the cluster member interfaces and the virtual cluster interfaces, run
the following command on the cluster member:

```
cphaprob [-a] if
```

The output of this command must be identical to the configuration in the cluster object Topology page. For example:

```
cphaprob -a if

Required interfaces: 4
Required secured interfaces: 1

qfe4 UP  (secured, unique, multicast)
qfe5 UP  (non secured, unique, multicast)
qfe6 DOWN (4810.2 secs) (non secured, unique, multicast)
qfe7 UP  (non secured, unique, multicast)

Virtual cluster interfaces: 2
qfe5 30.0.1.130
qfe6 30.0.2.130
```

The interfaces are ClusterXL critical devices. ClusterXL checks the number of good interfaces and sets a value of Required interfaces to the maximum number of good interfaces seen since the last reboot. If the number of good interfaces is less than the Required number, ClusterXL initiates failover. The same for secured interfaces, where only the good synchronization interfaces are counted.

An interface can be:
- Non-secured or Secured. A secured interface is a synchronization interface.
- Shared or unique. A shared interface applies only to High Availability Legacy mode.
- Multicast or broadcast. The Cluster Control Protocol (CCP) mode used in the cluster. CCP can be changed to use broadcast instead. To toggle between these two modes use the command `cphaconf set_ccp <broadcast|multicast>`

For third-party clustering products, except in the case of Nokia IP Clustering, `cphaprob -a if` should always show virtual cluster IP addresses.

When an interface is DOWN, it means that the interface can neither receive or transmit CCP packets. This may happen when an interface is malfunctioning, is connected to an incorrect subnet, is unable to pick up Multicast Ethernet packets and so on. The interface may also be able to receive but not transmit CCP packets, in which case the status field is read. The displayed time is the number of seconds that have elapsed since the interface was last able to receive/transmit a CCP packet.

See “Defining Disconnected Interfaces” on page 122 for additional information.
Monitoring Critical Devices (cphaprob list)

When a critical device fails, the cluster member is considered to have failed. To see the list of critical devices on a cluster member, and of all the other machines in the cluster, run the following command on the cluster member:

```
cphaprob [-i[a]] [-e] list
```

There are a number of built-in critical devices, and the administrator can define additional critical devices. The default critical devices are:

- The cluster interfaces on the cluster members.
- Synchronization—full synchronization completed successfully.
- Filter—the Security Policy, and whether it is loaded.
- cphad—which follows the ClusterXL process called cphamcset.
- fwd—the VPN-1 Pro daemon.

For Nokia IP Clustering, the output is the same as for ClusterXL Load Sharing. For other third-party products, this command produces no output. The following example output shows that the fwd process is down:

```
cphaprob list

Built-in Devices:
Device Name: Interface Active Check
Current state: OK

Registered Devices:
Device Name: Synchronization
Registration number: 0
Timeout: none
Current state: OK
Time since last report: 15998.4 sec

Device Name: Filter
Registration number: 1
Timeout: none
Current state: OK
Time since last report: 15644.4 sec

Device Name: fwd
Registration number: 3
Timeout: 2 sec
Current state: problem
Time since last report: 4.5 sec
```
Registering a Critical Device (cphaprob -d ... register)

It is possible to add a user defined critical device to the default list of critical devices. Use this command to register <device> as a critical process, and add it to the list of devices that must be running for the cluster member to be considered active. If <device> fails, then the cluster member is considered to have failed.

If <device> fails to contact the cluster member in <timeout> seconds, <device> will be considered to have failed. For no timeout, use the value 0.

Define the status of the <device> that will be reported to ClusterXL upon registration. This initial status can be one of:

- **ok** — <device> is alive.
- **init** — <device> is initializing. The machine is down. This state prevents the machine from becoming active.
- **problem** — <device> has failed.

[-p] makes these changes permanent. This means that after performing a reboot or after removing the VPN-1 Pro kernel module (on Linux or IPSO for example) and re-attaching it, the status of critical devices that were registered with this flag will be saved.

Registering Critical Devices Listed in a File (cphaprob -f <file> register)

Register all the user defined critical devices listed in <file>. <file> must be an ASCII file, with each device on a separate line. Each line must list three parameters, which must be separated by at least a space or a tab, as follows:

- **<device> <timeout> <status>**

- **<device>** — The name of the critical device. It must have no more than 15 characters, and must not include white spaces.
- **<timeout>** — If <device> fails to contact the cluster member in <timeout> seconds, <device> will be considered to have failed. For no timeout, use the value 0.
- **<status>** — can be one of
  - **ok** — <device> is alive.
  - **init** — <device> is initializing. The machine is down. This state prevents the machine from becoming active.
Unregistering a Critical Device (cphaprob -d ... unregister)

Unregister a user defined <device> as a critical process. This means that this device is no longer considered critical. If a critical device (and hence a cluster member) was registered as “problem” before running this command, then after running this command the status of the cluster will depend only on the remaining critical devices.

[-p] makes these changes permanent. This means that after performing a reboot or after removing the kernel (on Linux or IPSO for example) and re-attaching it, these critical devices remain unregistered.

Reporting Critical Device Status to ClusterXL (cphaprob -d ... report)

Use this command to report the status of a user defined critical device to ClusterXL. <device> is the device that must be running for the cluster member to be considered active. If <device> fails, then the cluster member is considered to have failed.

The status to be reported. The status can be one of:

ok — <device> is alive
init — <device> is initializing. The machine is down. This state prevents the machine from becoming active.
problem — <device> has failed. If this status is reported to ClusterXL, the cluster member will immediately failover to another cluster member.

If <device> fails to contact the cluster member within the timeout that was defined when the it was registered, <device> and hence the cluster member, will be considered to have failed. This is true only for critical devices with timeouts. If a critical device is registered with the -t 0 parameter, there will be no timeout, and until the device reports otherwise, the status is considered to be the last reported status.

Example cphaprob Script

Predefined cphaprob scripts are located on the location $FWDIR/bin. Two scripts are available

   clusterXL_monitor_ips
   clusterXL_monitor_process
The `clusterXL_monitor_process` script in the Appendix chapter “Example cphaprob Script” on page 143 has been designed to provide a way to check end-to-end connectivity to routers or other network devices and cause failover if the ping fails. The script monitors the existence of given processes and cause failover if the processes die. This script uses the normal `pnote` mechanism.

See “Example cphaprob Script” on page 143.

**Monitoring Cluster Status using SmartConsole Clients**

In This Section

- *SmartView Status* page 76
- *SmartView Tracker* page 79

**SmartView Status**

SmartView Status displays a snapshot of all ClusterXL cluster members in the enterprise, enabling real-time monitoring and alerting. For each cluster member, state change and critical device problem notifications are displayed. You can also use the SmartView Monitor to specify the action to be taken if the status of a cluster member changes. For example, VPN-1 Pro can issue an alert notifying you of suspicious activity.

FIGURE 5-1 is an example of the ClusterXL module details in the SmartView Status window.
FIGURE 5-1 ClusterXL Module Details in the SmartView Status Window

The ClusterXL module details in the SmartView Status are explained in TABLE 5-2.
To collapse or expand ClusterXL module data, click the toolbar icon, or select ClusterXL from the Products menu.

**TABLE 5-2  ClusterXL Module Details in SmartView Status**

<table>
<thead>
<tr>
<th>Column</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>The status of ClusterXL modules on this object:</td>
</tr>
<tr>
<td></td>
<td><strong>Note</strong> - For ClusterXL the status reflects the state of the cluster member, while for OPSEC clusters only the current state of the Synchronization Mechanism is displayed.</td>
</tr>
<tr>
<td></td>
<td><strong>OK</strong> — The module installed and active.</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="OK" /></td>
</tr>
<tr>
<td></td>
<td><strong>Attention</strong> — The module is active even though all the members of the cluster have some problem. Despite this, the gateway with the least problems and the next highest priority level is active and working as a backup until the highest priority level gateway can be restored.</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Attention" /></td>
</tr>
<tr>
<td>Description</td>
<td>Text representing the status of the ClusterXL module installed on this object: OK, Attention, Problem. See above for meaning.</td>
</tr>
<tr>
<td></td>
<td><strong>Note</strong> - When using OPSEC clusters the Status Description does not show the state of the cluster member.</td>
</tr>
<tr>
<td>Working Mode</td>
<td>The cluster member’s working mode as follows:</td>
</tr>
<tr>
<td></td>
<td>• High Availability (Active Up)</td>
</tr>
<tr>
<td></td>
<td>• High Availability (Primary Up)</td>
</tr>
<tr>
<td></td>
<td>• Load Sharing (Unicast)</td>
</tr>
<tr>
<td></td>
<td>• Load Sharing (Multicast)</td>
</tr>
<tr>
<td>Started</td>
<td>“Yes” if the module is running. “No” if the module is not running.</td>
</tr>
<tr>
<td>Running Mode</td>
<td>Describes the current mode of the cluster member: “Active”, “Standby” or “Down”.</td>
</tr>
<tr>
<td></td>
<td><strong>Note</strong> - In Load Sharing mode there is no “Stand By” mode.</td>
</tr>
</tbody>
</table>
Starting and Stopping ClusterXL Using SmartView Status

To stop ClusterXL on the machine and cause failover to another machine, open SmartView Status, click the cluster object, select one of the member gateway branches, right click a cluster member, and select **Down**.

To initiate a restart of ClusterXL, open SmartView Status, click the cluster object, select one of the member gateway branches, right click a cluster member, and select **Up**.

**Note** - SmartView Status does not initiate full synchronization, so that some connections may be lost. To initiate full synchronization, perform `cpstart`, or start the cluster member using the `cphaprob` command.

### SmartView Tracker

Every change in status of a cluster member is recorded in SmartView Tracker according to the choice in the **Fail-Over Tracking** option of the cluster object **ClusterXL** page.

### ClusterXL Log Messages

The following conventions are used in this section:

1. Square brackets are used to indicate place holders, which are substituted by relevant data when an actual log message is issued (for example, `[NUMBER]` will be replaced by a numeric value).

2. Angle brackets are used to indicate alternatives, one of which will be used in actual log messages. The different alternatives are separated with a vertical line (for example, `<up|down>` indicates that either “up” or “down” will be used).

3. The following place holders are frequently used:

<table>
<thead>
<tr>
<th>Column</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Priority</strong></td>
<td>The index of the cluster member.</td>
</tr>
<tr>
<td><strong>Interfaces</strong></td>
<td>Details of the open interfaces including the IP Address, Status (Up or Down) and Sync mode (State or No). In Legacy High Availability an additional field called Shared (Yes or No) is displayed. <strong>Note</strong> - This information is not shown for OPSEC cluster.</td>
</tr>
<tr>
<td><strong>Problem</strong></td>
<td>ClusterXL Module details, including Synchronization details. Filter details and cphad information.</td>
</tr>
</tbody>
</table>

**TABLE 5-2** ClusterXL Module Details in SmartView Status
- ID: A unique cluster member identifier, starting from “1”. This corresponds to the order in which members are sorted in the cluster object's GUI.
- IP: Any unique IP address that belongs to the member.
- MODE: The cluster mode (for example, New HA, LS Multicast, and so on).
- STATE: The state of the member (for example, active, down, standby).
- DEVICE: The name of a pnote device (for example, fwd, Interface Active Check).

**General logs**

*Starting <ClusterXL|State Synchronization>.*
Indicates that ClusterXL (or State Synchronization, for 3rd party clusters) was successfully started on the reporting member. This message is usually issued after a member boots, or after an explicit call to `cphastart`.

*Stopping <ClusterXL|State Synchronization>.*
Informs that ClusterXL (or State Synchronization) was deactivated on this machine. The machine will no longer be a part of the cluster (even if configured to be so), until ClusterXL is restarted.

*Unconfigured cluster Machines changed their MAC Addresses. Please reboot the cluster so that the changes take affect.*
This message is usually issued when a machine is shut down, or after an explicit call to `cphastop`.

**State logs**

*Mode inconsistency detected: member [ID] ([IP]) will change its mode to [MODE]. Please re-install the security policy on the cluster.*
This message should rarely happen. It indicates that another cluster member has reported a different cluster mode than is known to the local member. This is usually the result of a failure to install the security policy on all cluster members. To correct this problem, install the Security Policy again.

**Note** - The cluster will continue to operate after a mode inconsistency has been detected, by altering the mode of the reporting machine to match the other cluster members. However, it is highly recommended that the policy will be re-installed as soon as possible.

*State change of member [ID] ([IP]) from [STATE] to [STATE] was cancelled, since all other members are down. Member remains [STATE].*
When a member needs to change its state (for example, when an active member encounters a problem and needs to bring itself down), it first queries the other members for their state. If all other members are down, this member cannot change its
state to a non-active one (or else all members will be down, and the cluster will not function). Thus, the reporting member continues to function, despite its problem (and will usually report its state as “active attention”).

This message is issued whenever a cluster member changes its state. The log text specifies the new state of the member.

Pnote logs

PNote log messages are issued when a pnote device changes its state.

• [DEVICE] on member [ID] ([IP]) status OK ([REASON]).
  The pnote device is working normally.

• [DEVICE] on member [ID] ([IP]) detected a problem ([REASON]).
  Either an error was detected by the pnote device, or the device has not reported its state for a number of seconds (as set by the “timeout” option of the pnote)

• [DEVICE] on member [ID] ([IP]) is initializing ([REASON]).
  Indicates that the device has registered itself with the pnote mechanism, but has not yet determined its state.

• [DEVICE] on member [ID] ([IP]) is in an unknown state ([STATE ID]) ([REASON]).
  This message should not normally appear. Contact Check Point Support.

Interface logs

• interface [INTERFACE NAME] of member [ID] ([IP]) is up.
  Indicates that this interface is working normally, meaning that it is able to receive and transmit packets on the expected subnet.

• interface [INTERFACE NAME] of member [ID] ([IP]) is down (receive <up|down>, transmit <up|down>).
  This message is issued whenever an interface encounters a problem, either in receiving or transmitting packets. Note that in this case the interface may still be working properly, as far as the OS is concerned, but is unable to communicate with other cluster members due to a faulty cluster configuration.

• interface [INTERFACE NAME] of member [ID] ([IP]) was added.
  Notifies the user that a new interface was registered with FireWall-1 (meaning that packets arriving on this interface are filtered by FireWall-1). Usually this message is the result of activating an interface (such as issuing an ifconfig up command on Unix systems). The interface will now be included in the ClusterXL reports (such as in SmartView Status, or in the output of cphaprobs -a if). Note that the interface may still be reported as “Disconnected”, in case it was so configured for ClusterXL.

• interface [INTERFACE NAME] of member [ID] ([IP]) was removed.
Indicates that an interface was detached from FireWall-1, and is therefore no longer monitored by ClusterXL.

**SecureXL logs**

- SecureXL device was deactivated since it does not support CPLS. This message is the result of an attempt to configure a ClusterXL in Load Sharing Multicast mode over Firewall-1 modules using an acceleration device that does not support Load Sharing. As a result, acceleration will be turned off, but the cluster will work in Check Point Load Sharing mode (CPLS).

**Reason Strings**

- member [ID] ([IP]) reports more interfaces up. This text can be included in a pnote log message describing the reasons for a problem report: Another member has more interfaces reported to be working, than the local member does. This means that the local member has a faulty interface, and that its counterpart can do a better job as a cluster member. The local member will therefore go down, leaving the member specified in the message to handle traffic.

- member [ID] ([IP]) has more interfaces - check your disconnected interfaces configuration in the <discntd.if file|registry>. This message is issued when members in the same cluster have a different number of interfaces. A member having less interfaces than the maximal number in the cluster (the reporting member) may not be working properly, as it is missing an interface required to operate against a cluster IP address, or a synchronization network. If some of the interfaces on the other cluster member are redundant, and should not be monitored by ClusterXL, they should be explicitly designated as “Disconnected”. This is done using the file $FWDIR/conf/dicntd.if (under Unix systems), or the Windows Registry.

- [NUMBER] interfaces required, only [NUMBER] up. ClusterXL has detected a problem with one or more of the monitored interfaces. This does not necessarily mean that the member will go down, as the other members may have less operational interfaces. In such a condition, the member with the highest number of operational interfaces will remain up, while the others will go down.
ClusterXL Configuration Commands (cphaconf, cphastart, cphastop)

The cphaconf Command

Running this command is not recommended. It should be run only by Firewall-1.

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cphaconf [-i &lt;machine id&gt;] [-p &lt;policy id&gt;] [-b &lt;db_id&gt;] [-n &lt;cluster num&gt;] [-c &lt;cluster size&gt;] [-m &lt;service &gt;] [-t &lt;secured IF 1&gt;... ] start</td>
<td></td>
</tr>
<tr>
<td>cphaconf [-t &lt;secured IF 1&gt;... ] [-d &lt;disconnected IF 1&gt;... ] add</td>
<td></td>
</tr>
<tr>
<td>cphaconf clear-secured</td>
<td></td>
</tr>
<tr>
<td>cphaconf clear-disconnected</td>
<td></td>
</tr>
<tr>
<td>cphaconf stop</td>
<td></td>
</tr>
<tr>
<td>cphaconf init</td>
<td></td>
</tr>
<tr>
<td>cphaconf forward &lt;on/off&gt;</td>
<td></td>
</tr>
<tr>
<td>cphaconf debug &lt;on/off&gt;</td>
<td></td>
</tr>
<tr>
<td>cphaconf set_ccp &lt;broadcast/multicast&gt;</td>
<td></td>
</tr>
<tr>
<td>cphaconf mc_reload</td>
<td></td>
</tr>
<tr>
<td>cphaconf debug_data</td>
<td></td>
</tr>
</tbody>
</table>

The cphastart and cphastop Commands

Running cphastart on a cluster member activates ClusterXL on the member. It does not initiate full synchronization. cphastart is the recommended way to start a cluster member.

Running cphastop on a cluster member stops the cluster member from passing traffic. State synchronization also stops. It is still possible to open connections directly to the cluster member. In High Availability Legacy mode, running cphastop may cause the entire cluster to stop functioning.

These commands should only be run by Firewall-1, and not directly by the user.

How to Initiate Failover

In This Section

Stopping the Cluster Member  page 84
Starting the Cluster Member  page 84
The state of a cluster member can be manually controlled in order to take down the cluster member. This initiates failover to the other cluster member(s), in the case of Load Sharing, or failover to the next highest priority cluster member in the case of High Availability.

**Stopping the Cluster Member**

To stop ClusterXL on the machine and cause failover to another machine, do one of the following:

- Register a dummy critical device (`faildevice` for example) using the command `cphaprob -d faildevice ok register`, and then run the following command to report to ClusterXL that the critical device `faildevice` has a problem. Failover will immediately occur to another cluster member:
  
  - `cphaprob -d faildevice -s problem report`

- Open SmartView Status, click the cluster object, select one of the member gateway branches, right click a cluster member, and select **Down**.

**Starting the Cluster Member**

ClusterXL starts automatically when VPN-1 Pro is started on the cluster member (`cpstart`). To initiate a restart of ClusterXL, do one of the following:

- To reactivate a cluster member that was downed using the command `cphaprob -d faildevice -s problem report`, and to initiate full synchronization, run either of the following commands:
  
  - `cphaprob -d faildevice -s ok report`
  - `cphaprob -d faildevice unregister`

- Open SmartView Status, click the cluster object, select one of the member gateway branches, right click a cluster member, and select **Up**.

**Note** - Starting the Cluster member from SmartView Status does not initiate full synchronization, so that some connections may be lost. To initiate full synchronization, perform `cpstart`, or start the cluster member using the `cphaprob` command.

**Monitoring Synchronization (fw ctl pstat)**

To monitor the synchronization mechanism on ClusterXL or third-party OPSEC certified clustering products, run the following command on a cluster member:

```
fw ctl pstat
```
The output of this command is a long list of statistics for the VPN-1 Pro Gateway. At the end of the list there is a section called “Synchronization” that applies per Gateway Cluster member. Many of the statistics are counters that can only increase. A typical output is as follows:

<table>
<thead>
<tr>
<th>Version: new</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status: Able to Send/Receive sync packets</td>
</tr>
<tr>
<td>Sync packets sent:</td>
</tr>
<tr>
<td>total: 3976, retransmitted: 0, retrans reqs: 58, acks: 97</td>
</tr>
<tr>
<td>Sync packets received:</td>
</tr>
<tr>
<td>total: 4290, were queued: 58, dropped by net: 47</td>
</tr>
<tr>
<td>retrans reqs: 0, received 0 acks</td>
</tr>
<tr>
<td>retrans reqs for illegal seq: 0</td>
</tr>
<tr>
<td>Callback statistics: handled 3 cb, average delay: 1, max delay: 2</td>
</tr>
<tr>
<td>Delta Sync memory usage: currently using XX KB mem</td>
</tr>
<tr>
<td>Callback statistics: handled 322 cb, average delay: 2, max delay: 8</td>
</tr>
<tr>
<td>Number of Pending packets currently held: 1</td>
</tr>
<tr>
<td>Packets released due to timeout: 18</td>
</tr>
</tbody>
</table>

The meaning of each line in this printout is explained below.

This line must appear if synchronization is configured. It indicates that new sync is working (as opposed to old sync from version 4.1).

| Status: Able to Send/Receive sync packets |

If sync is unable to either send or receive packets, there is a problem. Sync may be temporarily unable to send or receive packets during boot, but this should not happen during normal operation. When performing full sync, sync packet reception may be interrupted.

**Sync packets sent:**

| total: 3976, retransmitted: 0, retrans reqs: 58, acks: 97 |

The total number of sync packets sent is shown. Note that the total number of sync packets is non-zero and increasing.

The cluster member sends a retransmission request when a sync packet is received out of order. This number may increase when under load.
Acks are the acknowledgements sent for received sync packets, when an acknowledgement was requested by another cluster member.

Sync packets received:
   total: 4290, were queued: 58, dropped by net: 47

The total number of sync packets received is shown. The queued packets figure increases when a sync packet is received that complies with one of the following conditions:

1. The sync packet is received with a sequence number that does not follow the previously processed sync packet.
2. The sync packet is fragmented. This is done to solve MTU restrictions.

This figure never decreases. A non-zero value does not indicate a problem.

The dropped by net number may indicate network congestion. This number may increase slowly under load. If this number increases too fast, a networking error may interfere with the sync protocol. In that case, check the network.

retrans reqs: 0, received 0 acks
retrans reqs for illegal seq: 0
Callback statistics: handled 3 cb, average delay: 1, max delay: 2

This message refers to the number of received retransmission requests, in contrast to the transmitted retransmission requests in the section above. When this number grows very fast, it may indicate that the load on the machine is becoming too high for sync to handle.

Acks refer to the number of acknowledgements received for the “cb request” sync packets, which are sync packets with requests for acknowledgments.

Retrans reqs for illegal seq displays the number of retransmission requests for packets which are no longer in this member's possession. This may indicate a sync problem.

Callback statistics relate to received packets that involve Flush and Ack. This statistic only appears for a non-zero value.

The callback average delay is how much the packet was delayed in this member until it was released when the member received an ACK from all the other members. The delay happens because packets are held until all other cluster members have acknowledged reception of that sync packet.
This figure is measured in terms of numbers of packets. Normally this number should be small (~1-5). Larger numbers may indicate an overload of sync traffic, which causes connections that require sync acknowledgements to suffer slight latency.

| dropped updates as a result of sync overload: 0 |

In a heavily loaded system, the cluster member may drop synchronization updates sent from another cluster member.

| Delta Sync memory usage: currently using XX KB mem |

Delta Sync memory usage only appears for a non-zero value. Delta sync requires memory only while full sync is occurring. Full sync happens when the system goes up—after reboot for example. At other times, Delta sync requires no memory because Delta sync updates are applied immediately. For information about Delta sync see “How State Synchronization Works” on page 17.

| Number of Pending packets currently held: 1 |
| Packets released due to timeout: 18 |

Number of Pending packets currently held only appears for a non-zero value. ClusterXL prevents out-of-state packets in non-sticky connections. It does this by holding packets until a SYN-ACK is received from all other active cluster members. If for some reason a SYN-ACK is not received, FireWall-1 on the cluster member will not release the packet, and the connection will not be established.

Packets released due to timeout only appears for a non-zero value. If the Number of Pending Packets is large (more than 100 pending packets), and the number of Packets released due to timeout is small, you should take action to reduce the number of pending packets. To tackle this problem, see “Reducing the Number of Pending Packets” on page 121.
Troubleshooting Synchronization (cphaprob [-reset] syncstat)

Introduction to cphaprob [-reset] syncstat

Heavily loaded clusters and clusters with geographically separated members pose special challenges. High connection rates, and large distances between the members can lead to delays that affect the operation of the cluster.

The cphaprob [-reset] syncstat command is a tool for monitoring the operation of the State Synchronization mechanism in highly loaded and distributed clusters. It can be used for both ClusterXL and third-party OPSEC certified clustering products.

The troubleshooting process is as follows:

1. Run the cphaprob syncstat command.
2. Examine and understand the output statistics.
3. Tune the relevant synchronization global configuration parameters.
4. Rerun the command, resetting the statistics counters using the -reset option:
   cphaprob -reset syncstat
5. Examine the output statistics to see if the problem is solved.

The section “Output of the cphaprob [-reset] syncstat command” on page 89 explains each of the output parameters, and also explains when the output represents a problem.

Any identified problem can be solved by performing one or more of tips described in “Synchronization Troubleshooting Options” on page 97.
Output of the cphaprob [-reset] syncstat command

The output parameters of the cphaprob syncstat command are shown in TABLE 5-3. The values (not shown) give an insight into the state and characteristics of the synchronization network. Each parameter and the meaning of its possible values is explained in the following sections.

TABLE 5-3  cphaprob syncstat command output parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sync Statistics (IDs of F&amp;A Peers - 1)</td>
<td>90</td>
</tr>
<tr>
<td>Other Member Updates</td>
<td>90</td>
</tr>
<tr>
<td>Sent retransmission requests</td>
<td>90</td>
</tr>
<tr>
<td>Avg missing updates per request</td>
<td>90</td>
</tr>
<tr>
<td>Old or too-new arriving updates</td>
<td>90</td>
</tr>
<tr>
<td>Unsynced missing updates</td>
<td>91</td>
</tr>
<tr>
<td>Lost sync connection (num of events)</td>
<td>91</td>
</tr>
<tr>
<td>Timed out sync connection</td>
<td>91</td>
</tr>
<tr>
<td>Local Updates</td>
<td>92</td>
</tr>
<tr>
<td>Total generated updates</td>
<td>92</td>
</tr>
<tr>
<td>Recv Retransmission requests</td>
<td>92</td>
</tr>
<tr>
<td>Recv Duplicate Retrans request</td>
<td>92</td>
</tr>
<tr>
<td>Blocking Scenarios</td>
<td>92</td>
</tr>
<tr>
<td>Blocked packets</td>
<td>93</td>
</tr>
<tr>
<td>Max length of sending queue</td>
<td>94</td>
</tr>
<tr>
<td>Avg length of sending queue</td>
<td>94</td>
</tr>
<tr>
<td>Hold Pkts events</td>
<td>95</td>
</tr>
<tr>
<td>Unhold Pkt events</td>
<td>95</td>
</tr>
<tr>
<td>Not held due to no members</td>
<td>95</td>
</tr>
<tr>
<td>Max held duration (ticks)</td>
<td>95</td>
</tr>
<tr>
<td>Avg held duration (ticks)</td>
<td>96</td>
</tr>
<tr>
<td>Timers:</td>
<td>96</td>
</tr>
<tr>
<td>Sync tick (ms)</td>
<td>96</td>
</tr>
<tr>
<td>CPHA tick (ms)</td>
<td>96</td>
</tr>
<tr>
<td>Queues:</td>
<td>96</td>
</tr>
<tr>
<td>Sending queue size</td>
<td>97</td>
</tr>
<tr>
<td>Receiving queue size</td>
<td>97</td>
</tr>
</tbody>
</table>
Sync Statistics (IDs of F&A Peers - 1):
These statistics relate to the state synchronization mechanism. The F&A (Flush and Ack) peers are the cluster members that this member recognizes as being part of the cluster. The IDs correspond to IDs and IP addresses generated by the `cphaprobs` state command.

Other Member Updates:
The statistics in this section relate to updates generated by other cluster members, or to updates that were not received from the other members. Updates inform about changes in the connections handled by the cluster member, and are sent from and to members. Updates are identified by sequence numbers.

Sent retransmission requests
The number of retransmission requests, which were sent by this member. Retransmission requests are sent when certain packets (with a specified sequence number) are missing, while the sending member already received updates with advanced sequences.

A high value can imply connectivity problems.

*Tip* - Compare the number of retransmission requests to the Total Regenerated Updates of the other members (see “Total generated updates” on page 92).

If its value is unreasonably high (more than 30% of the Total Generated Updates of other members), contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration.

Avg missing updates per request
Each retransmission request can contain up to 32 missing consecutive sequences. The value of this field is the average number of requested sequences per retransmission request.

More than 20 missing consecutive sequences per retransmission request can imply connectivity problems.

*Tip* - If this value is unreasonably high, contact Technical Support, equipped with the entire output and a detailed description of the network topology and configuration.

Old or too-new arriving updates
The number of arriving sync updates where the sequence number is too low, which implies it belongs to an old transmission, or too high, to the extent that it cannot belong to a new transmission.
Large values imply connectivity problems.

**Tip** - See "Enlarging the Receiving Queue" on page 97. If this value is unreasonably high (more than 10% of the total updates sent), contact Technical Support, equipped with the entire output and a detailed description of the network topology and configuration.

**Unsynced missing updates**

The number of missing sync updates for which the receiving member stopped waiting. It stops waiting when the difference in sequence numbers between the newly arriving updates and the missing updates is larger than the length of the receiving queue.

This value should be zero. However, the loss of some updates is acceptable as long as the number of lost updates is less than 1% of the total generated updates.

**Tip** - To decrease the number of lost updates, expand the capacity of the Receiving Queue. See "Enlarging the Receiving Queue" on page 97.

**Lost sync connection (num of events)**

The number of events in which synchronization with another member was lost and regained due to either Security Policy installation on the other member, or a large difference between the expected and received sequence number.

The value should be zero. A positive value indicates connectivity problems.

**Tip** - Allow the sync mechanism to handle large differences in sequence numbers by expanding the Receiving Queue capacity. See "Enlarging the Receiving Queue" on page 97.

**Timed out sync connection**

The number of events in which the member declares another member as not connected. The member is considered as disconnected because no ACK packets were received from that member for a period of time (one second), even though there are Flush and Ack packets being held for that member.

The value should be zero. Even with a round trip time on the sync network as high as 100ms, one second should be enough time to receive an ACK. A positive value indicates connectivity problems.

**Tip** - Try enlarging the Sync Timer (see “Enlarging the Sync Timer” on page 98). However, you may well have to contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration.
Local Updates:
The statistics in this section relate to updates generated by the local cluster member. Updates inform about changes in the connections handled by the cluster member, and are sent from and to members. Updates are identified by sequence numbers.

Total generated updates
The number of sync update packets generated by the sync mechanism since the statistics were last reset. Its value is the same as the difference between the sequence number when applying the \(-\text{reset}\) option, and the current sequence number.

Can have any value.

Recv Retransmission requests
The number of received retransmission requests. A member requests retransmissions when it is missing specified packets with lower sequence numbers than the ones already received.

A large value can imply connectivity problems.

Tip - If this value is unreasonably high (more than 30% of the “Total generated updates” on page 92) contact Technical Support, equipped with the entire output and a detailed description of the network topology and configuration.

Recv Duplicate Retrans request
The number of duplicated retransmission requests received by the member. Duplicate requests were already handled, and so are dropped.

A large value may indicate network problem or storms on the sync network.

Tip - If this value is unreasonably high (more than 30% of the “Total generated updates” on page 92) contact Technical Support, equipped with the entire output and a detailed description of the network topology and configuration.

Blocking Scenarios
Under extremely heavy load conditions, the cluster blocks new connections. This parameter shows the number of times that the cluster member started blocking new connections due to sync overload.

The member starts to block connections when its Sending Queue has reached its capacity threshold. The capacity threshold is calculated as 80% of the difference between the current sequence number and the sequence number for which the member received an ACK from all the other operating members.
A positive value indicates heavy load. In this case, observe the “Blocked packets” on page 93 to see how many packets we blocked. Each dropped packet means one blocked connection.

This parameters is only measured if the Block New Connections mechanism (described in “Blocking New Connections Under Load” on page 119) is active. To activate the Block New Connections mechanism, apply the following command on all the cluster members:

```bash
fw ctl set int fw_sync_block_new_conns 0
```

**Tip** - The best way to handle a severe blocking connections problem is to enlarge the sending queue. See “Enlarging the Sending Queue” on page 97.

Another possibility is to decrease the timeout after which a member initiates an ACK. See “Reconfiguring the Acknowledgment Timeout” on page 99. This updates the sending queue capacity more accurately, thus making the blocking process more precise.

### Blocked packets

The number of packets that were blocked because the cluster member was blocking all new connections (see “Blocking Scenarios” on page 92). The number of blocked packets is usually one packet per new connection attempt.

A value higher than 5% of the Sending Queue see “Avg length of sending queue” on page 94) can imply a connectivity problem, or that ACKs are not being sent frequently enough.

This parameters is only measured if the Block New Connections mechanism (described in “Blocking New Connections Under Load” on page 119) is active. To activate the Block New Connections mechanism, apply the following command on all the cluster members:

```bash
fw ctl set int fw_sync_block_new_conns 0
```

**Tip** - The best way to handle a severe blocking connections problem is to enlarge the sending queue. See “Enlarging the Sending Queue” on page 97.

Another possibility is to decrease the timeout after which a member initiates an ACK. See “Reconfiguring the Acknowledgment Timeout” on page 99. This updates the sending queue capacity more accurately, thus making the blocking process more precise.
Max length of sending queue

The size of the Sending Queue is fixed. By default it is 512 sync updates. As newer updates with higher sequence numbers enter the queue, older updates with lower sequence numbers drop off the end of the queue. An older update could be dropped from the queue before the member receives an ACK about that update from all the other members.

This parameter is the difference between the current sync sequence number and the last sequence number for which the member received an ACK from all the other members. The value of this parameter can therefore be greater than 512.

The value of this parameter should be less than 512. If larger than 512, there is not necessarily a sync problem. However, the member will be unable to answer retransmission request for updates which are no longer in its queue.

This parameter is only measured if the Block New Connections mechanism (described in “Blocking New Connections Under Load” on page 119) is active. To activate the Block New Connections mechanism, apply the following command on all the cluster members:

```
fw ctl set int fw_sync_block_new_conns 0
```

**Tip** - Enlarge the Sending Queue to value larger than this value. See “Enlarging the Sending Queue” on page 97.

Avg length of sending queue

The average value of the Max length of sending queue parameter, since reboot or since the Sync statistics were reset.

The value should be up to 80% of the size of the Sending Queue.

This parameter is only measured if the Block New Connections mechanism (described in “Blocking New Connections Under Load” on page 119) is active. To activate the Block New Connections mechanism, apply the following command on all the cluster members:

```
fw ctl set int fw_sync_block_new_conns 0
```

**Tip** - Enlarge the Sending Queue so that this value is not larger than 80% of the new queue size. See “Enlarging the Sending Queue” on page 97.
Hold Pkts events
The number of occasions where the sync update required Flush and Ack, and so was kept within the system until an ACK arrived from all the other functioning members. Should be the same as the number of Unhold Pkt events.

Tip · Contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration.

Unhold Pkt events
The number of occasions when the member received all the required ACKS from the other functioning members. Should be the same as the number of Hold Pkts events.

Tip · Contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration.

Not held due to no members
The number of packets which should have been held within the system, but were released because there were no other operating members.

When the cluster has at least two live members, the value should be 0.

Tip · The cluster has a connectivity problem. Examine the values of the parameters: “Lost sync connection (num of events)” on page 91 and “Timed out sync connection” on page 91 to find out why the member thinks that it is the only cluster member.

You may also need to contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration.

Max held duration (ticks)
The maximum time in ticks (one tick equals 100ms) for which a held packet was delayed in the system for Flush and Ack purposes.
Troubleshooting Synchronization (cphaprob [-reset] syncstat)

It should not be higher than 50 (5 seconds), because of the pending timeout mechanism which releases held packets after a certain timeout. By default, the release timeout is 50 ticks. A high value indicates connectivity problem between the members.

**Tip** - Optionally change the default timeout by changing the value of the fwldbcast_pending_timeout global variable. See “How to Configure Module Configuration Parameters” on page 117 and “Reducing the Number of Pending Packets” on page 121.

Also, examine the parameter “Timed out sync connection” on page 91 to understand why packets were held for a long time.

You may also need to contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration.

**Avg held duration (ticks)**

The average duration in ticks (tick equals 100ms) that held packets were delayed within the system for Flush and Ack purposes.

The average duration should be about the round-trip time of the sync network. A larger value indicates connectivity problem.

**Tip** - If the value is high, contact Technical Support equipped with the entire output and a detailed description of the network topology and configuration in order to examine the cause to the problem.

**Timers:**

The Sync and CPHA timers perform sync and cluster related actions every fixed interval.

**Sync tick (ms)**

The *Sync timer* performs cluster related actions every fixed interval. By default, the Sync timer interval is 100ms. The base time unit is 100ms (or 1 tick), which is also the minimum value.

**CPHA tick (ms)**

The *CPHA timer* performs cluster related actions every fixed interval. By default, the CPHA timer interval is 100ms. The base time unit is 100ms (or 1 tick), which is also the minimum value.

**Queues:**

Each cluster member has two queues. The Sending Queue and the Receiving Queue.
Sending queue size

The Sending Queue on the cluster member stores locally generated sync updates. Updates in the Sending Queue are replaced by more recent updates. In a highly loaded cluster, updates are therefore kept for less time. If a member is asked to retransmit an update, it can only do so if the update is still in its Sending Queue. The default (and minimum) size of this queue is 512. Each member has one sending queue.

Receiving queue size

The Receiving Queue on the cluster member keeps the updates from each cluster member until it has received a complete sequence of updates. The default (and minimum) size of this queue is 128. Each member keeps a Receiving Queue for each of the peer members.

Synchronization Troubleshooting Options

The following options specify the available troubleshooting options. Each option involves editing a global system configurable parameter to reconfigure the system with different value than the default.

Enlarging the Sending Queue

The Sending Queue on the cluster member stores locally generated sync updates. Updates in the Sending Queue are replaced by more recent updates. In a highly loaded cluster, updates are therefore kept for less time. If a member is asked to retransmit an update, it can only do so if the update is still in its Sending Queue. The default (and minimum) size of this queue is 512. Each member has one sending queue.

To enlarge the sending queue size, change the value of the global parameter `sending_queue_size`. See “How to Configure Module Configuration Parameters” on page 117. You must also make sure that the required queue size survives boot. See “How to Configure Module Configuration Parameters to Survive a Boot” on page 117.

Enlarging this queue allows the member to save more updates from other members. However, be aware that each saved update consumes memory. When changing this variable you should consider carefully consider the memory implications.

Enlarging the Receiving Queue

The Receiving Queue on the cluster member keeps the updates from each cluster member until it has received a complete sequence of updates. The default (and minimum) size of this queue is 128. Each member keeps a Receiving Queue for each of the peer members.
To enlarge the receiving queue size, change the value of the global parameter `recv_queue_size`. See “How to Configure Module Configuration Parameters” on page 117. You must also make sure that the required queue size survives boot. See “How to Configure Module Configuration Parameters to Survive a Boot” on page 117.

Enlarging this queue means that the member can save more updates from other members. However, be aware that each saved update consumes memory. When changing this variable you should carefully consider the memory implications.

**Enlarging the Sync Timer**

The `sync timer` performs sync related actions every fixed interval. By default, the sync timer interval is 100ms. The base time unit is 100ms (or 1 tick), which is therefore the minimum value.

To enlarge the sync timer, change the value of the global parameter `fwha_timer_sync_res`. See “How to Configure Module Configuration Parameters” on page 117. The value of this variable can be changed while the system is working. A reboot is not needed.

By default, `fwha_timer_sync_res` has a value of 1, meaning that the sync timer operates every base time unit (every 100ms). If you configure this variable to n, the timer will be operated every n*100ms.

**Enlarging the CPHA Timer**

The `CPHA timer` performs cluster related actions every fixed interval. By default, the CPHA timer interval is 100ms. The base time unit is 100ms (or 1 tick), which is also the minimum value.

If the cluster members are geographically separated from each other, set the CPHA timer to be around 10 times the round-trip delay of the sync network.

Enlarging this value increases the time it takes to detect a failover. For example, if detecting interface failure takes 0.3 seconds, and the timer is doubled to 200ms, the time needed to detect an interface failure is doubled to 0.6 seconds.

To enlarge the sync timer, change the value of the global parameter `fwha_timer_cpha_res`. See “How to Configure Module Configuration Parameters” on page 117. The value of this variable can be changed while the system is working. A reboot is not needed.

By default, `fwha_timer_cpha_res` has a value of 1, meaning that the CPHA timer operates every base time unit (every 100ms). If you configure this variable to n, the timer will be operated every n*100ms.
Reconfiguring the Acknowledgment Timeout

A cluster member deletes updates from its Sending Queue (described in “Sending queue size” on page 97) on a regular basis. This frees up space in the queue for more recent updates.

The cluster member deletes updates from this queue if it receives an ACK about the update from the peer member.

The peer member sends an ACK in one of two circumstances—on condition that the Block New Connections mechanism (described in “Blocking New Connections Under Load” on page 119) is active:

- After receiving a certain number of updates.
- If it didn’t send an ACK for a certain time. This is important if the sync network has a considerable line delay, which can occur if the cluster members are geographically separated from each other.

To reconfigure the timeout after which the member sends an ACK, change the value of the global parameter `fw_sync_ack_time_gap`. See “How to Configure Module Configuration Parameters” on page 117. The value of this variable can be changed while the system is working. A reboot is not needed.

The default value for this variable is 10 ticks (10 * 100ms). Thus, if a member didn’t send an ACK for a whole second, it will send an ACK for the updates it received.

Contact Technical Support

If the other recommendations do not help solve the problem, contact Technical Support for further assistance.

ClusterXL Error Messages

In This Section

General ClusterXL Error Messages  page 100
SmartView Tracker Active Mode Messages  page 101
Sync Related Error Messages  page 102
TCP Out-of-State Error Messages  page 103
Platform Specific Error Messages  page 104

This section lists the ClusterXL error messages. For other, less common error messages, see SecureKnowledge solution sk23642 at http://support.checkpoint.com/kb/.
General ClusterXL Error Messages

- **FW-1**: changing local mode from <mode1> to <mode2> because of ID <machine_id>
  
  This log message can happen if the working mode of the cluster members is not the same, for example, if one machine is running High Availability, and another Load Sharing Multicast or Unicast mode. In this case, the internal ClusterXL mechanism tries to synchronize the configuration of the cluster members, by changing the working mode to the lowest common mode. The order of priority of the working modes (highest to lowest) is: 1. Synchronization only 2. Load Sharing 3. High Availability (Active Up) 4. High Availability (Primary Up).

- **CPHA**: Received confirmations from more machines than the cluster size
  
  This log message can occur during policy installation on the cluster. It means that a serious configuration problem exists in that cluster. Probably some other cluster has been configured with identical parameters and both of them have common networks.

- **fwldbcast_timer**: peer X probably stopped...
  
  This is caused when the member that printed this message stops hearing certain types of messages from member X. Verify that cphaprob state shows all members as active and that fw_ctl_pstat shows that sync is configured correctly and working properly on all members. In such a case it is fair to assume that there was a temporary connectivity problem that was fixed in the meantime. There may be several connections may suffer from connectivity problems due to that temporary synchronization problem between the two members. On the other hand, this can indicate that the other member is really down.

- **FW-1**: fwha_notify_interface: there are more than 4 IPs on interface <interface name> notifying only the first ones
  
  A member of the same cluster as the reporting machine has more than three virtual IP addresses defined on the same interface. This is not a supported configuration and will harm ClusterXL functionality.

- **Sync could not start because there is no sync license**
  
  This is a license error message: If you have a basic VPN-1 Pro license then sync is also licensed. Check the basic VPN-1 Pro license using cplic print and cplic check.

- **FW-1**: h_slink: an attempt to link to a link kbuf id not found
  
  fw_conn_post_inspect: fwconn_init_links failed
  
  Several problems of this sort can happen during a full sync session when there are connections that are opened and closed during the full sync process. Full sync is automatic as far as possible, but it is not fully automatic for reasons of performance. A gateway continues to process traffic even when it is serving as a full sync server.
This can cause some insignificant problems, such as a connection that is being deleted twice, a link to an existing link, and so forth. It should not affect connectivity or cause security issues.

- **Error SEP_IKE_owner_outbound**: other cluster member packet in outbound Cluster in not synchronized. Usually happens in OPSEC certified third-party load sharing products for which **Support non-sticky connections** is unchecked in the cluster object **3rd Party Configuration** page. (Or equivalently, in NG FP3 clusters, where the property `use_limited_flushnack` is set to false).

- **FW-1: fwha_pnote_register**: too many registering members, cannot register

  The critical device (also known as Problem Notification, or pnote) mechanism can only store up to 16 different devices. An attempt to configure the 17th device (either by editing the `cphaprof.conf` file or by using the `cphaprof -d ... register command`) will result in this message.

- **FW-1: fwha_pnote_register**: `<NAME>` already registered (# `<NUMBER>`)

  Each device registered with the pnote mechanism must have a unique name. This message may happen when registering new pnote device, and means that the device `<NAME>` is already registered as with pnote number `<NUMBER>`.

- **FW-1: fwha_pnote_unregister**: attempting to unregister an unregistered device `<DEVICE NAME>`

  Indicates an attempt to unregister a device which is not currently registered.

- **FW-1: alert_policy_id_mismatch**: failed to send a log

  A log indicating that there is a different policy id between the two or more members was not sent. Verify all cluster members have the same policy (using `fw stat`). It is recommended to re-install the policy.

- **FW-1: fwha_receive_fwhap_msg**: received incomplete HAP packet (read `<number>` bytes)

  This message can be received when ClusterXL hears CCP packets of clusters of version 4.1. In that case it can be safely ignored.

### SmartView Tracker Active Mode Messages

The following error messages can appear in SmartView Tracker Active mode. These errors indicate that some entries may not have been successfully processed, which may lead to missing synchronization information on a cluster member and inaccurate reports in SmartView Tracker.

- **FW-1: fwlddist_adjust_buf**: record too big for sync. update Y for table `<id>` failed. fwlddist_state=`<val>`

  Indicates a configuration problem on a clustered machine. Either synchronization is misconfigured, or there is a problem with transmitting packets on the sync interface. To get more information on the source of the problem...
• Run `fw ctl pstat` (described in “Monitoring Synchronization (fw ctl pstat)” on page 84).

• In ClusterXL clusters, run `cphaprob -a if` to get the statuses of the interfaces (see “Monitoring Cluster Interfaces (cphaprob [-a] if)” on page 71).

To solve this problem, see “Working with SmartView Tracker Active Mode” on page 120.

• `FW-1: fwldbcast_flush`: active connections is currently enabled and due to high load it is making sync too slow to function properly. X active updates were dropped

Indicates that a clustered machine has dropped SmartView Tracker Active mode updates in order to maintain sync functionality. To solve this problem, see “Working with SmartView Tracker Active Mode” on page 120.

### Sync Related Error Messages

• `FW-1: fwldbcast_retreq`: machine `<MACHINE_ID>` sent a retrans request for seq `<SEQ_NUM>` which is no longer in my possession (current seq `<SEQ_NUM>`)

This message appears when the local member receives a retransmission request for a sequence number which in no longer in its sending window. This message can indicate a sync problem if the sending member didn’t receive the requested sequence.

• `FW-1: fwlddist_save`: WARNING: this member will not be fully synchronized!

`FW-1: fwlddist_save`: current delta sync memory during full sync has reached the maximim of `<MEM_SIZE>` MB

`FW-1: fwlddist_save`: it is possible to set a different limit by changing `fw_sync_max_saved_buf_mem` value

These messages may appear only during full sync. While performing full sync the delta sync updates are being saved and are applied only after the full sync process has finished. It is possible to limit the memory used for saving delta sync updates by setting the `fw_sync_max_saved_buf_mem` variable to this limit.

• `FW-1: fwldbcast_flush`: `fwlddist_buf_ldbcast_unread` is not being reset fast enough (`ur=<UNREAD_LOC>,fwlddist_buflen=<BUFFER_LEN>`)  

This message may appear due to high load resulting in the sync buffer being filled faster than it is being read. A possible solution is to enlarge `fwlddist_buf_size`, as described in the “Working with SmartView Tracker Active Mode” on page 120.

• `FW-1: fwlddist_mode_change`: Failed to send trap requesting full sync

This message may appear due to a problem starting the full sync process, and indicates a severe problem. Contact Technical Support.
FW-1: State synchronization is in risk. Please examine your synchronization network to avoid further problems!

This message could appear under extremely high load, when a synchronization update was permanently lost. A synchronization update is considered to be permanently lost when it cannot be retransmitted because it is no longer in the transmit queue of the update originator. This scenario does not mean that VPN-1 Pro will malfunction, but rather that there is a potential problem. The potential problem is harmless if the lost sync update was to a connection that runs only on a single member as in the case of unencrypted (clear) connections (except in the case of a failover when the other member needs this update).

The potential problem can be harmful when the lost sync update refers to a connection that is non-sticky (see “Non-Sticky connections” on page 19), as is the case with encrypted connections. In this case the other cluster member(s) may start dropping packets relating to this connection, usually with a TCP out of state error message (see “TCP Out-of-State Error Messages” on page 103). In this case it is important to block new connections under high load, as explained in “Blocking New Connections Under Load” on page 119.

The following error message is related to this one.

FW-1: fwdbcast_recv: delta sync connection with member <MACHINE_ID> was lost and regained. <UPDATES_NUM> updates were lost.
FW-1: fwdbcast_recv: received sequence <SEQ_NUM> (fragm <FRAG_NUM>, index <INDEX_NUM>), last processed seq <SEQ_NUM>

These messages appear when there was a temporary sync problem and some of the sync updates were not synchronized between the members. As a result some of the connections might not survive a failover.

The previous error message is related to this one.

FW-1: The use of the non_sync_ports table is not recommended anymore. Refer to the user guide for configuring selective sync instead

Previous versions used a kernel table called non_sync_ports to implement selective sync, which is a method of choosing services that don’t need to be synchronized. Selective sync can now be configured from SmartDashboard. See “Choosing Services That Don’t Require Synchronization” on page 18.

**TCP Out-of-State Error Messages**

When the synchronization mechanism is under load, TCP packet out-of-state error messages may appear in the Information column of SmartTracker. This section explains how to resolve each error.
ClusterXL Error Messages

- **TCP packet out of state - first packet isn't SYN**
  tcp_flags: FIN-ACK
  TCP packet out of state - first packet isn't SYN tcp_flags: FIN-PUSH-ACK

  These messages occur when a FIN packet is retransmitted after deleting the connection from the connection table. To solve the problem, in SmartDashboard **Global properties** for **Stateful Inspection**, enlarge the TCP end timeout from 20 seconds to 60 seconds. If necessary, also enlarge the connection table so it won't fill completely.

- **SYN packet for established connection**

  This message occurs when a SYN is received on an established connection, and the sequence verifier is turned off. The sequence verifier is turned off for a non-sticky connection in a cluster (or in SecureXL). Some applications close connections with a RST packet (in order to reuse ports). To solve the problem, enable this behavior to specific ports or to all ports. For example, run the command:
  
  `fw ctl set -1 fw_trust_rst_on_port <port>`

  Which means that VPN-1 Pro should trust a RST coming from every port, in case a single port is not enough.

Platform Specific Error Messages

**Solaris Specific Error Messages**

- **WARNING: IP: Proxy ARP problem? Hardware address <MULTICAST MAC ADDRESS> think it is <VIRTUAL CLUSTER IP>**

  This Solaris console (or `/var/adm/messages`) message can appear in a Load Sharing Multicast mode cluster. It can be safely disregarded.

  This message can occur when a CCP packet is received by the IP stack but is not processed by the clustering mechanism. This can happen when the clustering mechanism is not fully initialized during boot, but another member is already transmitting these kind of packets.

**Nokia Specific Error Messages**

- **FW-1: fwha_nok_get_mc_mac_by_ip: received a NULL query**
  FW-1: fwha_nok_get_mc_mac_by_ip: nokcl_get_clustermac returned unknown type <TYPE>

  These messages mean that automatic proxy ARP entries for static NAT configuration might not be properly installed.

- **FW-1: fwha_nokcl_sync_rx_f: received NULL mbuf from ipso. Packet dropped.**
  FW-1: fwha_nokcl_sync_rx_f: received packet with illegal flag=<FLAG>. drop packet.

  These messages mean that an illegal CPHA packet was received and will be dropped. If this happens more than few times during boot, the cluster malfunctions.
Platform Specific Error Messages

- FW-1: fwha_nokcl_reregister_rx: unregister old magic mac values with IPSO.
  FW-1: fwha_nokcl_reregister_rx: new magic mac values <MAC,FORWARD MAC> registered successfully with IPSO.
  A notification that the operation fw ctl set int fwha_magic_mac succeeded.

- FW-1: fwha_nokcl_reregister_rx: error in de-registration to the sync_rx (<ERR NUM>) new magic macs values will not be applied
  A notification that the operation fw ctl set int fwha_magic_mac failed.
  Previous MAC values will be retained.

- FW-1: fwha_nokcl_creation_f: error in registration ...
  FW-1: fwha_nok_init: NOT calling nokcl_register_creation since did not de-register yet.
  FW-1: fwha_nok_fini: failed nokcl_deregister_creation with rc=<ERROR NUM>
  These messages mean that an internal error in registration to the IPSO clustering mechanism. Verify that the IPSO version is supported with this FireWall-1 version and that the Nokia IP Clustering or VRRP cluster is configured properly.

- FW-1: successfully (dis)connected to Nokia Clustering
  A notification that should be normally received during FireWall-1 enforcement module initialization and removal.

- FW-1: fwha_pnote_register: noksr_register_with_status failed
  FW-1: fwha_nokia_pnote_expiration: mismatch between nokia device to ckp device <DEVICE NAME>
  FW-1: fwha_nokia_pnote_expiration: can not find the device nokia claims to be expired
  FW-1: fwha_noksr_report_wrapper: attempting to report an unregistered device <DEVICE NAME>
  These messages may appear as a result of a problem in the interaction between the Nokia and ClusterXL device monitoring mechanisms. A reboot should solve this problem. Should this problem repeat itself contact Check Point Technical support.

Solaris Platform Specific Issues: VLAN Switch Port Flapping

When a Solaris cluster member has the same MAC address for all its interfaces (which is the default for some network cards), and each interface is connected to a different VLAN on the same switch, the switch ports may flap.

This can occur if the switch does not have the intelligence to allow packets with identical MAC addresses to come in to more than one switch port, even though those ports are in different VLANs.

To solve this, configure the eeprom on every cluster members by setting local-mac-addresses?=true
**Preventing Broadcast Storms on Switches**

Certain switches and routers have an upper limit on the number of broadcasts and multicasts that they can pass, in order to prevent broadcast storms. This limit is usually a percentage of the total interface bandwidth.

In Load Sharing Unicast or Load Sharing Multicast mode, if the Cluster Control Protocol uses multicasts, the switch or router may consider that a broadcast storm is occurring, in which case it will drop packets.

It is possible to either turn off broadcast storm control, or to allow a higher level of broadcasts or multicasts through the switch.

For information about configuring traffic-storm control on Cisco Catalyst 6500 Series Switch routers, see:

For information about configuring port-based traffic control on Cisco Catalyst 4500 Series Switches, see

**Member Fails to Start After Reboot**

If a reboot (or `cpstop` followed by `cpstart`) is performed on a cluster member while the cluster is under severe load, the member may fail to start correctly. The starting member will attempt to perform a full sync with the existing active member(s) and may in the process use up all its resources and available memory. This can lead to unexpected behavior.
To overcome this problem, define the maximum amount of memory that the member may use when starting up for synchronizing its connections with the active member. By default this amount is not limited. Estimate the amount of memory required as follows:

**TABLE 5-4 Memory required (MB) for Full Sync.**

<table>
<thead>
<tr>
<th>Number of open Connections</th>
<th>100</th>
<th>1000</th>
<th>5000</th>
<th>10,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>1.1</td>
<td>6.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10000</td>
<td>11</td>
<td>69</td>
<td>329</td>
<td></td>
</tr>
<tr>
<td>20000</td>
<td>21</td>
<td>138</td>
<td>657</td>
<td>1305</td>
</tr>
<tr>
<td>50000</td>
<td>53</td>
<td>345</td>
<td>1642</td>
<td>3264</td>
</tr>
</tbody>
</table>

*Note* - These figures were derived for cluster members using the Windows platform, with Pentium 4 processors running at 2.4 GHz.

For example, if the cluster holds 10,000 connections, and the connection rate is 1000 connections/sec you will need 69 MB for full sync.

Define the maximum amount of memory using the module global parameter: `fw_sync_max_saved_buf_mem`.

The units are in megabytes. For details, see “Advanced Cluster Configuration using Module Configuration Parameters” on page 117.
Member Fails to Start After Reboot
ClusterXL Advanced Configuration

In This Chapter

- Upgrading ClusterXL Clusters page 110
- Working with NAT and Clusters page 111
- Working with VLANS and Clusters page 113
- Advanced Cluster Configuration using Module Configuration Parameters page 117
  - Controlling the Clustering and Synchronization Timers page 119
  - Blocking New Connections Under Load page 119
  - Working with SmartView Tracker Active Mode page 120
  - Reducing the Number of Pending Packets page 121
  - Configuring Full Synchronization Advanced Options page 121
- Defining Disconnected Interfaces page 122
- Enhanced Enforcement of the TCP 3-Way Handshake page 123
- Configuring Cluster Addresses on Different Subnets page 124
- Moving from High Availability Legacy to High Availability New Mode or Load Sharing with Minimal Effort page 129
- Moving from High Availability Legacy to High Availability New Mode or Load Sharing with Minimal Downtime page 131
- Moving from a Single Gateway to a ClusterXL Cluster page 133
- Adding Another Member to an Existing Cluster page 134
- Configuring ISP Redundancy on a Cluster page 135
Upgrading ClusterXL Clusters

For detailed information about how to upgrade a ClusterXL or OPSEC certified gateway cluster, see The Upgrade Guide.

Working with VPN-1 Pro and Clusters

In This Section

How to Configure VPN and Clusters page 110

How to Define a Cluster Object for a VPN Peer with a Separate Manager page 111

How to Configure VPN and Clusters

Configuring a VPN-1 Pro Gateway cluster in SmartDashboard is very similar to configuring a single VPN-1 Pro Gateway. All attributes of the VPN are defined in the Gateway Cluster object, except for two attributes that are defined per cluster member.

1 Go to the Gateway Cluster Properties window, Cluster Members page. For each cluster member, in the Cluster member Properties window, configure the VPN tab:
   - Office Mode for Remote access — If you wish to use Office Mode for remote access, define the IP pool allocated to each cluster member.
   - Hardware Certificate Storage List — If your cluster member supports hardware storage for IKE certificates, define the certificate properties. In that case, SmartCenter Server directs the cluster member to create the keys and supply only the required material for creation of the certificate request. The certificate is downloaded to the cluster member during policy install.

2 In a VPN cluster, IKE keys are synchronized. In the Synchronization page of the Gateway Cluster Properties window, make sure that Use State Synchronization is selected, even for High Availability configurations.

3 In the Topology page of the Gateway Cluster Properties window, define the encryption domain of the cluster. Under VPN Domain, choose one of the two possible settings:
   - All IP addresses behind cluster members based on topology information. This is the default option.
   - Manually Defined. Use this option if the cluster IP address is not on the member network, in other words, if the cluster virtual IP address is on a different subnet than the cluster member interfaces. In that case, select a network or group of networks, which must include the virtual IP address of the cluster, and the network or group of networks behind the cluster.
How to Define a Cluster Object for a VPN Peer with a Separate Manager

When working with a VPN peer that is a Check Point Gateway cluster, and the VPN peer is managed by a different SmartCenter Server, do NOT define another cluster object. Instead:

1. In the objects tree, **Network Objects** branch, right click and select **New Check Point Externally Managed Gateway**.

2. In the **Topology** page, add the external and internal *cluster* interface addresses of the VPN peer. Do not use the cluster member interface addresses, except in the following cases:
   - If the external cluster is of version 4.1, add the IP addresses of the cluster member interfaces.
   - If the cluster is an OPSEC certified product (excluding Nokia), you may need to add the IP addresses of the cluster members.

   When adding cluster member interface IP addresses, in the interface **Topology** tab, define the interface as **Internal**, and the **IP Addresses behind this interface** as **Not defined**.

3. In the **VPN Domain** section of the page, define the encryption domain of the externally managed gateway to be behind the internal virtual IP address of the gateway. If the encryption domain is just one subnet, choose **All IP addresses behind cluster members based on topology information**. If the encryption domain includes more than one subnet, it must be **Manually Defined**.

Working with NAT and Clusters

In This Section

- **Cluster Fold and Cluster Hide**
- **Configuring NAT on the Gateway Cluster**
- **Configuring NAT on a Cluster Member**

**Cluster Fold and Cluster Hide**

Network Address Translation (NAT) is a fundamental aspect of the way ClusterXL works.
When a cluster member establishes an *outgoing* connection towards the Internet, the source address in the outgoing packets, is the physical IP address of the cluster member interface. The source IP address is changed using NAT to that of the external virtual IP address of the cluster. This address translation is called “Cluster Hide”.

For OPSEC certified clustering products, this corresponds to the default setting in the 3rd Party Configuration page of the cluster object, of *Hide Cluster Members’ outgoing traffic behind the Cluster's IP address* being checked.

When a client establishes an *incoming* connection to external (virtual) address of the cluster, ClusterXL changes the destination IP address using NAT to that of the physical external address of one of the cluster members. This address translation is called “Cluster Fold”.

For OPSEC certified clustering products, this corresponds to the default setting in the 3rd Party Configuration page of the cluster object, of *Forward Cluster's incoming traffic to Cluster Members’ IP addresses* being checked.

### Configuring NAT on the Gateway Cluster

Network Address Translation (NAT) can be performed on a Gateway Cluster, in the same way as it is performed on a Gateway. This NAT is in addition to automatic the “Cluster Fold” and “Cluster Hide” address translations.

To configure NAT, edit the Gateway Cluster object, and in the Gateway Cluster Properties window, select the NAT page. Do NOT configure the NAT tab of the Cluster Member object.

### Configuring NAT on a Cluster Member

It is possible to perform Network Address Translation (NAT) on a non-cluster interface of a Cluster Member.

A possible scenario for this is if the non-Cluster interface of the Cluster Member is connected to another (non-cluster) internal VPN-1 Pro Gateway, and you wish to hide the address of the non-Cluster interface of the Cluster Member.

Performing this NAT means that when a packet originates behind or on the non-Cluster interface of the Cluster Member, and is sent to a host on the other side of the internal VPN-1 Pro Gateway, the source address of packet will be translated.

Configure NAT on a non-cluster interface of a Cluster Member Gateway as follows:

1. Edit the Gateway Cluster object.
2. In the Cluster Member page of the Gateway Cluster Properties window, edit the Cluster Member object.
3. In the **Cluster Member Properties** window, click the **NAT** tab.

4. Configure Static or Hide NAT as desired.

### Working with VLANS and Clusters

In This Section

- **VLAN Support in ClusterXL**
- **Connecting Several Clusters on the Same VLAN**

#### VLAN Support in ClusterXL

A VLAN switch tags packets that originate in a VLAN with a four-byte header that specifies which switch port it came from. No packet is allowed to go from a switch port in one VLAN to a switch port in another VLAN, apart from ports (“global” ports) that are defined so that they belong to all the VLANS.

The cluster member is connected to the global port of the VLAN switch, and this logically divides a single physical port into many VLAN ports.

When configuring virtual interfaces on Solaris GigaSwift interfaces, and no corresponding physical interface is defined, ClusterXL may not recognize the virtual interfaces. If the virtual interface is not recognized, it will not run a monitoring mechanism on them and eventually it will not perform a failover. To make ClusterXL work properly on such virtual interfaces, the correspondent physical interface must be defined. For example, when a CE device with an instance 0 is defined on the system, the `/etc/hostname.ce0` file must be created and must contain some arbitrary IP address that will be assigned to the physical interface.

![Note]

- **Note**: ClusterXL does not support VLANS on Windows 2000 or Windows 2003 Server.

#### Connecting Several Clusters on the Same VLAN

It is not recommended to connect the non-secured interfaces (the internal or external cluster interfaces, for example) of multiple clusters to the same VLAN. A separate VLAN, and/or switch is needed for each cluster.

Connecting the secured interfaces (the synchronization interfaces) of multiple clusters is also not recommended for the same reason. Therefore, it is best to connect the secured interfaces of a given cluster via a crossover link when possible, or to an isolated VLAN.
If there is a need to connect the secured or the non-secured interfaces of multiple clusters to the same VLAN you need to make changes to:

- The destination MAC address, to enable communication between the cluster and machines outside the cluster (for ClusterXL Load Sharing Multicast Mode clusters only).
- The source MAC address of the cluster, to enable Cluster Control Protocol communication between cluster members.

**Changes to the Destination MAC Address**

This section applies to ClusterXL Load Sharing Multicast Mode only.

**How the Destination Cluster MAC Address is Assigned in Load Sharing Multicast Mode**

When a machine that is outside the cluster wishes to communicate with the cluster, it sends an ARP query with the cluster (virtual) IP address. The cluster replies to the ARP request with a multicast MAC address, even though the IP address is a unicast address.

This destination multicast MAC address of the cluster is based on the unicast IP address of the cluster. The upper three bytes are 01.00.5E, and they identify a Multicast MAC in the standard way. The lower three bytes are the same as the lower three bytes of the IP address. An example MAC address based on the IP address 10.0.10.11 is shown in FIGURE 6-1.

**Duplicate Multicast MAC Addresses: The Problem**

When more than one cluster is connected to the same VLAN, the last three bytes of the IP addresses of the cluster interfaces connected to the VLAN must be different. If they are the same, then communication from outside the cluster that is intended for one of the clusters will reach both clusters, which will cause communication problems.
For example, it is OK for the cluster interface of one of the clusters connected to the VLAN to have the address 10.0.10.11, and the cluster interface of a second cluster to have the address 10.0.10.12. However, the following addresses for the interfaces of the first and second clusters will cause complications: 10.0.10.11 and 20.0.10.11.

**Duplicate Multicast MAC Addresses: The Solution**

The best solution is to change to the last three bytes of the IP address of all but one of the cluster interfaces that share the same last three bytes of their IP address.

If the IP address of the cluster interface cannot be changed, you must change the automatically assigned multicast MAC address of all but one of the clusters and replace it with a user-defined multicast MAC address. Proceed as follows:

1. In the **ClusterXL** page of the cluster object, select **Load Sharing>Multicast Mode**. In the **Topology** tab, edit the cluster interface that is connected to same VLAN as the another cluster.

2. In the **Interface Properties** window, **General** tab, click **Advanced**.

3. Change the default MAC address, and carefully type the new user defined MAC address. It must be of the form 01:00:5e:xy:yy:yy where x is between 0 and 7 and y is between 0 and f (hex).

**Changes to the Source MAC Address**

This section applies to all ClusterXL modes, both High Availability and Load Sharing, and to OPSEC certified clustering products.

**How the Source Cluster MAC Address is Assigned**

Cluster members communicate with each other using the Cluster Control Protocol (CCP). CCP packets are distinguished from ordinary network traffic by giving CCP packets a unique source MAC address.

- The first four bytes of the source MAC address are all zero: 00.00.00.00
- The fifth byte of the source MAC address is a magic number. Its value indicates its purpose

**TABLE 6-1**

<table>
<thead>
<tr>
<th>Default value of fifth byte</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>0xfe</td>
<td>CCP traffic</td>
</tr>
<tr>
<td>0xfd</td>
<td>Forwarding layer traffic</td>
</tr>
</tbody>
</table>

- The sixth byte is the ID of the sending cluster member
Duplicate Source Cluster MAC Addresses: The Problem

When more than one cluster is connected to the same VLAN, if CCP and forwarding layer traffic uses multicast, this traffic reaches only the intended cluster.

However, if broadcast is used for CCP and forwarding layer traffic (and in certain other cases), cluster traffic intended for one cluster is seen by all connected clusters, and is processed by the wrong cluster, which causes communication problems.

Duplicate Source Cluster MAC Addresses: The Solution

To ensure that the source MAC address in packets from different clusters that are connected to the same VLAN can be distinguished, change the MAC source address of the cluster interface that is connected to the VLAN in all but one of the clusters.

Use the following module configuration parameters to set more than one cluster on the same VLAN. These parameters apply to both ClusterXL and OPSEC certified clustering products.

<p>| TABLE 6-2 |</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>fwha_mac_magic</td>
<td>0xfe</td>
</tr>
<tr>
<td>fwha_mac_forward_magic</td>
<td>0xfd</td>
</tr>
</tbody>
</table>

Changing the values of these module configuration parameters alters the fifth part of the source MAC address of Cluster Control Protocol and forwarded packets. Use any value as long as the two module configuration parameters are different. To avoid confusion, do not use the value 0x00.

When Performance Pack is used to enhance the performance of ClusterXL Load Sharing Multicast Mode, the values of fwha_mac_magic and fwha_mac_forward_magic, it is recommended that the chosen numbers be consecutive, with the lower one being even (for example 0x10 and 0x11, or 0xBE and 0xBF).

For instruction about how to change these parameters, see “How to Configure Module Configuration Parameters” on page 117.
Advanced Cluster Configuration using Module Configuration Parameters

In This Section

- How to Configure Module Configuration Parameters  page 117
- How to Configure Module Configuration Parameters to Survive a Boot  page 117
- Controlling the Clustering and Synchronization Timers  page 119
- Blocking New Connections Under Load  page 119
- Working with SmartView Tracker Active Mode  page 120
- Reducing the Number of Pending Packets  page 121
- Configuring Full Synchronization Advanced Options  page 121

How to Configure Module Configuration Parameters

A number of synchronization and ClusterXL capabilities are controlled by means of VPN-1 Pro enforcement module configuration parameters. Run these commands on the VPN-1 Pro Gateway machine as follows:

```
fw ctl set int Parameter <value>
```

Parameter is any of the parameters described in the following sections.

These configuration parameters are only available for version NG with Application Intelligence and higher clusters.

Changes to their default values must be implemented on all cluster members. Setting different values on cluster members can cause configuration problems and possibly connection failures.

All these module configuration parameters can be configured to survive a boot. The way to do this varies with the operating system.

How to Configure Module Configuration Parameters to Survive a Boot

Module configuration parameters that are changed using the `fw ctl set int` command do not survive reboot. The way to do make them survive a reboot varies with the operating system. In the following instructions, Parameter is any of the parameters described in the following sections.
Linux/SecurePlatform
1  Edit the file $FWDIR/boot/modules/fw Kern.conf.
2  Add the line Parameter=<value in hex>.
3  Reboot.

Solaris
1  Edit the file /etc/system.
2  Add the line set fw:Parameter=<value in hex>.
3  Reboot.

Windows
1  Edit the registry.
2  Add a DWORD value named Parameter under the key
   HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\FW1\Parameters\G
   lobals.
3  Reboot.

On Nokia
Run the command
modzap _Parameter $FWDIR/boot/modules/fwmod.o <value in hex>.
Note that the underscore before the parameter is not a mistake.
Controlling the Clustering and Synchronization Timers

The following module configuration parameters are used to control the clustering and synchronization timers. Changing the default values is not recommended.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>fwha_timer_cpha_res</td>
<td>The frequency of ClusterXL operations on the cluster. Operations occur every: 10 multiplied by fwha_timer_cpha_res multiplied by fwha_timer_base_res milliseconds</td>
<td>1</td>
</tr>
<tr>
<td>fwha_timer_sync_res</td>
<td>The frequency of synch flush operations on the cluster. Operations occur every: 10 multiplied by fwha_timer_sync_res multiplied by fwha_timer_base_res milliseconds</td>
<td>1</td>
</tr>
<tr>
<td>fwha_timer_base_res</td>
<td>Must be divisible by 10 with no remainders.</td>
<td>10</td>
</tr>
</tbody>
</table>

Blocking New Connections Under Load

The reason for blocking new connections is that new connections are the main source of new synchronization traffic, and synchronization may be put at risk if new traffic continues to be processed at this rate.

A related error message is “FW-1: State synchronization is in risk. Please examine your synchronization network to avoid further problems!” on page 103.

Reducing the amount of traffic passing through VPN-1 Pro protects the synchronization mechanism.
Advanced Cluster Configuration using Module Configuration Parameters

- `fw_sync_block_new_conns` allows FireWall-1 to detect heavy loads and start blocking new connections. Load is considered heavy when the synchronization transmit queue of a Firewall-1 starts to fill beyond the `fw_sync_buffer_threshold`. Set to -1 (the default) to disable load detection. Set to 0 to enable load detection. This parameter can be set to survive boot using the mechanism described in “How to Configure Module Configuration Parameters to Survive a Boot” on page 117.

- `fw_sync_buffer_threshold` is the maximum percentage of the buffer that may be filled before new connections are blocked. By default it is set to 80, with a buffer size of 512. By default, if more than 410 consecutive packets are sent without getting an ACK on any one of them, new connections are dropped. When blocking starts, `fw_sync_block_new_conns` is set to 1. When the situation stabilizes it is set back to 0.

- `fw_sync_allowed_protocols` is used to determine the type of connections that can be opened while the system is in a blocking state. Thus, the user can have better control over the system's behavior in cases of unusual load. The `fw_sync_allowed_protocols` variable is a combination of flags, each specifying a different type of connection. The required value of the variable is the result of adding the separate values of these flags. For example, the default value of this variable is 24, which is the sum of `TCP_DATA_CONN_ALLOWED` (8) and `UDP_DATA_CONN_ALLOWED` (16), meaning that the default allows only TCP and UDP data connections to be opened under load.

<table>
<thead>
<tr>
<th>Flag Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICMP_CONN_ALLOWED</td>
<td>1</td>
</tr>
<tr>
<td>TCP_CONN_ALLOWED</td>
<td>2 (except for data connections)</td>
</tr>
<tr>
<td>UDP_CONN_ALLOWED</td>
<td>4 (except for data connections)</td>
</tr>
<tr>
<td>TCP_DATA_CONN_ALLOWED</td>
<td>8 (the control connection should be established or allowed)</td>
</tr>
<tr>
<td>UDP_DATA_CONN_ALLOWED</td>
<td>16 (the control connection should be established or allowed)</td>
</tr>
</tbody>
</table>

**Working with SmartView Tracker Active Mode**

Active mode in SmartView Tracker shows connections currently open through any of the VPN-1 Pro enforcement modules that are sending logs to the currently active Log File on the Management Server.

Active mode tends to slow down synchronization. If that happens, the synchronization mechanism randomly drops Active connection updates in order to maintain synchronization. The drop will be accompanied by one of the error message described in “SmartView Tracker Active Mode Messages” on page 101.
Active mode view is not recommended on a heavily loaded cluster. To obtain a more accurate report of Active connections under load, two solutions are available. They apply both to a cluster and to a single VPN-1 Pro Gateway:

1. **Enlarge fwldist_buf_size**

   The fwldist_buf_size parameter controls the size of the synchronization buffer in words. (Words are used for both synchronization and in SmartView Tracker Active mode. 1 word equals 4kbytes). The default is 16k words. The maximum value is 64k words and the minimum value is 2k words.

   If changing this parameter, make sure that it survives boot, because the change is only applied after a reboot. Use the mechanism described in “How to Configure Module Configuration Parameters to Survive a Boot” on page 117.

2. **Obtain a Hotfix from Technical Support**

   Obtain a Check Point Technical Support Hotfix. This Hotfix has a variable that controls the rate at which Active connections are read by fwd on the enforcement module before being sent to the Management Server. Note that this solution requires additional CPU resources.

**Reducing the Number of Pending Packets**

ClusterXL prevents out-of-state packets in non-sticky connections. It does this by holding packets until a SYN-ACK is received from all other active cluster members. If for some reason a SYN-ACK is not received, VPN-1 Pro on the cluster member will not release the packet, and the connection will not be established.

To find out if held packets are not being released, run the fw ctl pstat command. If the output of the command shows that the Number of Pending Packets is large under normal loads (more than 100 pending packets), and this value does not decrease over time, use the fwldbcast_pending_timeout parameter to reduce the number of pending packets.

Change the value of fwldbcast_pending_timeout from the default value of 50 to a value lower than 50.

The value is in ticks units, where each tick is equal to 0.1 sec, so that 50 ticks is 5 seconds.

The value represents the time after which packets are released even if SYN-ACKs are not received.
**Configuring Full Synchronization Advanced Options**

When a cluster member comes up after being rebooted (or after `cpstart`), it has to perform Full Synchronization. As a first step in the Full Synchronization process, it performs a handshake with one of the other active cluster members. Only if this handshake succeeds does the cluster member continue with the Full Synchronization process.

The extended handshake that takes place (by default) from NG with Application Intelligence (R55), exchanges information between cluster members. This information includes version information, information about the installed Check Point products, and can include information about which Firewall-1 kernel tables are currently active. The extended handshake is unrelated to the exchange of kernel table information that happens later in the Full Synchronization.

All cluster members must have the same Check Point products and versions installed. The extended handshake identifies when different products are installed on the cluster members. When different products are installed, a console warning and a log message are issued.

In order to support backward compatibility, it is possible to change the behavior of the extended handshake by means of the following Module Configuration Parameters. How to edit these parameters is explained in “Advanced Cluster Configuration using Module Configuration Parameters” on page 117:

- `fw_sync_simplified_fullsync` has the default value of 0. It is used in NG with Application Intelligence (R54) and previous versions. The default value is required when performing the Full Connectivity Upgrade (described in *The Upgrade Guide*), because this upgrade requires an extended handshake to overcome version differences.

  Set to 1 in order for Full Synchronization to use the simplified handshake as it did in NG AI (R54).

- `fw_sync_no_ld_trans` has the default the value of 1. Set to 0 in order to exchange kernel table information between members in the first phase of the Full Synchronization process.

- `fw_sync_no_conn_trans` has the default value of 0. Set to 1 in order not to exchange installed product information between members in the first phase of the Full Synchronization process.

- `fw_sync_fcu_ver_check` has the default value of 1. Set to 0 to allow Full Connectivity Upgrade for versions that do not comply with the version requirements. Read about these requirements in *The Upgrade Guide*.
Defining Disconnected Interfaces

Disconnected interfaces are cluster member interfaces that are not monitored by the ClusterXL mechanism. If a disconnected interface fails, failover does not occur.

You may wish to define an interface as disconnected if the interface is down for a long time, and you wish the cluster member to continue to be active.

Defining a Disconnected Interface on Unix

Create a file in under \$FWDIR/conf/discntd.if and write the name of each interface that you do not want monitored by ClusterXL on a separate line.

Defining a Disconnected Interface on Windows

1. Open the regedit32 registry editor. Do not use regedit.
2. Under HKEY_LOCAL_MACHINES\System\CurrentControlSet\Services\CPHA create a new value with the following characteristics:
   Value Name: DisconnectedInterfaces
   Data Type: REG_MULTI_SZ
3. Open regedit registry editor. This time, do not use regedit32.
4. Add the interface name. To obtain the interface system name run the command: fw getifs
5. Add this name to the list of disconnected interfaces using the following format: \device\<System Interface Name>
6. Run cphastop and then cphastart to apply the change.

Enhanced Enforcement of the TCP 3-Way Handshake

The standard enforcement on the 3-way handshake that initiates a TCP connection provides good security enforcement by guaranteeing one-directional stickiness. This means that it ensures that the SYN-ACK will always arrive after the SYN. However, it does not guarantee that the ACK will always arrive after the SYN-ACK, or that the first data packet will arrive after the ACK.

If you wish to have an extra strict policy that denies all out-of-state packets, it is possible to configure the synchronization mechanism so that all the TCP connection initiation packets arrive in the right sequence (SYN, SYN-ACK, ACK, followed by the data). The price to be paid for this extra security is a considerable slowdown in connection establishment.
To configure enhanced enforcement, use the Database Tool to change the global property `sync_tcp_handshake_mode` from the default value of `minimal_sync` to `complete_sync`. 
Configuring Cluster Addresses on Different Subnets

In This Section

- Introduction page 124
- Configuration page 124
- Example page 126
- Limitations page 127

Introduction

Cluster IPs are virtual IP addresses given to ClusterXL objects, which differ from the unique IPs of the separate cluster machines. These addresses enable the cluster to be seen as a single gateway, thus allowing it to serve as a router in a network that is unaware of the cluster's internal structure and status.

Previously, cluster IP addresses had to be configured on the same subnets as those used by the unique addresses of the cluster members. As of NG with Application Intelligence, cluster IPs can reside on subnets other than those of the members. The advantage of this is that it enables a multi-machine cluster to replace a single-machine gateway in a pre-configured network, without the need to allocate new addresses to the cluster members.

Note · Cluster IP addresses can reside on subnets other than those of the members for ClusterXL only. For details about OPSEC certified clusters, see the vendor documentation.

An important aspect of this is that packets sent from cluster members (as opposed to packets routed through the members) are hidden behind the cluster IP and MAC addresses. The cluster MAC is the:

- MAC of the active machine, in High Availability New mode.
- Multicast MAC, in Load Sharing Multicast mode.
- Pivot member MAC in Load Sharing Unicast mode.

This enables the members to communicate with the surrounding networks, but also has certain limitations, as described in “Limitations” on page 127.

Configuration

There are two major steps required in order for ClusterXL to function correctly with cluster IPs on different subnets.
The first step is to create static routes on each cluster member, which determine the interface connected to the cluster's network (the subnet to which the cluster IP belongs). Unless these entries are created, the OS cannot route packets destined to the cluster's network. No additional configuration is required for the cluster members. It is, however, important to note that the unique IPs given to the members must share common subnets on each “side” of the cluster (meaning, each interface on each machine must have an interface on every other machine using the same subnet).

The second step relates to the configuration of the cluster topology. Here the cluster IPs are determined, and associated with the interfaces of the cluster members (each member must have an interface responding to each cluster IP). Normally, cluster IPs are associated with an interface based on a common subnet. In this case these subnets are not the same. It must be explicitly specified which member subnet is associated with the cluster IP. The **Member Network** tab in the Interface Properties window enables you to specify the member network (FIGURE 6-2).

Note that this interface actually refers to the cluster's virtual IP address, as determined in the cluster topology.

**FIGURE 6-2** Interface Properties - Member Network Tab
Example

In this example, a single-gateway firewall separating network 172.16.6.0 (Side “A”) from network 172.16.4.0 (Side “B”) is to be replaced with a ClusterXL cluster. The cluster members, however, will use networks 192.168.1.0 for Side “A”, 192.168.2.0 for Side “B” and 192.168.3.0 for the synchronization network (all network addresses given in this example are of class “C”). The addresses in italics are the cluster IP addresses. The resulting configuration is depicted in FIGURE 6-3:

FIGURE 6-3 Cluster addresses on different subnets

Configuring Static Routes on the Members

Each member should be configured with two static routes:

- One setting its 192.168.1.x IP address as the gateway for network 172.16.6.0
- One setting its 192.168.2.x IP address as the gateway for network 172.16.4.0.

The commands are platform specific. Examples follow:

**Linux**

```
route add -net 172.16.6.0 gw 192.168.1.1 netmask 255.255.255.0
```

**Windows**

```
routing 172.16.6.0 mask 255.255.255.0 192.168.1.1
```

**Solaris**

```
route add -net 172.16.6 192.168.1.1
```

If this does not work, and packets destined for the 172.16.6.x network are routed through the wrong interface after running the command, try the following instead:

```
rout add -net 172.16.6.0 127.0.0.1 -ifp <interface name of 192.168.1.1>
```

Chapter 6 ClusterXL Advanced Configuration 127
IPSO

Use Voyager to configure the static routes.

Configuring Cluster IP Addresses in SmartDashboard

The cluster IP addresses in this example should be configured in the Gateway cluster object Topology > Interface Properties window, as follows:

<table>
<thead>
<tr>
<th>General tab</th>
<th>Cluster Interface A IP address</th>
<th>Cluster Interface B IP address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member Networks tab</td>
<td>192.168.1.0</td>
<td>192.168.2.0</td>
</tr>
</tbody>
</table>

All IP addresses have the Netmask 255.255.255.0

Note - Do not define Cluster IP addresses for the synchronization interfaces. Define the synchronization interfaces in the Synchronization page of the Gateway Cluster object.

Limitations

The new feature does not yet support all the capabilities of ClusterXL. Some of those require additional configuration to work properly, while others do not work at all.

Connectivity between Cluster Members

Since ARP requests issued by cluster members are hidden behind the cluster IP and MAC, requests sent by one cluster member to the other may be ignored by the destination machine. To allow cluster members to communicate with each other, a static ARP should be configured for each cluster member, stating the MAC addresses of all other machines in the cluster. IP packets sent between members are not altered, and therefore no changes should be made to the routing table.

Note - Static ARP is not required in order for the machines to work properly as a cluster, since the cluster synchronization protocol does not rely on ARP.
Limitations

Chapter 6 ClusterXL Advanced Configuration

Load Sharing Multicast Mode with “Semi-Supporting” Hardware

Although not all types of network hardware work with multicast MAC addresses, some routers can pass such packets, even though they are unable to handle ARP replies containing a multicast MAC address. Where a router semi-supports Load sharing Multicast mode, it is possible to configure the cluster MAC as a static ARP entry in the router's internal tables, and thus allow it to communicate with the cluster.

When different subnets are used for the cluster IPs, static ARP entries containing the router's MAC need to be configured on each of the cluster members. This is done because this kind of router will not respond to ARP requests containing a multicast source MAC. These special procedures are not required when using routers that fully support multicast MAC addresses.

Automatic Proxy ARP

When using static NAT, the cluster can be configured to automatically recognize the hosts hidden behind it, and issue ARP replies with the cluster MAC, on their behalf. This process is known as Automatic Proxy ARP. If you use different subnets for the cluster IPs, this mechanism will not work, and you must configure the proxy ARP manually. This is done by creating a file called `local.arp`, under the firewall's configuration directory (`$FWDIR/conf`). In SmartDashboard, uncheck Automatic proxy arp.

Each entry in this file is a triplet, containing the:

- host address to be published
- MAC address that needs to be associated with the IP address
- unique IP of the interface that responds to the ARP request.

The MAC address that should be used is the cluster's multicast MAC defined on the responding interface, when using multicast LS, or this interface's unique IP, for all other modes.

For example, if host `172.16.4.3` is to be hidden using the address `172.16.6.25`, and the cluster uses Load Sharing Multicast mode, add the following line to the `local.arp` file of Member 1:

```
172.16.6.25 00:01:5e:10:06:64 192.168.1.1
```

The second parameter in this line is the multicast MAC address of cluster IP `172.16.6.100`, through which ARP requests for `172.16.6.25` will be received. On Member 2, this line will be:

```
172.16.6.25 00:01:5e:10:06:64 192.168.1.2
```

If the cluster is in unicast LS mode, or in HA mode, the entries on Member 1 and 2 will be:

```
172.16.6.25 00:A0:C9:E8:C7:7F 192.168.1.1
```
- And -
172.16.6.25 00:A0:C9:E8:CB:3D 192.168.1.2
where the second entry in each line is the unique MAC address of the matching local interface.

Connecting to the Cluster Members from the Cluster Network

Since the unique IPs may be chosen arbitrarily, there is no guarantee that these addresses are accessible from the cluster IP's subnet. In order to access the members through their unique IPs, you must configure routes on the accessing machine, such that the cluster IP is the gateway for the subnet of the unique IPs. Following the above example, 172.16.6.100 should be the gateway for subnet 192.168.1.0.

Default Gateway on SecurePlatform

Run `sysconfig > routing > add network route` > add the routable network with its subnet, and choose the correct physical interface in this direction.

Now go to `routing > add default gateway` and add the IP address of the default (routable) gateway. This will usually be the IP address of the router external.

If you have the different subnets feature configured on more than one interface, repeat the addition of the network address (as above) for all these interfaces. (It is NOT required to define a default gateway for the other subnets as well.)

Anti-Spoofing

When the different subnets feature is defined on a non-external interface, the cluster IP in the Cluster Topology tab should not be defined with the this network Anti-spoofing definition. You must add a group of networks that contain both the routable network and the non-routable network, and define the Anti-spoofing for this interface as specific network with this new group.

In the example shown on page 19, suppose side “B” is the internal network, you must define a group which contains both 172.16.4.0 and 192.168.2.0 networks, and define the new group in the specific field of the Topology tab.
The shared internal and external interfaces become cluster interfaces. The general IP address of the cluster therefore stays as an external cluster IP address.

**On the Modules**

1. Run `cpstop` on all members (all network connectivity will be lost).
2. Reconfigure the IP addresses on all the cluster members, so that unique IP addresses are used instead of shared (duplicate) IP addresses.

   **Note** - SecurePlatform only: These address changes delete any existing static routes. Copy them down for restoration in step 4.

3. Remove the shared MAC addresses by executing the command:
   `cphaconf uninstall_macs`
4. SecurePlatform cluster members only: Redefine the static routes deleted in step 2.
5. Reboot the members.

**On the Management**

1. In SmartDashboard, open the cluster object, select the **ClusterXL** tab, change the cluster mode from **Legacy mode** to **new mode** or to **Load sharing mode**. Click **OK** in the message that appears.
2. In the **Cluster Members** tab of the cluster, for each cluster member, get the interfaces which have changed since the IP addresses were changed. The interfaces which were previously used as **Shared Interfaces** should now be marked on the topology as **Cluster Interfaces** (if they are not already marked as such).
3. In the **Topology** tab of the cluster object, define the cluster IP addresses of the cluster. The cluster interfaces' names may be defined as you wish as they will be bound to physical interfaces according to the IP addresses. If the new IP addresses of the cluster members on a specific interface reside on different subnet than the cluster IP address in this direction, the cluster members' network should be defined in the **Members Network** fields of the cluster interface.
4. Install the policy on the new cluster object (Security policy, QOS policy and so on).
Moving from High Availability Legacy to High Availability New Mode or Load Sharing with Minimal Downtime

This procedure describes how to move from Legacy Check Point High Availability to New Check Point High Availability or to Load Sharing while minimizing the downtime of the cluster.

The shared internal and external interfaces become the cluster interfaces. As the cluster members will need additional IP addresses these must be prepared in advance.

If downtime of the cluster during the change is not a major issue, it is recommended to use the easier process described in “Moving from High Availability Legacy to High Availability New Mode or Load Sharing with Minimal Effort” on page 129.

**Note** -

1. Make sure that you have all the IP addresses needed before you start implementing the changes described here.

2. Backup your configuration before starting this procedure, because this procedure deletes and recreates the objects in SmartDashboard.

In this procedure we use the example of machines 'A' and 'B', with the starting point being that machine 'A' is active, and machine 'B' is on standby.

1. Disconnect machine 'B' from all interfaces except the interface connecting it to the management (the management interface).

2. Run `cphastop` on machine 'B'.

3. Change the IP addresses of machine 'B' (as required by the new configuration).

**Note** - SecurePlatform only: These address changes delete any existing static routes. Copy them down for restoration in step 6.

4. Reset the MAC addresses on machine 'B' by executing `cphaconf uninstall_macs`.

5. Reboot the machine for the MAC address change to take affect.

6. SecurePlatform cluster members only: Redefine the static routes deleted in step 3.

7. In SmartDashboard, right-click member 'A' and select **Detach from cluster**.

8. In the **Topology** tab of the **Cluster Member Properties** window, define the topology of cluster member 'B' by clicking **Get**. Make sure to mark the appropriate interfaces as **Cluster Interfaces**.
9. In the ClusterXL page, change the cluster’s High Availability mode from Legacy Mode to New Mode or select Load Sharing mode.

10. In the Cluster Object, define the new topology of the cluster (define the cluster interfaces in the cluster’s Topology tab).

11. Verify that the other pages in the Cluster Object (NAT, VPN, Remote Access and so on) are correct. In Legacy Check Point High Availability, the definitions were per cluster member, while now they are on the cluster itself.

12. Install the policy on the cluster, which now only comprises cluster member 'B'.

13. Reconnect machine 'B' (which you disconnected in step 1) to the networks.

14. In this example the cluster comprises only two members, but if the cluster comprises more then two members, repeat steps 1-9 for each cluster member.

15. For Load Sharing Multicast mode, configure the routers as described in “Multicast MAC” on page 42.

16. Disconnect machine 'A' from all networks except the management network. The cluster stops processing traffic.

17. Run cphastop on machine 'A'.

18. Run cpstop and then cpstart on machine 'B' (if there are more the two machines, run these commands on all machines except 'A').

19. Machine 'B' now becomes active and starts processing traffic.

20. Change the IP addresses of machine 'A' (as required by the new configuration).

21. Reset the MAC addresses of machine 'A' by executing cphaconf uninstall_macs.

22. Reboot the machine for the MAC address change to take affect.

23. In SmartDashboard, open the Cluster Object and select the Cluster Members page. Click Add > Add Gateway to Cluster and select member 'A' to re-attach it to the cluster.

24. Redefine the member with the new IP addresses.

25. Reconnect machine 'A' to the networks from which it was disconnected in step 13.

26. Install the security policy on the cluster.

27. Run cpstop and then cpstart on machine 'A'.

28. Redefine static routes

The cluster now operates in the new mode.
Moving from a Single Gateway to a ClusterXL Cluster

This procedure describes how to add a new gateway module (Machine 'B') to a standalone FireWall-1 enforcement module (Machine 'A') to create a cluster.

On the Single Gateway Machine

If your single gateway installation uses the same machine for the SmartCenter Server and the enforcement module:

1. Separate the SmartCenter Server from the enforcement module, and place them on two machines.
2. Initialize SIC on the separated enforcement module (Machine 'A').

On Machine 'B'

1. Define an interface on machine 'B' for each proposed cluster interface and synchronization interface on machine 'A', with the same subnet.
2. Install VPN-1 Pro on the machine. During the installation you must enable ClusterXL.

On the Management

1. Create a ClusterXL object.
2. In the Cluster Members page, click Add, and select New Cluster Member.
3. Connect to machine 'B', and define its topology.
4. Define the Synchronization networks for the cluster.
5. Define the cluster topology. The cluster IP addresses should be the same as the addresses of machine 'A', on its proposed cluster interfaces.
6. Install the policy on the cluster, currently including member 'B' only.

On Machine 'A'

1. Disconnect all proposed cluster and Synchronization interfaces. New connections now open through the cluster, instead of through machine 'A'.
2. Change the addresses of these interfaces to some other unique IP address (preferably on the same subnet as before.).
3 Connect each pair of interfaces of the same subnet using a dedicated network. Any hosts or gateways previously connected to the single gateway must now be connected to both machines, using the hub/VLAN.

**Note** - Starting NG with Application Intelligence (RS5), it is possible to run synchronization across a WAN. For details, see “Synchronizing Clusters over a Wide Area Network” on page 22.

**On the Management**

1 In the **Cluster Members** page, click **Add**, and select **Add Gateway to Cluster**.
2 Select machine 'A' in the window.
3 Open the **Member's Topology** page, and determine which interface is a cluster interface, and which is an internal or an external interface.
4 Install the policy on the cluster.

**Adding Another Member to an Existing Cluster**

1 On the cluster member, run `cpconfig` to enable ClusterXL.
2 Change the IP addresses of the new cluster member to reflect the correct topology (either shared IP addresses or unique IP addresses, depending on the clustering solution).
3 Ensure that all required Check Point products are installed on the new cluster member.
4 In the **Cluster Members** page of the Gateway Cluster object, either create a new cluster member (if it is a new FireWall-1 machine) with the appropriate properties, or convert an existing Gateway to a cluster member.
5 If this is a new FireWall-1 machine, ensure that SIC is initialized, and that the topology is correctly defined.
6 If the **Cluster Mode** is **Load Sharing** or **New HA**, ensure that the proper interfaces on the new cluster member are configured as **Cluster Interfaces**.
7 Install the security policy on the cluster.
8 The new member is now part of the cluster.
Configuring ISP Redundancy on a Cluster

If you have a ClusterXL Gateway cluster, connect each cluster member to both ISPs via a LAN using two interfaces. The cluster-specific configuration is illustrated in FIGURE 6-4.

Note that the member interfaces must be on the same subnet as the cluster external interfaces.

Configure ClusterXL in the usual way.

To configure ISP Redundancy, see the FireWall-1 guide, Chapter 7, “ISP Redundancy”, in the section “Configuring ISP Link Redundancy” on page 165.

FIGURE 6-4 Gateway Cluster Connected to Two ISP links
Appendix A

High Availability Legacy Mode

In This Appendix

Introduction to High Availability Legacy Mode       page 137
Example of High Availability HA Legacy Mode Topology page 138
Implementation Planning Considerations for HA Legacy Mode page 139
Configuring High Availability Legacy Mode           page 141

Introduction to High Availability Legacy Mode

In High Availability configurations, only one machine is active at any one time. A failure of the active machine causes a failover to the next highest priority machine in the cluster.

High Availability Legacy mode was the only available High Availability mode before NG FP3. When setting up High Availability for the first time, High Availability New mode is recommended.

In Legacy Mode the cluster members share identical IP and MAC addresses, so that the active cluster member receives from a hub or switch all the packets that were sent to the cluster IP address. A shared interface is an interface with MAC and IP addresses that are identical to those of another interface.

Moving from a single gateway configuration to a High Availability Legacy Mode cluster requires no changes to IP addresses, or routing, and any switch or hub can be used to connect interfaces. However, configuring this mode is complicated, and must be performed in a precise sequence in order to be successful. The SmartCenter Server has to be connected to non-shared cluster network, in other words, the synchronization network of the cluster, or to a dedicated management network.
Example of High Availability HA Legacy Mode Topology

In This Section

- Shared Interfaces IP and MAC Address Configuration
- The Synchronization Interface

FIGURE A-1 shows an example ClusterXL Topology for High Availability Legacy mode. The diagram relates the physical cluster topology to the required SmartDashboard configuration. It shows two cluster members: Member_A (the primary) and Member_B (the secondary) each with three interfaces. One for synchronization, one external shared interface, and one internal shared interface.

FIGURE A-1  Example High Availability Legacy Mode Topology

Shared Interfaces IP and MAC Address Configuration

High Availability Legacy mode uses identical IP and MAC addresses on all cluster members, on interfaces that face the same direction. Shared interfaces are configured with the same IP address, and they automatically obtain identical MAC addresses. One shared interface on each cluster member faces the Internet through a hub or switch, and one or more interfaces face the local networks through a hub or switch.
Only one cluster member is active at any given time, so that the outside world can see only the shared interfaces on one machine at any given time.

FIGURE A-1 shows the shared interfaces. The EXT interface, facing the Internet, has IP address 192.168.0.1 on both Member_A and Member_B. The INT interface facing the local network has IP address 172.20.10.1 on both Member_A and Member_B.

**The Synchronization Interface**

State Synchronization between cluster members ensures that if there is a failover, connections that were handled by the failed machine will be maintained. The synchronization network is used to pass connection synchronization and other state information between cluster members. This network therefore carries the most sensitive security policy information in the organization, and so it is important to make sure the network is secure. It is possible to define more than one synchronization network for backup purposes.

To secure the synchronization interfaces, they should be directly connected by a cross-cable, or in the case of a three of more member cluster, by means of a dedicated hub, switch, or VLAN.

Machines in a High Availability cluster do not have to be synchronized, though if they are not, connections may be lost upon failover.

FIGURE A-1 shows a SYNC interface with a unique IP address on each machine. 10.0.10.1 on Member_A and 10.0.10.2 on Member_B.

**Implementation Planning Considerations for HA Legacy Mode**

In This Section

- *IP Address Migration*  
- *SmartCenter Server Location*  
- *Routing Configuration*  
- *Switch (Layer 2 Forwarding) Considerations*

**IP Address Migration**

Many ClusterXL installations are intended to provide High Availability or Load Sharing to an existing single gateway configuration. In those cases, it is recommended to take the existing IP addresses from the current gateway, and make these the cluster addresses...
(cluster virtual addresses) when feasible. Doing so will avoid altering current IPSec endpoint identities, and in many cases will make it unnecessary to change Hide NAT configurations.

**SmartCenter Server Location**

The SmartCenter Management Server must be able to download a Security Policy to all cluster members. This is only possible if the SmartCenter Server can “see” them all at any given time. Therefore, in High Availability Legacy mode, the SmartCenter Server must be connected to non-shared cluster network.

The SmartCenter Server cannot be connected to any network that includes the cluster interfaces with shared IP addresses, because they are configured with identical IP and MAC addresses.

The SmartCenter Server must therefore be connected to the cluster synchronization network of the cluster, because the SYNC interface on each cluster member must have a unique IP address, or to a dedicated management network attached to the cluster.

**Routing Configuration**

Configure routing so that communication with the opposite side of the cluster is via the cluster IP address on the near side of the cluster.

For example, in FIGURE A-1, configure routing as follows:

- On each machine on the internal side of the router, define 172.20.0.1 as the default gateway.
- On external router, configure a static route such that network 172.20.0.1 is reached via 192.168.10.1.

**Switch (Layer 2 Forwarding) Considerations**

The Cluster Control Protocol (CCP), used by both High Availability New Mode and Load Sharing configurations, makes use of layer two multicast. In keeping with multicast standards, this multicast address is used only as the destination, and is used in all CCP packets sent on “non-secured” interfaces.

A layer two switch connected to non-secured interfaces, must be capable of forwarding multicast packets to ports of the switch, or within a VLAN, if it is a VLAN switch. It is acceptable that the switch forward such traffic to all ports, or to ports within the given VLAN. However, it is considered more efficient to forward to only those ports connecting cluster members.

Most switches support multicast by default. Please check your switch documentation for details.
If the connecting switch is incapable of forwarding multicast, CCP can be changed to use broadcast instead. To toggle between these two modes use the command:

`cphaconf set_ccp broadcast/multicast`

**Configuring High Availability Legacy Mode**

See FIGURE A-1 on page 138 for an example configuration.

1. Obtain and install a Central license for ClusterXL on the SmartCenter Server.
2. Disconnect the machines that are to participate in the High Availability Legacy configuration from the hub/switch.
3. Define the same IP addresses for each machine participating in the High Availability Legacy configuration, only for the interfaces that will be shared. To avoid network conflicts due to the sharing of MAC addresses, define the IP addresses before connecting the machines into the High Availability Legacy topology.
4. Install the same version (and build number) of VPN-1 Pro on each cluster member. During the configuration phase, enable ClusterXL/State Synchronization. Do NOT reboot the machines after the configuration phase.
5. Connect (or reconnect) the machines participating in the High Availability Legacy configuration to the hub/switch. Make sure you connect the configured interfaces to the matching physical network outlet. Connect each network (internal, external, Synchronization, DMZ, etc.) to a separate VLAN, switch or hub. No special configuration of the switch is needed.

**Routing Configuration**

6. Configure routing so that communication with the networks on the internal side of the cluster is via the cluster IP address on the external side of the cluster. For example, in FIGURE A-1, on the external router, configure a static route such that network 10.255.255.100 is reached via 192.168.10.100.
7. Configure routing so that communication with the networks on the external side of the cluster is via the cluster IP address on the internal side of the cluster. For example, in FIGURE A-1, on each machine on the internal side of the router, define 10.255.255.100 as the default gateway.
8. Reboot the cluster members. MAC address configuration will take place automatically.
SmartDashboard configuration

When configuring the Gateway Cluster object for HA Legacy mode, you need only configure the topology of the cluster members in the Cluster Members page. The cluster topology information in the Topology tab will be unavailable, or will be deleted if converting from another cluster mode. Cluster members exist solely inside the Gateway Cluster object. One (or more) interfaces of each cluster member will be the synchronization interface, which must be in the Synchronization network that is defined in the Synchronization page of the Gateway Cluster object.

Proceed as follows:

1. Using SmartDashboard, define the Gateway Cluster object. In the General Properties page of the Gateway Cluster object, assign the routable external IP address of the cluster as the general IP address of the cluster. Check ClusterXL as a product installed on the cluster.

2. In the Cluster Members page, click Add to add all cluster members to the cluster. Cluster members exist solely inside the Gateway Cluster object. For each cluster member,
   - initialize SIC (in the Cluster Members Properties > General tab),
   - define all interfaces (in the Topology tab),
   - define NAT and VPN tabs, if required.

3. In the ClusterXL page, check High Availability Legacy Mode, specify the action Upon Gateway Recovery, and define the Fail-over Tracking method.

4. In the Synchronization page, click Add to define the cluster synchronization network(s).
   If you do not use State Synchronization, existing connections will be closed when failover occurs. If you choose not to use this feature, deselect Use State Synchronization and skip to the next step.

5. Define the other pages in the Gateway Cluster object as required (NAT, VPN, Remote Access, etc.).


7. Reboot all the cluster members in order to activate the MAC address configuration on the cluster members.
Example cphaprob Script

The clusterXL_monitor_process script shown below has been designed to provide a way to check end-to-end connectivity to routers or other network devices and cause failover if the ping fails. The script monitors the existence of given processes and cause failover if the processes die. It uses the normal pnote mechanism.

The clusterXL_monitor_process script is located in $FWDIR/bin.

More information

- The cphaprob command is described in “How to Verify the Cluster is Working Properly (cphaprob)” on page 68.
- Chapter 5, “Monitoring and Troubleshooting Gateway Clusters” on page 67.

The clusterXL_monitor_process script

```sh
#!/bin/sh
#
# This script monitors the existence of processes in the system. The process names
# should be written
# in the $FWDIR/conf/cpha_proc_list file one every line.
#
# USAGE :
# cpha_monitor_process X silent
# where X is the number of seconds between process probings.
# if silent is set to 1, no messages will appear on the console.
#
# We initially register a pnote for each of the monitored processes
# (process name must be up to 15 characters) in the problem notification mechanism.
```
# when we detect that a process is missing we report the pnote to be in "problem"
# state.
# when the process is up again - we report the pnote is OK.

if [ "$2" -le 1 ]
then
  silent=$2
else
  silent=0
fi

if [ -f $FWDIR/conf/cpha_proc_list ];
then
  procfile=$FWDIR/conf/cpha_proc_list
else
  echo "No process file in $FWDIR/conf/cpha_proc_list "
  exit 0
fi

arch=`uname -s`

for process in `cat $procfile`
do
  $FWDIR/bin/cphaprob -d $process -t 0 -s ok -p register > /dev/null 2>&1
  done

while [ 1 ]
do

  result=1

  for process in `cat $procfile`
  do
    ps -ef | grep $process | grep -v grep > /dev/null 2>&1
    status=$?
    if [ $status = 0 ]
    then
      if [ $silent = 0 ]
      then
        echo " $process is alive"
        fi
      fi
  # echo "3, $FWDIR/bin/cphaprob -d $process -s ok report"
      $FWDIR/bin/cphaprob -d $process -s ok report
      else
      if [ $silent = 0 ]
      then
        echo " $process is down"
        fi
  #
```bash
$FWDIR/bin/cphaprob -d $process -s problem report
result=0
fi
done
if [ $result = 0 ]
then
  if [ $silent = 0 ]
  then
    echo "One of the monitored processes is down!"
  fi
else
  if [ $silent = 0 ]
  then
    echo "All monitored processes are up"
  fi
fi
if [ "$silent" = 0 ]
then
  echo "sleeping"
fi
sleep $1
done
```
Appendix C

ClusterXL Command Line Interface

The following command line commands relate to ClusterXL and are documented in the Command Line Interface (CLI) Guide.

**TABLE C-1** Cluster-XL Command Line Interface

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cphaconf</td>
<td>Used to configure ClusterXL. Running this command is not recommended. It should be run only by Firewall-1. See “The cphaconf Command” on page 83.</td>
</tr>
<tr>
<td>cphaprob</td>
<td>Verifies that the cluster and the cluster members are working properly. See “How to Verify the Cluster is Working Properly (cphaprob)” on page 68. On Nokia VRRP and other OPSEC certified clusters, this command behaves differently. See “The cphaprob Command in OPSEC Clusters” on page 65.</td>
</tr>
<tr>
<td>cphastart</td>
<td>Running cphastart on a cluster member activates ClusterXL on the member. It does not initiate full synchronization. cpstart is the recommended way to start a cluster member. See “The cphastart and cphastop Commands” on page 83.</td>
</tr>
<tr>
<td>cphastop</td>
<td>Running cphastop on a cluster member stops the cluster member from passing traffic. State synchronization also stops. It is still possible to open connections directly to the cluster member. In High Availability Legacy mode, running cphastop may cause the entire cluster to stop functioning. See “The cphastart and cphastop Commands” on page 83.</td>
</tr>
</tbody>
</table>
Index

A
accounting
   and synchronized FireWalls 23

B
Broadcast storms 106

C
cable failure 40
Client Authentication
   High Availability 23
cphastart 84
cphastop 84

F
failover
   definition 12
   when does it occur 40
fw_sync_allowed_protocols 119
fw_sync_block_new_conns 119
fw_sync_buffer_threshold 119
fw_sync_max_saved_buf_mem 106
fw_sync_simplified_fullsync 121
fwha_timer_base_res 119
fwha_timer_cpha_res 119
fwha_timer_sync_res 119
fwldbcast_pending_timeout 121

H
High Availability
   and SmartView Tracker 79
   resources 23
   Security Servers 23

I
interface failure 40
IP address
   unique 139

M
Module Configuration
   Parameters 117
Module Configuration
   Parameters 117

S
synchronized FireWall Modules
   restrictions on implementation 22
synchronized FireWalls
   restrictions 22
synchronizing VPN/FireWall
   Modules
   on different platforms 22

U
unique IP address 139
User Authentication
   High Availability 23

V
VPN/FireWall Modules
   restrictions on synchronization 22