Consider the following scenario: A network uses one server that cannot handle the amount of traffic flowing through it. The Security Administrator obtains additional servers to handle the load, but this creates another problem. Most of the network traffic is still flowing through only one of the servers. To alleviate this problem, VPN-1/FireWall-1 load balancing redirects traffic, and distributes it among several servers. This reduces the load on any one server.

VPN-1/FireWall-1 load-balancing tools include the following:

- HTTP redirect
- Load measuring
- HTTP logical server
- Load-balancing algorithms
- Rule Base order

**Objectives**

1. Describe the purpose of load balancing.
2. Define the different methods for load balancing.

**Key Terms**

- ConnectControl
- Load balancing
- Address Resolution Protocol
- Load-balancing algorithm
- Load Measuring Agent
- Server Load, Round Trip, Round Robin, Random, and Domain algorithms
LOAD BALANCING

The Need for Server Load Balancing

THE NEED FOR SERVER LOAD BALANCING

Increasing dependence on the Internet to deliver important resources requires server infrastructures that can reliably support a burgeoning volume of traffic, and an ever-increasing number of client connections. For many network environments, reliance on a single server results in poor response times, or even connection time-outs. Unreliable connectivity with potential or existing customers can mean decreased revenues and lost business opportunities.

VPN-1/FireWall-1 can add ConnectControl (licensed separately), which incorporates advanced traffic-control functionality, to ensure the highest degree of network connectivity and optimal server response times. With ConnectControl, a single server providing Web (or any other) services, can be replaced with a logical pool of servers sharing a common IP address. Connection requests are load balanced among multiple servers. Network users experience noticeably improved response times, and are unaware of any ConnectControl intervention. At the same time, corporations are relieved of the need to continually upgrade to more expensive servers, to meet the increasing demand for service. Instead, existing hardware can be fully utilized to deliver a completely scalable traffic-management solution.

How Load Balancing Works

Load balancing allows several servers to share and distribute network load. The VPN-1/FireWall-1 Security Administrator does this by creating a logical server on the Enforcement Module. The logical server has a unique IP address, through which packets are routed for load balancing. Traffic directed to this logical server is then load shared among the physical servers in a logical-server group. Using Address Resolution Protocol, which is a method for finding a host’s Ethernet address from its Internet address, VPN-1/FireWall-1 load balancing ensures packets destined to the IP address of the logical server are passed to the appropriate physical server.
This figure illustrates how load balancing works. Using a specific load-balancing method, VPN-1/FireWall-1 routes packets to a specific physical server.
VPN-1/FireWall-1 load-balancing algorithms determine which physical server is best able to handle new communication requests. When a communication request to a logical server’s IP address reaches an Enforcement Module, the VPN-1/FireWall-1 load-balancing algorithm determines which physical server will fulfill the request. VPN-1/FireWall-1 includes five load-balancing algorithms. Each algorithm prevents any server from handling a disproportionate volume of traffic.

The five VPN-1/FireWall-1 load-balancing algorithms are:

- **Server Load** – determines the load of each physical server; Server Load requires a Load Measuring Agent, which is installed on each server and automatically reports current system load to the ConnectControl module.

- **Round Trip** – determines round trips between an Enforcement Module and each physical server; the Enforcement Module Pings each server three times, and averages the round-trip time. The server with the lowest average gets the packets.

- **Round Robin** – chooses the next physical server in a server group

- **Random** – chooses a physical server randomly

- **Domain** – chooses a physical server based on domain name; this is only useful for HTTP load balancing.
## ConnectControl Properties

ConnectControl Properties are set using the Global Properties screen. To access this screen from SmartDashboard, select > Policy > Global Properties > ConnectControl.

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Server availability check interval</strong></td>
<td>the interval, in seconds, in which an Enforcement Module will Ping a physical server, to determine if it is available</td>
</tr>
<tr>
<td><strong>Server check retries</strong></td>
<td>the number of consecutive times the server-availability check must fail, before an Enforcement Module considers a physical server unavailable, and no longer direct connections to it</td>
</tr>
<tr>
<td><strong>Persistent server timeout</strong></td>
<td>the length of time during which connections will be redirected to the same physical server, when Persistent Server Mode is enabled for a logical server in the Logical Server Properties</td>
</tr>
<tr>
<td><strong>Load agents port</strong></td>
<td>the port on which the Load Measuring Agent communicates</td>
</tr>
<tr>
<td><strong>Load measurement interval</strong></td>
<td>the intervals at which the Load Measuring Agent measures a load</td>
</tr>
</tbody>
</table>
LOAD BALANCING
Logical-Server Types

LOGICAL-SERVER TYPES

HTTP

When HTTP is chosen under Server Types in the Logical Server Properties screen, VPN-1/FireWall-1 detects any service requests for a logical server. This figure illustrates what happens when a client initiates an HTTP session on an HTTP logical server:

HTTP Redirect

1. VPN-1/FireWall-1 detects an HTTP request to a logical server, and redirects the request to the load-balancing application on the enforcement module.
2. The load-balancing application notifies the client that the request is being redirected to the destination physical (HTTP) server, based on the selected load-balancing algorithm.
3. The rest of the session is conducted between the client and the destination server, without the intervention of the load-balancing application.

If you are using HTTP redirect on an Enforcement Module, you should create one rule that specifies the logical server for the HTTP session to connect. Create a second rule that specifies the physical server group that will communicate directly with the client, throughout the rest of the session.
Other Load Balancing

“Other” load balancing places entries in the VPN-1/FireWall-1 address-translation tables for a connection. Other load balancing allows a server’s IP address to be a logical server’s address (from an Enforcement Module to a client), and a physical server’s IP address (from a server to the Enforcement Module). Each HTTP connection is then handled separately, and connections may be redirected to different servers. This may cause problems in some cases: for example, in an application where a user fills in a number of HTTP forms, and a single server is expected to process all data.

Non-HTTP

When a non-HTTP service request (such as FTP) is detected, a client starts an FTP session on an FTP logical server. VPN-1/FireWall-1 determines that the session is to be redirected to a particular FTP server. Using a mechanism similar to reverse Hide Network Address Translation, VPN-1/FireWall-1 modifies the destination IP address of incoming packets. If a return connection is opened, the connection is automatically established between the server and the client. The source IP address of the packets reverts to the logical server’s IP address.
LOAD BALANCING

Logical-Server Types