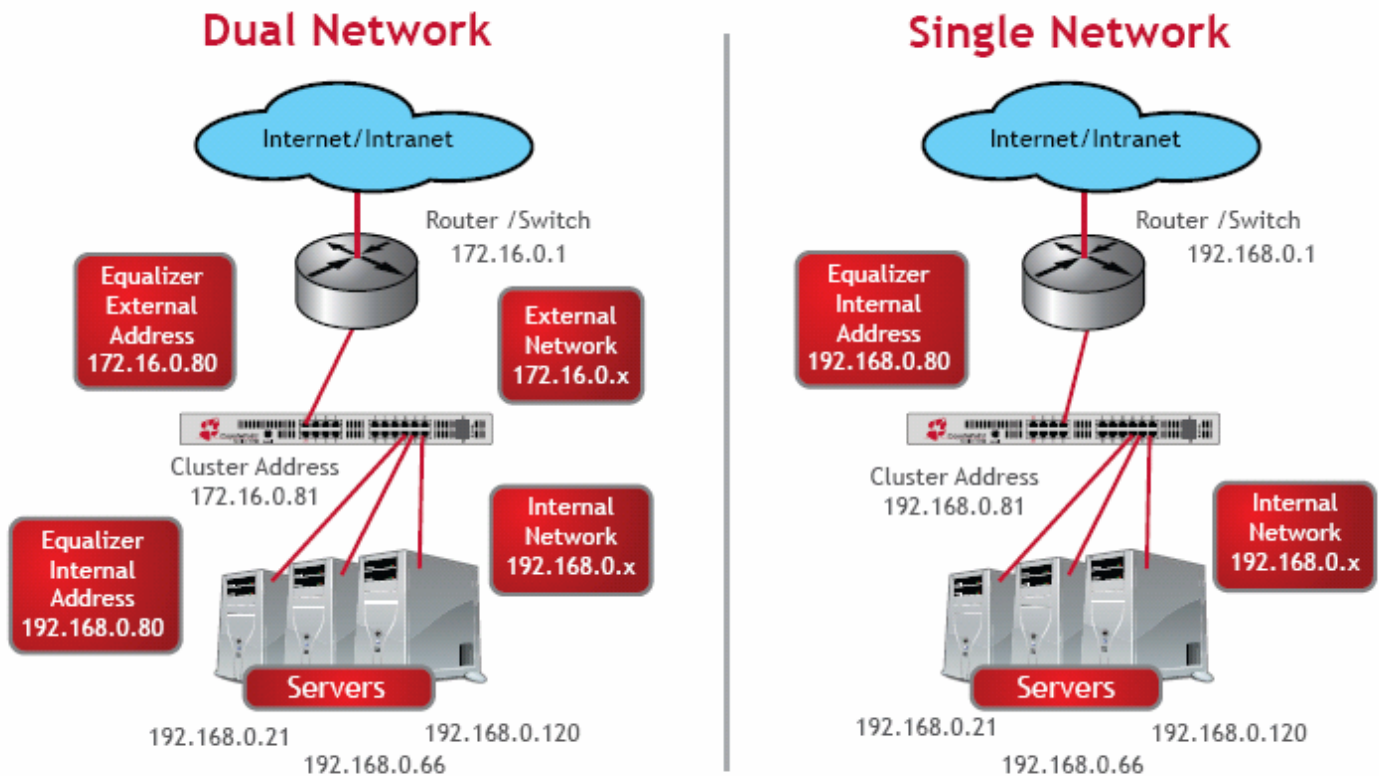


Five Easy Steps to Implementing Application Load Balancing for Non-Stop Availability and Higher Performance.



Coyote Point Systems White Paper

Load Balancing Guide for Application Server Administrators

<http://www.coyotepoint.com>



The idea of load balancing is well defined in the IT world: A network device accepts traffic on behalf of a group of servers, and distributes that traffic according to load balancing algorithms and the availability of the services that the servers provide. From network administrators to server administrators to application developers, this is a generally well understood concept.

The implementation of load balancing, however, is another matter. There are often many questions regarding how load balancing is deployed, how the servers are configured, and how the overall network architecture may need to change to accommodate load balancing appliances.

The following is a five step guide and introduction to the process of implementing application and server load balancing.

The good news is that deploying a load balancer needn't be perplexing or difficult. In fact, installing a Coyote Point Equalizer™ load balancer into an existing web server infrastructure can easily be done with minimal changes to your existing configuration. This document outlines how a fairly common web server installation can be outfitted with an Equalizer to provide load balancing with minimal changes to your network architecture using a simple “drop-in” deployment strategy. And best of all, **you don't need to be a networking guru to install an Equalizer.**

On the cover of this document is a diagram illustrating the two major network configuration scenarios: dual network and single network. In this paper, we'll show you how to set up a single network, drop-in configuration in five easy steps. The five steps are:

1. Preparation
2. Configure Equalizers on the network
3. Configure Virtual Clusters
4. Configure Server Gateways
5. Changeover DNS

If you understand networking from a server perspective, then you've got the knowledge necessary to drop a Coyote Point Equalizer into a network and configure a fully load balanced, even redundant load balanced, environment. The following section explains how Equalizer can easily fit into your existing network.

The Network

The Coyote Point Equalizer series of load balancers are flexible, and can be implemented to provide non-stop availability and higher application performance in a wide variety of different network configurations, depending on your infrastructure needs. The most common network implementation, known as single-network configuration, Coyote Point load balancers provide a simple “drop-in” implementation.

The drop-in implementation has several advantages that make Equalizer particularly simple to implement. These advantages include:

- No need for additional subnets or physical networks
- The servers do not need to have their IP addresses changed
- Only one small change is needed on the servers to fully implement load balancing
- Works without changing existing network infrastructure
- Cut over is seamless -- does not interrupt site traffic, even if connections go to the old IP



To show how this drop-in can easily be added into an existing infrastructure, take as an example a very common installation illustrated here in Diagram 1

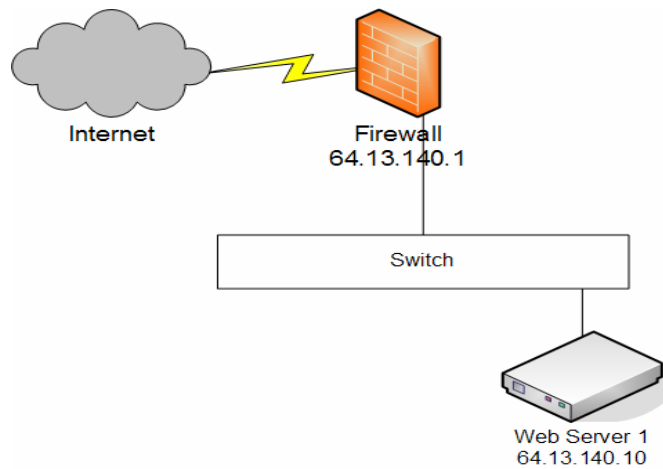


Diagram 1: Simple Web Serving Scenario

This is a very simple single-web server configuration: The domain name points in DNS to the web server with the IP address 64.13.140.10. The firewall, located at 64.13.140.1, acts as the default gateway for the single web server.

If your business or customers depend on this web site, this configuration trades off simplicity for high vulnerability to service outages. There is no redundancy in case the web server were to suffer a failure, and expanding the capacity would require either upgrading memory/processors for the system, or replacing it entirely with a more powerful system.

With a drop-in configuration, a single, or redundant, load balancing Equalizer(s) sits on a single network, on one subnet, the same network and subnet that the web servers sit on. You don't need to add additional networks, change the IP addresses of your servers, or add any extra networking gear. The servers will still be accessible the same way they were before a load balancer was implemented.



Pictured below are two Coyote Point Equalizer load balancers added to the network providing a redundant load balancing configuration (Diagram 2). In this example we can add two more web servers, bringing the number of web servers to three and dramatically increasing the performance and availability of the web services applications.

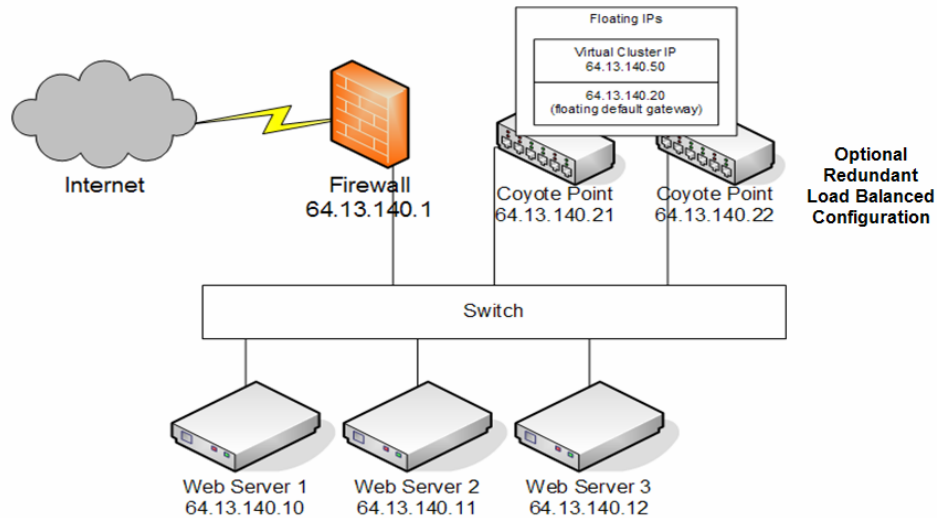


Diagram 2: New Load Balanced Environment

While the servers can still be individually accessed, all web traffic will be directed to a separate IP address, called a Virtual Cluster. The Virtual Cluster will accept traffic and distribute it to the available web servers. In the case of a redundant load balancing configuration as shown, if the active Equalizer were to go off-line, the Virtual Cluster IP would automatically switch to the standby Equalizer.

Beyond the load balancing the Equalizer has additional capabilities that ensure the highest application availability. By performing health checks on the three servers, the Equalizer ensures that they are capable of serving up traffic. If one web server goes down, Equalizer stops sending traffic to that server and routes traffic to the remaining active servers. Once the server comes back up, Equalizer automatically resumes sending traffic to it.

Each Equalizer has an individual IP address, which is used for management. In addition, both Equalizers share a floating IP address (the failover address). Like the Virtual Cluster address, the floating address exists only on the active Equalizer. This floating IP also serves as the default gateway for the web servers.

While the servers change their default gateway to the floating IP address, both Equalizers have their default gateway set to the firewall; the effect is that the outbound gateway for the entire configuration is still the firewall. The web servers have inbound and outbound Internet access just as they did before Equalizer was installed, and are limited only by the firewall's security profile. Equalizers need to be the default gateway to ensure traffic passes through Equalizer in both inbound and outbound directions.

Inbound traffic will be changed to use the Virtual Cluster IP Address instead of the server address previously used (64.13.140.10). This change will be made in DNS once installation is complete so that there is no interruption of service.

With this explanation of the network, the actual process of installation can begin.



Step 1: Preparation

There are a few minor preparatory steps that you'll want to take before implementation to ensure a successful deployment. First, you'll need two additional IP addresses on your network if you're running a single Equalizer in stand-alone mode or four additional IP addresses if you're running Equalizers in redundant mode.

You'll also want to change the TTL (Time To Live) on your domain name (or names) to zero. This will make the cut-over from the single web server to the load balancer quicker. Your DNS provider (whoever shows up in a WHOIS for the domain, typically your ISP) should be able to accommodate this request.

Step 2: Adding Equalizer to the network

Adding an Equalizer to your network is very simple and begins with the physical connection. In the example single network configuration from Diagram 2, Equalizer, the application servers, and the upstream router are all on the same subnet, so we'll only be using Equalizer's configured Internal interface ports.

All Equalizer models are equipped with two network interfaces and gigabit switch ports on the front panel. The ports on Equalizer E350GX, E450GX, and E650GX models can be configured for either interface. By default, ports 1 and 2 are configured for the External network interface, while ports 2 and above are configured for the Internal interface. The factory installed port configuration for an E350GX is shown below:

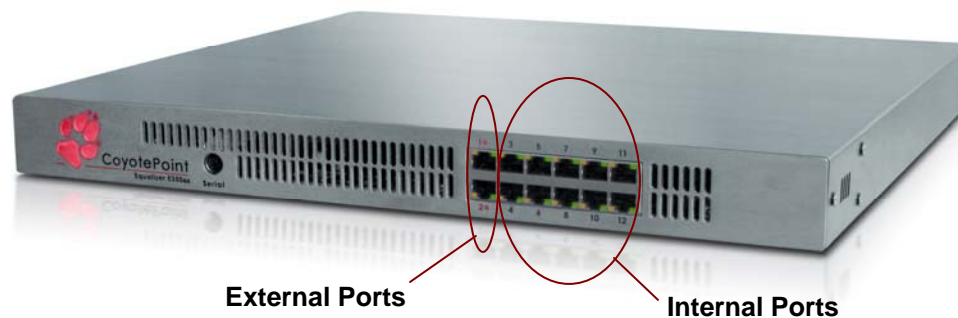


Diagram 3: Port Layout of an Equalizer E350GX

The ports on Equalizer E450GX and E650GX models are similarly configured.

Equalizer E250GX models have two front panel gigabit ports, one for the Internal interface and one for the External interface. These ports cannot be re-assigned to the other interface.

Connect one of the Internal interface ports (or the single Internal port on an E250GX) to the switch or hub infrastructure for the upstream router and the servers. On E350GX and higher models, you can connect your servers directly to Equalizer's other unused gigabit Internal ports.

Power up each Equalizer and give them IP addresses and hostnames. The IP addresses for eq1 and eq2 are assigned to the internal network cards on the Equalizers using the character line



interface through Equalizer's serial port. Initial configuration of the box is done using the included serial cable and a serial terminal; or, a terminal emulator application (such as HyperTerminal, which is included with most versions of Microsoft Windows) can also be used. The installation manual describes other freely available terminal emulator options, and contains the information needed to make a connection to Equalizer.

Below is the IP scheme for this particular configuration. Again, three IP addresses (Table 1) will be needed in a redundant scenario.

<i>Equalizer Hostname</i>	<i>IP Address</i>
eq1	64.13.140.21
eq2	64.13.140.22
Floating	64.13.140.20

Table 1: IP Addressing for Equalizers

By simply assigning IP addresses to the Equalizers, you will be able to reach them (using the **ping** command, for example) from other systems on the same subnet. You can now finish configuring Equalizer via a web browser over secure HTTP. If Equalizer's IP is 64.13.140.21, the URL would be <https://64.13.140.21>. This brings up Equalizer's Administration Interface. The Administrative Interface is designed to work with any Java-enabled browser. FireFox 2, Internet Explorer 6, or Internet Explorer 7 are recommended.

Step 3: Configure a Cluster

The next step is to configure an Equalizer Cluster. The Cluster is what accepts connections on behalf of the servers. Eventually, the site's DNS entry will point to the Virtual Cluster IP.

In the left frame object tree of the Administration Interface, right-click on **Equalizer** and select **Add Cluster** from the popup menu (Diagram 4):

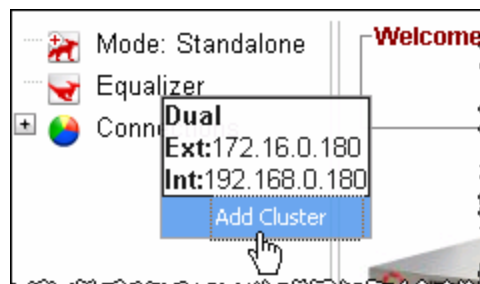


Diagram 4: Add Cluster Command

Clicking on this starts the Add Cluster Wizard. The Wizard prompts you for the information required to create the cluster, as shown in Diagram 5:



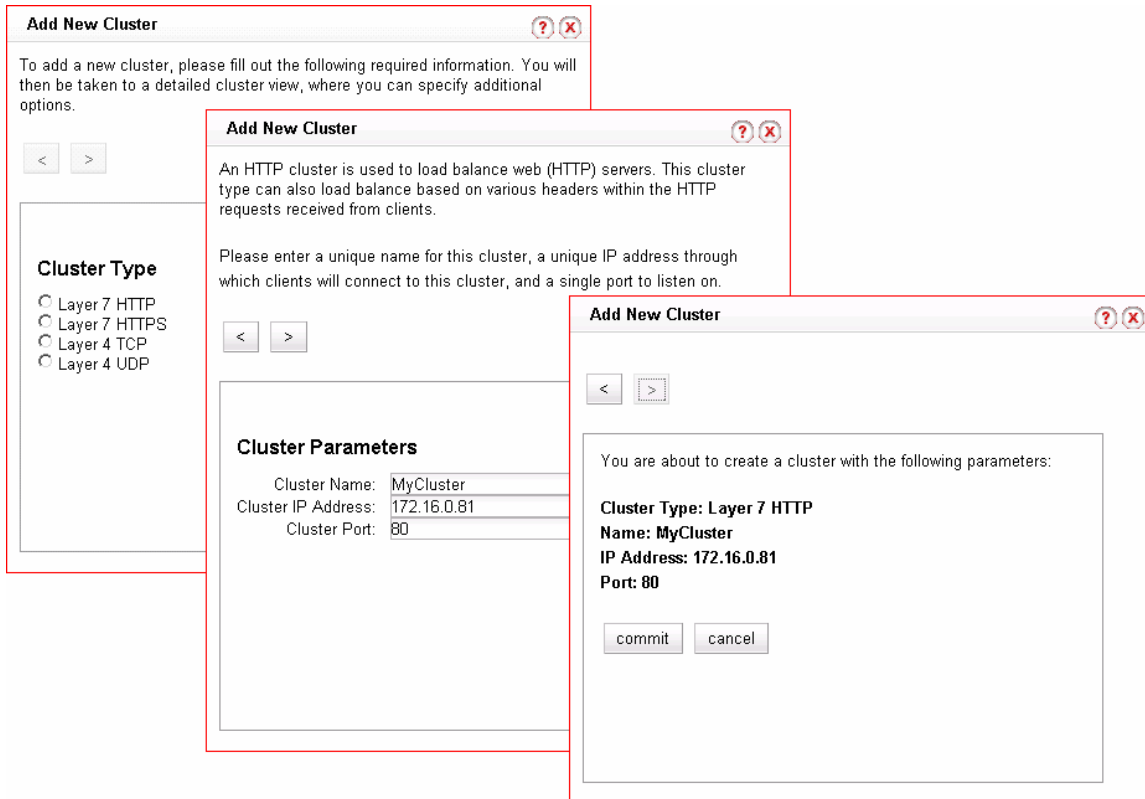


Diagram 5: Add Cluster Wizard

First, you select the type of cluster you want to create, then click the Next (>) button to enter the name, IP address, and port for the cluster. Click Next (>) and a final screen summarizes the information you have entered; click **commit** to create the cluster. In this example, the virtual cluster is given the name **MyCluster** and configured for the HTTP protocol on port 80 at the IP address shown. Once the cluster is created, Equalizer opens the cluster's **Configuration > Required** tab (Diagram 6):

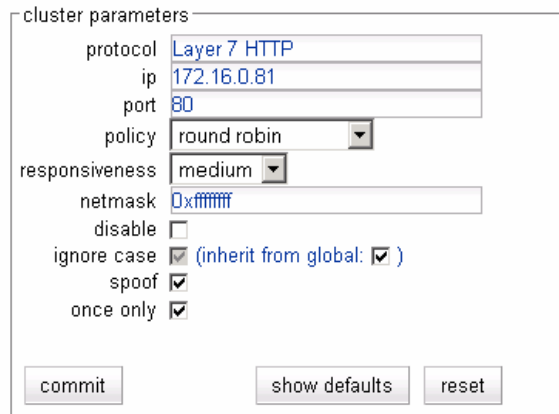


Diagram 6: Cluster Configuration > Required Tab



We'll leave the cluster parameters set to their default values, and discuss a few of the more important ones briefly before we configure the servers in the cluster.

The load balancing **policy** is the algorithm used by Equalizer to distribute incoming requests to a cluster's servers. The default is **round robin**, which distributes incoming requests to each server in the cluster one at a time, then loops back to the beginning of the list of servers. Other available load balancing policies include **static weight**, **adaptive**, **fastest response**, **least connections**, and **server agent**. You can also select the custom policy and set the sliders below to configure the load balancing behavior you desire.

The **spoof** flag, when enabled (the default), preserves the IP addresses of the client requests. When disabled, all the servers in the cluster see all inbound connections as coming from Equalizer's IP address.

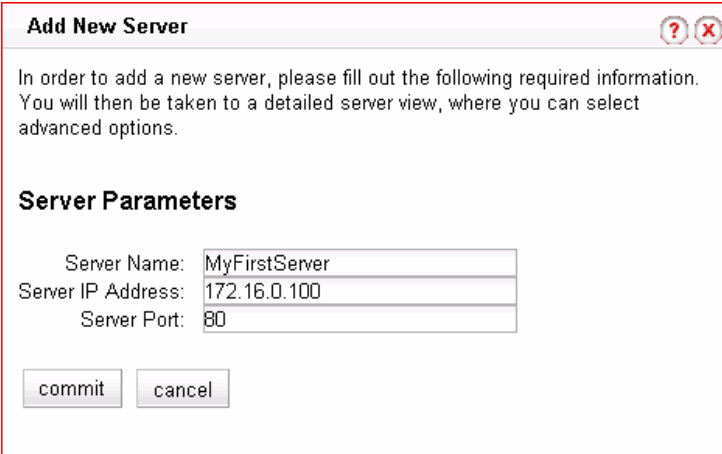
The **once only** flag, when enabled, improves performance by only examining the first set of headers in a connection. Most applications that require persistence or match rules will require that once only is disabled.

Another important option configured by default is **persist**, located on the **Persistence** tab. When a Virtual Cluster is configured, persistence through active cookies is setup by default. This will keep a client tied to a specific server for the duration of their session. This is typically a requirement for interactive web sites. Even if it's not a requirement, persistence is generally benign, and will not adversely affect a site that doesn't require persistence.

Next, we'll add some servers to our new cluster.

Add Servers to the Cluster

Now that the virtual cluster is configured, it's time to add the web servers to the cluster. Right-click on the name of the new cluster in the left frame object tree and select **Add Server** from the popup menu. A popup dialog is displayed that prompts you for the settings required to create the new server (Diagram 7):



Add New Server ? X

In order to add a new server, please fill out the following required information. You will then be taken to a detailed server view, where you can select advanced options.

Server Parameters

Server Name: MyFirstServer
 Server IP Address: 172.16.0.100
 Server Port: 80

commit cancel

Diagram 7: Add Server to Virtual Cluster

Give the server a name, add the IP address and the port, and click **commit** to add the server to the cluster. This opens the server's **Configuration > Required** tab (Diagram 8):



Server parameters	
ip	172.16.0.100
port	80
probe port	0
max connections	0
weight	100
hot spare	<input type="checkbox"/>
quiesce	<input type="checkbox"/>
dont probe	<input type="checkbox"/>
dont persist	<input type="checkbox"/>
<input type="button" value="commit"/> <input type="button" value="show defaults"/> <input type="button" value="reset"/>	

Diagram 8: Server Configuration Tab

Repeat this with all your web servers.

Step 4: Configure Default Gateways on Servers and Test

There is only one configuration change that must be made on the server systems, and that's the default gateway. In Microsoft Windows, this setting can be found under the TCP/IP settings of a given network connection's control panel (Diagram 9).

<input type="radio"/>	Obtain an IP address automatically
<input checked="" type="radio"/>	Use the following IP address:
IP address:	64 . 13 . 140 . 10
Subnet mask:	255 . 255 . 255 . 0
Default gateway:	64 . 13 . 140 . 20

Diagram 9: Default Gateway Settings





When using a single network configuration (where all servers and Equalizer are on the same subnet) or when using the **spoof** option, the default gateway for all servers that are load balanced by Equalizer needs to be Equalizer's floating gateway IP address. All of your other settings, including the web server configuration, can remain unchanged. Equalizer will work seamlessly with this configuration.

Test Your Cluster

Now that your servers are configured, return to the Equalizer Administration Interface, click **Equalizer** in the left frame and then open the **Status** tab. A screen like the following is displayed:



The **Cluster Summary** table shows the server status and basic server statistics for all clusters. Click on the cluster name in the table to display the information for that cluster.

Server Status:  Up  Down  Quiesced  Hot Spare






MyCluster				
Servers	Status	Weight	Processed	Active
MyFirstServer		100	0	0
Match rules		Processed		
Default		0		

Diagram 10: Cluster Summary Screen

Using one or more clients that are not on the same subnet as your web servers, open a browser and type the Virtual Cluster IP address into the location bar and you'll see the web site coming up on each client. You should see the numbers in the screen above changing as you hit the website from different client machines, as shown below (note that the numbers on the screen are refreshed automatically about every 30 seconds):

The **Cluster Summary** table shows the server status and basic server statistics for all clusters. Click on the cluster name in the table to display the information for that cluster.

Server Status:  Up  Down  Quiesced  Hot Spare


MyCluster				
Servers	Status	Weight	Processed	Active
MyFirstServer		100	23134	24
Match rules		Processed		
Default		23134		

Diagram 11: Cluster Summary Screen during testing

(**Note:** if you don't have multiple clients available, then clear your browser's cookies after every web site access so you in effect open a new connection to the cluster every time you access it from the same client.)



Step 5: DNS Cut-over

By completing step one through four, all the setup and configuration are complete and tested, the site is now ready to switch over with a DNS cut-over.

Contact your DNS provider (again, typically your ISP) and have them switch the DNS for your site from the old IP (directly accessing the first web server) to the IP address of the Virtual Cluster on Equalizer. Also have them set the TTL for your DNS entry to 0.

Because DNS can (and will) be cached despite TTL, the effect will not be immediate for all clients. However, since both the old IP address and the Virtual Cluster are both serving up the website, this will not affect your overall site availability. Over a period of 24-48 hours, traffic will migrate to the Virtual Cluster without service interruption.

Drop In Simplicity – Added Capabilities

This solution can expand to many more web servers, Virtual Clusters, and serve multiple websites while providing advanced traffic management features such as persistence, health checking, and failover.

In addition, the Coyote Point Envoy solution is a software upgrade that adds global load balancing capabilities to provide failover and load distribution across multiple geographical locations.

With Coyote Point's Equalizer drop-in solution, load balancing needn't be difficult to implement or disruptive to your existing infrastructure. **With this easy to understand, easy to implement solution, there is no reason not to add the benefits of load balancing to your infrastructure today.**

Step 6: Equalizer Options

Here's a quick look at Equalizer options that further improve Equalizer's ability to deliver cutting edge application traffic management solutions:

Hardware SSL Acceleration

Equalizer E450 and E650 models contain the Xcel II SSL Accelerator Card, which not only dramatically improves Equalizer's ability to process HTTPS transactions, but also offloads processor-intensive SSL operations to the Xcel's on-card processor. Xcel lets Equalizer apply all its processing power to do what it does best: manage application traffic.

Hardware Data Compression

Equalizer E650 models are configured with the Express GZIP Compression Card, which provides data compression for server responses. This reduces the size of the payloads Equalizer sends back to clients over the network, making the most of expensive resources like bandwidth. Since all compression is performed on the Express card, compression operations are fast and don't interfere with normal Equalizer operation.

Geographic Server Load Balancing (GSLB)

Envoy is Equalizer's GSLB solution, providing scalable application load balancing and acceleration amongst geographically distributed data centers. Envoy empowers you to take full advantage of distributed application infrastructures, mitigates the risk of data center outages, provides disaster recovery for application servers, reduces bandwidth costs, and accelerates end user response time by directing user requests to the 'closest' data center.



Step 7: Contact Us!

Discuss with us the world-class benefits that Coyote Point's affordable, non-stop, application acceleration and load balancing technology can bring to your small to medium size business. Please see our [Contact Page](#) on the web.

