

HARDWARE

Coyote Point Offers Application Balancing for Virtual Servers

Stressed budgets and the high demand for Web applications have made efficiency in the data center priority number one, creating the need for inexpensive load balancing, application balancing and SSL-acceleration products. Does Coyote Point have what it takes to succeed in the enterprise? FRANK J. OHLHORST

Thanks to the low cost of open-source solutions and the falling prices of hardware, network managers are finding it easier than ever to build out networks and grow the capabilities of the data center. Network managers quickly can meet the increasing demands of Web services, remote access, VPNs and many other services by provisioning new inexpensive servers with open-source software, which has fueled the exponential growth of server clusters and Web applications. However, many network managers are discovering that just throwing additional servers into the mix is an inefficient way to balance loads across server farms. High server demand can create a cascade effect across the servers, with a services request sent to the primary server and moving on to the next server only after the primary server has become saturated.

What's more, server virtualization solutions magnify the problem. Virtualization makes it even easier and quicker to build out multiserver solutions, yet virtualized servers still rely on that same round-robin approach to meet high demand. That proves to be very inefficient and a waste of processor cycles, and it negates most energy savings offered by virtualization.

Many data-center administrators have turned to third-party products to help mitigate those load-related inefficiencies, creating a healthy market of load-balancing and traffic-acceleration solutions. Numerous products, ranging from open-source software to hardware appliances that cost tens of thousands of dollars, are all fighting for market share and promise to be the best way to manage loads across servers and sites. However, traffic-balancing products are not created equally and selecting the correct load balancer can be fraught with uncertainty.

Coyote Point Systems entered the fray more than 15 years ago, with the ideology that network traffic management was the key to maximizing bandwidth and services availability to endpoints. Contrary to what other vendors were doing in the 1990s, Coyote Point chose to go the route of building traffic management capabilities into an appliance. The goal was to replace bulky software solutions with an easy-to-manage device.

Although Coyote Point was a pioneer in the world of traffic-shaping and load-balancing appliances, other companies, such as F5 Networks, Barracuda Networks and Cisco also have focused on providing hardware-based load-balancing

Equalizer Model	E250Gx	E350Gx	E450Gx	E650Gx
Ethernet Connectivity	2 GigE (1000base-T)	12 GigE (1000base-T)	12 GigE (1000base-T)	20 GigE (1000base-T) + 2 SFP
Throughput	600 Mbps (L4)	850 Mbps	1.0 Gbps	1.8 Gbps
Layer 4 Load Balancing	✓	✓	✓	✓
Layer 7 Load Balancing	✓	✓	✓	✓
Layer 7 Persistence	✓	✓	✓	✓
Match Rules (content switching)	N/A	✓	✓	✓
SSL/HTTPS Performance	180 tp/s (Software)	500 tp/s (Software)	8,500 tp/s (Xcel-II hardware)	14,000 tp/s (Xcel-II hardware)
SSL/HTTPS Throughput	100 Mb/s	400 Mb/s	800 Mb/s	825 Mb/s
HTTP GZIP Compression	N/A	N/A	N/A	Express Hardware 1 Gb/s Throughput
Envoy (GLSB) Global Server Load Balancing	N/A	Optional	Optional	✓ Envoy GSLB Included
Equalizer VLB	VLB Basic Included	VLB Basic Included	VLB Basic Included	VLB Advanced Included
Supported Clusters	64	Unlimited	Unlimited	Unlimited
Real Servers per Cluster	16	32	128	512
Layer 7 Performance	8,500 tp/s	50,000 tp/s	75,000 tp/s	110,000 tp/s

Figure 1. Feature Chart

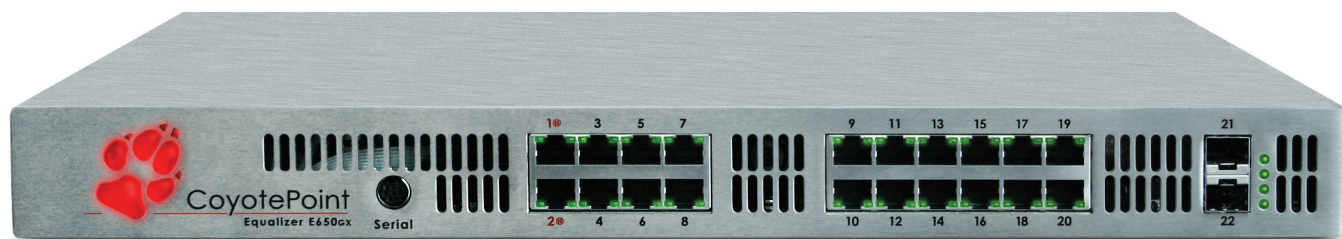


Figure 2. Coyote Point E650GX

solutions, creating a crowded field of contenders, where each vendor is looking to tout specialized capabilities to become the appliance of choice.

Coyote Point has chosen to up the ante with the launch of a new series of load-balancing appliances, which are virtual server-aware. What's more, the company is looking to shift the focus from layer 4 load balancing to application load balancing, where the appliance is aware of payload as well as raw traffic. Coyote Point's application load-balancing appliances can shape traffic and efficiently distribute loads across multiple servers, even if those servers are virtual in nature. The growth of virtualization solutions in the data center has made it critical for traffic-shaping and load-balancing appliances to integrate with virtual server solutions.

Coyote Point offers four different appliances. Those four appliances differ based upon design traffic load and sub-features, yet all share the same management console and basic feature set.

Application Load Balancing— the Coyote Point Way

I tested the E650GX (V8.6) load-balancing appliance for ease of use, feature set, performance and suitability to task. I found that the device is very simple to install; the physical portion of the installation consists of plugging in the device and routing the appropriate Ethernet cables to the unit. The E650GX is Coyote Point's top-of-the-line appliance and sports 22 Gigabit Ethernet interfaces for connecting server clusters.

I spent more time figuring out my cabling than I did configuring the device. Making sure your cabling goes to the appropriate servers is one of the most important steps for deploying a Coyote Point appliance. You have to be

certain that you are plugging your server farm in to a load-balancing port on the device. In complex environments, it is easy to forget that a particular network segment is plugged in to a different router or switch from what you originally thought. However, on smaller networks, you simply can plug in the connection from your firewall to the external port on the E650GX and then plug each segment of the LAN in to the internal ports on the device. All ports on the E650GX are Gigabit Ethernet and support full duplex operation. That means it is very unlikely the device will introduce any bottlenecks into the LAN or WAN connections, and none were detected during performance testing.

I found the rest of the setup process plug-and-play easy. After I connected all the cables, I was able to access the E650GX's management console using a Web browser. The management console is based on AJAX technologies, creating a rich user interface, which is easy to navigate and sports context-sensitive help. The interface was designed using the Dojo toolkit, a JavaScript development toolset, which helps give the management interface a professional look and feel.

The E650GX works pretty well right out of the box. All you need to do to start load balancing is to set up some basic parameters. One of the first steps to complete is the definition of your server clusters. For example, if you have nine servers running a Web application, you would plug each of those servers into an internal port on the E650GX appliance. The next step consists of defining your server clusters. For example, you may want to divide those nine servers into three clusters. That proves very easy to do and easy to modify, if you need to change anything. All clusters are defined logically, allowing a great deal of flexibility.

I found that the E650GX offers many options when it comes to load balancing. You can use "Match Rules and Custom Load Balancing Policies" to build policies based upon layer 4 requests, layer 7 requests or even create custom policies using Boolean logic. The layer 4 policies offer basic load-balancing capabilities, based on parameters, such as least connections, fastest response, adaptive and round-robin, as well as an agent-based algorithm that is accurate if the agent is run on each server. Layer 7 policies actually look at the content of the traffic to determine how to load balance it. For example, certain protocols or applications can be used to trigger a load-balancing policy to route traffic to a particular cluster. Policies based on Boolean logic take into account particular requests, based on a series of administrator-defined events. Those policies can be used to reroute traffic if a server fails to respond (failover routing) or to route based on a schedule.

Once you have defined your clusters, you then can define rules to handle traffic flow and load-balancing decisions. Coyote Point calls those definitions Smart Events. The rules are based on a number of parameters, such as server load, traffic type, server weighting and connection persistence. The underlying technology that allows the E650GX to make traffic-routing decisions is very complex. However, the E650GX does an excellent job of hiding that complexity by using rule-creation wizards and a commonsense procedural layout to make rule definition very easy. That ease of configuration is rarely found in software-only load-balancing solutions and allows even newbie network administrators to set up basic load balancing with the E650GX.

I found that one of the most

impressive features of the E650GX was the unit's ability to work with VMware's vSphere products. That brings application load balancing and traffic shaping to the world of virtual servers. Coyote Point has built support for VMware's APIs, allowing the E650GX to judge the load on a virtual server, then route requests based on virtual loads and administrator-defined load-balancing policies. What's more, Coyote Point has included support for IPMI-capable servers. The Intelligent Platform Management Interface (IPMI) is a specification that allows third-party products to power on and power off servers, as well as remotely execute

Coyote Point's application load-balancing appliances can shape traffic and efficiently distribute loads across multiple servers, even if those servers are virtual in nature.

other commands. Simply put, you can define a policy that automatically turns on a server when traffic loads hit a certain level, and then shut off that server once traffic load drops.

I also found it very easy to segment LANs using the product's VLAN capabilities. Administrators define VLANs based on IP address segments, and the unit's built-in routing capabilities keep traffic isolated on a VLAN for local requests. That can help reduce latency and speed up requests by keeping the appropriate

traffic on the same logical segment.

Ease of use permeates the interface, making it simple to set up many, if not all of the unit's capabilities. You also will find that ease of use present in the device's dashboards and reporting menus. The dashboards offer a quick snapshot of how the device is performing and what traffic is flowing across the device. Reports offer a historical reference of many monitored parameters and can be useful for fine-tuning the unit.

Although the E650GX's primary focus is on application load balancing, the unit also includes other features that help speed up network access

and reduce latency. Those features include SSL acceleration, HTTP compression and global/geographic load balancing. SSL acceleration helps reduce the latency found in HTTPS requests by offloading the packet encryption on to the device. HTTP compression helps reduce latency by compressing and optimizing HTTP requests, while global/geographic load balancing can be used to balance traffic across geographical clusters, placing requests on servers that are

closest to the user, as far as latency and bandwidth are concerned.

Administrators supporting e-commerce solutions will appreciate the E650GX's ability to deliver persistent connections. E-commerce transactions rely on a reliable connection between the client PC and the server providing the transaction—if either endpoint loses track of each other or is routed incorrectly, the e-commerce transaction will fail. The E650GX solves that problem by creating a persistent connection between the client PC and the server using cookies, which are inserted into the HTTP returned to the client. That ensures the client will return to the same server in the cluster.

The E650GX supports an active/passive failover model for sites that need guaranteed uptime. Failover works by transferring the Equalizer configuration to a backup device (which can be a lower-end model in the Coyote Point family), so that persistent client/server connections are maintained even when the primary unit fails.

Coyote Point has re-invented the idea of load balancing by shifting traffic shaping from basic layer 4 algorithms to layer 7, application-aware calculations. That approach has created a new market segment called application traffic shaping. Coyote Point also bundles in other advanced capabilities, ranging from SSL acceleration to VLAN definition to VMware vSphere support, making the device a complete traffic-acceleration solution. Coyote Point is very adept at providing an acceleration solution for most any server environment that can benefit from clustering and traffic management. The top-of-the-line E650GX has an MSRP of \$14,395 and comes with one year of support included. Although \$15K may seem like a big chunk of change, Coyote Point's price is less than half of what some larger competitors charge. ■

Performance was tested using a Spirent Avalanche 2900 configured to generate traffic load on the E650GX. The goal of the performance testing was to rate layer 7 HTTP transactions per second, layer 7 HTTPS throughput, layer 4 maximum concurrent connections and layer 4 connections per second. Those metrics are a good indicator of the processing power of the device, as well as the overall capacity of the device.

- Layer 4 connections per second: 160,000 CPS
- Layer 4 maximum concurrent connections: 17.5 million
- Layer 7 HTTP transactions per second: 110,000 TPS
- Layer 7 HTTP throughput: 1,300 Mbps
- Layer 7 HTTPS transactions per second: 14,000 TPS
- Layer 7 HTTPS throughput: 825 Mbps

Frank J. Ohlhorst is a freelance technology journalist, professional speaker and technology business consultant who covers several topics for many major publications. Frank currently writes *Virtualization and Application development* articles for *Tech Target*, *Operating Systems and Security* articles for *Computer World*, *Technology How-Tos* for *Ars Technica* and Channel-focused product reviews for *ChannelTechCenter.com*. Frank also has been the Executive Technology Editor for *eWeek* and the Lab Director of CRN's Test Center.