W H I T E P A P E R

Implementing High Availability Layer 4 Services Using VRRP and VRRP Extensions

- **VRRP OPERATION**
- **ALTEON EXTENSIONS TO VRRP**
- **A TYPICAL CONFIGURATION**
- **DEPLOYMENT CONSIDERATIONS**

Alteon WebSystems, Inc.

50 Great Oaks Boulevard

San Jose, California 95119

408-360-5500

408-360-5501 fax

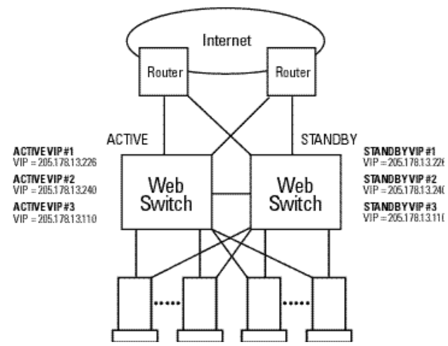
http://www.alteonwebsystems.com

94011.25/02-01

With service availability becoming a major concern on the Internet, service providers of all kinds are increasingly deploying Internet traffic control devices such as Web switches in redundant configurations. Traditionally, these configurations have been hot standby configurations where one switch is active and the other is in a standby mode. An example of the traditional hot standby configuration is shown in figure 1.

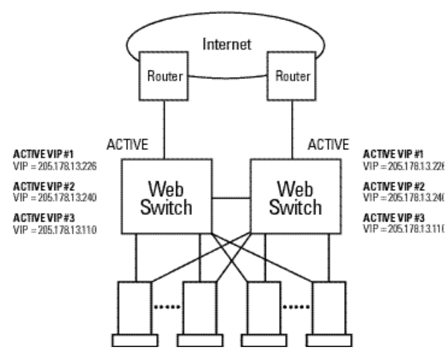
While hot standby configurations increase site availability by removing single points of failure, service providers increasingly view them as an inefficient use of investment dollars because one perfectly functional Web switch sits by idly until a failure calls it into action. Service providers now demand that vendors' equipment support redundant configurations where all devices can process traffic when they are healthy. This increases site throughput and capacity and decreases user response times when no device has failed.

Figure 1.
Hot Standby Redundancy



Alteon WebSystems Web switches, including the Alteon 180e, ACEdirector 3 and Alteon 700 Series, support active-active redundancy configurations. In active-active configurations, as shown in figure 2, both switches can process traffic for the same service at the same time. That is, both switches can be active simultaneously for a given IP routing interface or load balancing VIP address.

Figure 2.
Active/Active Redundancy



Alteon WebSystems' active-active redundancy configuration is based on the Virtual Router Redundancy Protocol (VRRP) described in RFC 2338. Alteon WebSystems has developed extensions to VRRP that allow it to be used to support layer 4 switching services such as server load balancing and to support active operation of interfaces (at layer 3) and services (at layer 4) across multiple switches at the same time.

This paper describes Alteon WebSystems' active-active redundancy capability. To give the reader the background necessary to understand active-active operation, the paper begins by reviewing VRRP operation and the Alteon-specific extensions to VRRP.

VRRP OPERATION

VRRP enables redundant router configurations within a LAN. It provides alternate router paths for a host. As a result, VRRP can eliminate single points of failure within a network.

The router associated with a given alternate path supported by VRRP uses the same IP address and MAC address as the routers for other paths. As a result, the host's gateway information does not change, no matter what path is used. Because of this, VRRP-based redundancy significantly reduces administrative overhead when compared to redundancy schemes that require hosts to be configured with multiple default gateways.

Terminology

Each physical router running VRRP is known as a VRRP Router. Two or more VRRP Routers can be configured to form a Virtual Interface Router. (Note that RFC 2338 calls this entity a Virtual Router. This paper uses the term Virtual Interface Router to refer to what RFC 2338 calls a Virtual Router to distinguish this type of entity from a Virtual Server Router, which is described in the section on Alteon WebSystems' extensions to VRRP. When the term Virtual Router is used in the paper, the described concept applies to both Virtual Interface Routers and Virtual Server Routers.) Each VRRP Router may participate in one or more Virtual Interface Routers.

A Virtual Interface Router acts as a default or next hop gateway for hosts on a LAN. Each Virtual Interface Router consists of a user-configured Virtual Router Identifier (VRID) and an IP address.

The VRID is used to build the Virtual Router MAC Address. The five highest order octets of the Virtual Router MAC Address are the standard MAC prefix (00-00-5E-00-01) defined in RFC 2338. The VRID is used to form the lowest order octet.

One, but not more than one, of the VRRP Routers in a Virtual Interface Router may be configured as the IP Address Owner. This router has the Virtual Interface Router's IP address as its real interface address. This router, when up, responds to packets addressed to the Virtual Interface Router's IP address for ICMP pings, TCP connections, etc.

There is no requirement for any VRRP Router to be the IP Address Owner. Virtual Routers may be implemented without an IP Address Owner.

For the purposes of this paper, VRRP Routers that are not the IP Address Owner are called Renters. While not part of RFC 2338, the Renter designation is a useful one.

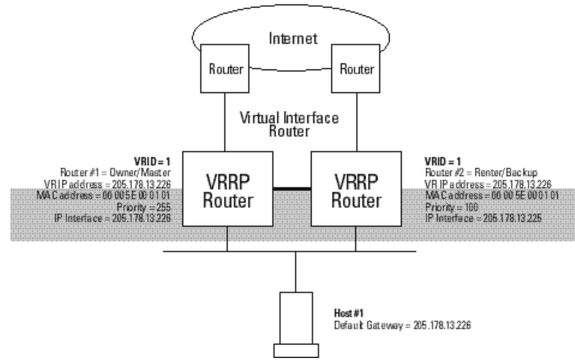
Within each Virtual Router, one of the VRRP routers is selected to be the Virtual Router Master. See the section "Determining which VRRP Router is the Master" for an explanation of how the selection process. Note that if the IP Address Owner is available, then it will always become the Master.

The Virtual Router Master forwards packets sent to the Virtual Interface Router. It also responds to ARP requests the Virtual Interface Router's IP address. Finally, it sends out periodic advertisements to let other VRRP Routers know it is alive and its priority (explained below).

Within a Virtual Router, the VRRP routers not selected to be the Master are known as Virtual Router Backups. Should the Virtual Router Master fail, one of the Virtual Router Backups becomes the Master and assumes its responsibilities.

The above points are illustrated in figure 3, Virtual Interface Router Example. The Web switches in the diagram have been configured as VRRP Routers. They form a Virtual Interface Router.

Figure 3.
Virtual Interface Router Example



The switch on the left has its real interface configured with the IP address of the Virtual Interface Router and is therefore the IP Address Owner. As a result, it is also the Virtual Router Master.

The switch on the right of the figure is a Virtual Router Backup. Its real interface is configured with an IP address that is on the same subnet as that of the Virtual Interface Router but that is not the IP address of the Virtual Interface Router. As a result, it is a Renter and is a Virtual Router Backup.

The Virtual Interface Router has been assigned a VRID = 1. Therefore, both of the VRRP Routers have a MAC address = 00-00-5E-00-01-01.

Operation

The host shown in figure 3 is configured with the Virtual Interface Router's IP address as its default gateway. The Master forwards packets destined to remote subnets and responds to ARP requests. Since, in this example, the Master is also the Virtual Interface Router's IP Address Owner, it also responds to ICMP ping requests and IP datagrams destined for the Virtual Interface Router's IP address. The Backup does not forward any traffic on behalf of the Virtual Interface Router, nor does it respond to ARP requests.

If the Owner is not available, the Backup becomes the Master and takes over responsibility for packet forwarding and responding to ARP requests. However, since this switch is not the Owner, it does not have a real interface configured with the Virtual Interface Router's IP address. Therefore, it does not respond to ICMP ping requests and IP datagrams destined to that address.

Determining which VRRP Router is the Master

Each VRRP Router that is a Renter is configured with a priority between 1–254. According to the VRRP standard, an Owner has a priority = 255.

A bidding process determines which VRRP Router is the Master. The VRRP Router with the highest priority becomes the Master. Since Owners have a priority higher than the range permitted for Renters, the IP Address Owner, if any, is always the Master for the Virtual Interface Router as long as it is available.

The Master periodically sends out advertisements to an IP multicast address. As long as the Backups receive these advertisements, they remain in the backup state. If a Backup does not receive an advertisement for three advertisement intervals, it kicks off a bidding process to determine which VRRP router has the highest priority. That VRRP router takes over as Master.

If, at any time, a Backup determines that it has higher priority than the current Master does, it can preempt the Master, unless it is configured not to do so. In preemption, the Backup begins to send its own Advertisements. The current Master will see that the Backup has higher priority and stop functioning as the Master. The Backup will then see that the Master has stopped sending advertisements and assume the role of Master.

ALTEON EXTENSIONS TO VRRP

Alteon WebSystems WebOS software implements important enhancements to VRRP. These enhancements are detailed in this section.

Virtual Server Routers

WebOS supports Virtual Server Routers. They extend the benefits of VRRP to VIP addresses used to perform server load balancing.

Virtual Server Routers operate for VIP addresses in much the same manner as Virtual Interface Routers operate for IP interfaces. A Master is negotiated via a bidding process during which information about each VRRP Router's priority is exchanged. Only the Master processes packets destined for the VIP address (unless sharing, explained below, is enabled) and responds to ARP requests. The Master sends periodic advertisements. If a Backup does not receive an advertisement within a specified period, it kicks off the bidding process to determine which VRRP Router takes over as Master. If, at any time, a Backup determines that it has higher priority than the current Master does, it can preempt the Master and become the Master itself, unless it is configured not to do so.

One difference between Virtual Server Routers and Virtual Interface Routers is that the concept of an IP Address Owner does not apply to Virtual Server Routers. All Virtual Server Routers are Renters. For a Virtual Server Router, the Master always responds to ICMP ping requests if sharing (explained below) is not enabled. If sharing is enabled, the switch where the ping request initially enters the Virtual Server Router responds.

Note that all Virtual Routers, whether Virtual Server Routers or Virtual Interface Routers operate independently from one another. Their priority assignments, advertisements and Master negotiations are separate. For example, when you configure a VRRP Router's priority in a Virtual Server Router, you are not affecting that VRRP Router's priority in any Virtual Interface Router or another other Virtual Server Router of which it is a part. However, because of the requirement that MAC addresses be unique on a LAN, VRIDs must be unique among all Virtual Routers whether Virtual Interface Routers or Virtual Server Routers.

Table 1 summarizes some of the differences between Virtual Interface Routers and Virtual Server Routers.

Virtual Router Type:	Virtual Interface Router	Virtual Server Router
Address:	IP address	Virtual IP (VIP) address
Owner/Renter:	The Owner, if any, has the Virtual Interface Router's IP address configured as its real interface address. There cannot be more the one Owner in a Virtual Interface Router. All VRRP Routers that do not have the Virtual Interface Router's IP address configured as their real interface address are Renters.	There is no Owner. All VRRP Routers in a Virtual Server Router are Renters.
Master/Backup:	The Master is determined by a bidding process that uses configured priorities. If there is an Owner and it is healthy, it always wins the bidding and becomes the Master.	The Master is determined by a bidding process that uses configured priorities.

Table 1: Differences Between Virtual Interface Routers and Virtual Server Routers

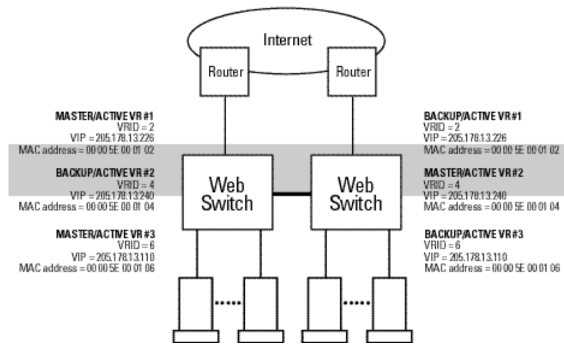
Sharing

WebOS supports sharing of interfaces at both layer 3 and layer 4, as shown in figure 4. With sharing, an IP interface or a VIP address can be active simultaneously on multiple switches, enabling active-active operation.

When sharing is used, incoming packets are processed by the switch on which they enter the Virtual Router. This is determined by external factors such as routing (e.g., OSPF) configuration and Spanning Tree configuration.

When sharing is enabled, the Master election process still occurs. Although the process does not affect which switch processes packets that must be routed or that are destined for the VIP address, it does determine which switch sends advertisements and responds to ARPs sent to the Virtual Router's IP address.

Figure 4.
Active/Active High Availability



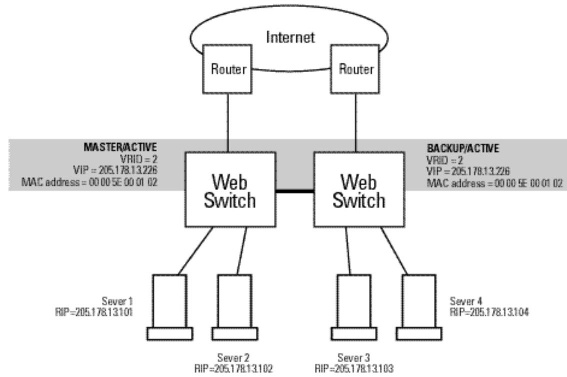
A TYPICAL CONFIGURATION

Alteon WebSystems Web switches offer many degrees of flexibility in implementing redundant configurations. This section discusses typical active-active Virtual Server Router configuration.

Figure 5 shows an example configuration where two Alteon 700 Series switches are used as VRRP Routers in an active-active configuration implementing a Virtual Server Router. Although this example shows only two switches, there is no limitation on the number of switches used in a redundant configuration. It's possible to implement an active-active configuration and perform load sharing between all of the VRRP-capable switches in a LAN.

In this configuration, when both switches are healthy, both load balance packets sent to the VIP. This results in higher capacity and performance than if the switches were used in a hot standby configuration.

Figure 5.
Active/Active High Availability



The switch on which a frame enters the Virtual Server Router is the one that processes that frame. The ingress switch is determined by external factors such as routing and Spanning Tree configuration.

Each VRRP-capable switch is autonomous. There is no requirement that the switches in a Virtual Router be identically configured. Different switch models with different numbers of ports and different enabled services may be used in a Virtual Router.

In this configuration, if a link between a switch and a server fails, the server will fail health checks and its backup attached to the other switch will be brought on line. If a link between a switch and its Internet router fails, the protocol used to distribute traffic between the routers, e.g., OSPF, will reroute traffic to the other router. Since all traffic now enters the Virtual Server Router on one switch, that switch will process all incoming connections. The same thing happens if an entire switch fails. If an entire switch fails and that switch is a Master, the Backup will detect this fact because it will stop receiving advertisements. In this case, the Backup will assume the Master's responsibility of responding to ARPs and issuing advertisements.

The behavior in the face of various failure conditions is summarized in Table 2.

Failure Condition	Behavior
Link between switch and server fails	Backup server on other switch added to load balancing mix
Server or application fails	Backup server on other switch added to load balancing mix
Link between switch and router fails	Routing protocol (e.g., OSPF) sends traffic to other router
Switch fails	Routing protocol (e.g., OSPF) sends traffic to router connected to other switch

Table 2: Summary of Resiliency Behavior

DEPLOYMENT CONSIDERATIONS

Users should consider the following questions when deciding how many switches to deploy in an active-active configuration:

1. How many switches are required to handle normal traffic? Consider both performance (sessions per second) and capacity (total simultaneous session) when answering.

2. How many switches are required to handle peak traffic? Again, consider both performance (sessions per second) and capacity (total simultaneous session) when answering.

3. Is performance degradation acceptable at peak load when all switches are running?

4. Is performance degradation acceptable at normal load if a switch fails?

5. Is performance degradation acceptable at peak load if a switch fails?

Obviously, a minimum of two switches is required for a redundant deployment. The answers to these questions determine if more than switches are required, if so, how many.

For example, say a network administrator determines that one switch is required to handle normal traffic and two are required to handle peak traffic. If the user decides that she wants a redundant configuration but it's not worth the investment needed to provide redundancy at peak load, then a two-switch configuration is sufficient. However, if she cannot tolerate performance degradation in the face of a failure at peak load, then three switches are required.

Or, let's say that another network administrator determines that normal traffic requires two switches and handling peak traffic, without a performance hit, requires three. If he wants to take the low-budget route and suffer slow response at peak loads and at normal loads when a switch fails, he can get away with two switches. If he determines that the site cannot afford slow response at peak loads when all switches are healthy and/or when a switch fails at normal loads, he'll need three switches. If he decides that he cannot tolerate performance degradation in the face of a failure at peak load, then four switches are required

The bottom line is that the number of switches that needs to be deployed is dictated by the site's requirements to handle normal and peak traffic and the cost of failures at various traffic levels.

CONCLUSIONS

Active-active redundant Web switch configurations increase application availability by removing single points of failure from networks. At the same time, when both switches are healthy, they increase performance and capacity by allowing two or more Web switches to support the same interface and service.

Alteon WebSystems' active-active redundancy capability is based on VRRP plus Alteon-specific extensions to VRRP. This capability is available in WebOS release 5.2 and later. It runs on the Alteon 180, ACEdirector and Alteon 700 series of products.