the Availability Digest

www.availabilitydigest.com

Redundant Load Balancing for High Availability

July 2013

A large data center can comprise hundreds or thousands of servers. These servers must not only be interconnected, but they also must connect to local users and to the Internet. Today's internal networks use Internet technology and are called Intranets. A company's Intranet must not represent a single point of failure if its systems are to provide high availability. There are many stories of critical systems that have been taken down by Intranet failures.[1]

The Need for High Availability in the Load Balancer

A data center is no good to anyone if it cannot be accessed by its users. To provide high capacity, performance, and availability, the customer-facing systems are often pools of servers. A request can be routed to any server in the pool so that the load can be shared by multiple servers. Furthermore, should a server fail, it is simply removed from the pool and the surviving servers continue to process requests.

Load balancers are used to distribute incoming traffic to the servers in the pool. However, a load balancer can represent a single point of failure in this critical portion of a company's Intranet. If the load balancer fails, the company's public services are down.

Load balancers from Loadbalancer.org (www.loadbalancer.org) solve this problem. Their load balancers are configured in active/passive pairs with instant failover so that customer-facing services are not interrupted should the active load balancer fail. The passive load balancer immediately takes over and continues service to the Internet.

The Need for Stateless Servers

Load balancing works best if the servers in the load-balanced pool are stateless. That is, there is no data that the servers must maintain between requests. For instance, a server pool that serves static web pages satisfies this requirement. In this case, a request may be routed to any server in the pool and be processed properly.

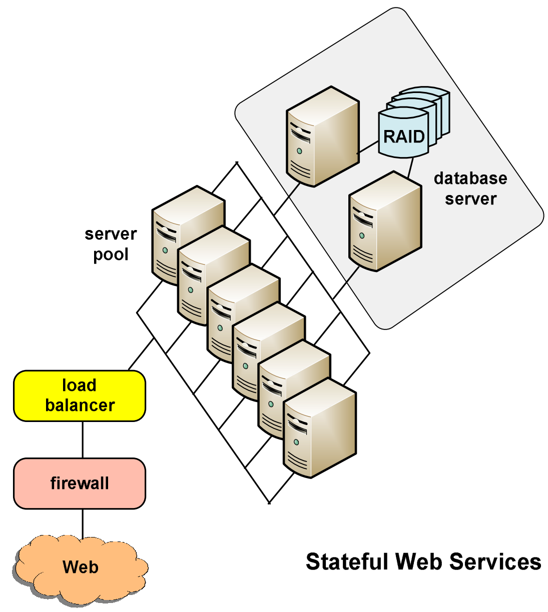
If servers must remember state, then failover to a backup load balancer is not transparent to the user. The user will in all likelihood lose his session and must re-logon. He might even lose his entire shopping cart and have to start over.

For stateful systems, it is important that session state data be maintained in a highly available database that is accessible to all servers rather than maintained in the servers themselves. For instance, the shared database may contain the shopping carts being filled by users. Any server can access the database to add to or modify the contents of a user's shopping cart. The shared database can also store

[1] Help! My Data Center is Down – Part 4: Intranet Outages, *Availability Digest*; January 2012. http://www.availabilitydigest.com/public_articles/0701/data_center_outages-intranet.pdf

session parameters so that current user sessions will be transferred to another server should a server fail without the users needing to re-logon to the application again.

A typical structure of a shared database is a pair of database servers, one acting as the active database server and the other acting as a backup database server. All servers in the pool route their requests to the active database server, which makes all updates to the database. Should it fail, the backup database server takes over and continues database services for the server pool.



This configuration requires that both database servers have access to the same database. The database itself must also be redundant so that it does not represent a single point of failure. This can be accomplished by using a single RAID database that will survive any single disk or controller failure. Alternatively, each database server can have its own disk subsystem with updates made to the primary disk replicated to the backup disk.

High Availability from Loadbalancer.org

Load balancers from Loadbalancer.org can be configured as a clustered master/slave pair to provide a highly available and resilient load balancing solution. The load balancing cluster is addressed via a virtual IP (VIP) address and converts incoming packets to the real IP address (RIP) of a selected server. The master and slave load balancers share a floating IP address. The network knows that the master controls the floating IP address, and all traffic will be sent to this address. If the master fails, the slave will take over the floating IP address and will seamlessly handle the load balancing for the application cluster.

All settings related to load balanced services are replicated from the master to the slave. This ensures that should the master node fail, the slave node is already configured to run the same services.

The pair communicates via a heartbeat to ensure that both the master and slave load balancers are performing properly. Should the master node fail, the slave immediately takes over all resources currently hosted on the shared floating IP address.

Supported Loadbalancer.org Configurations

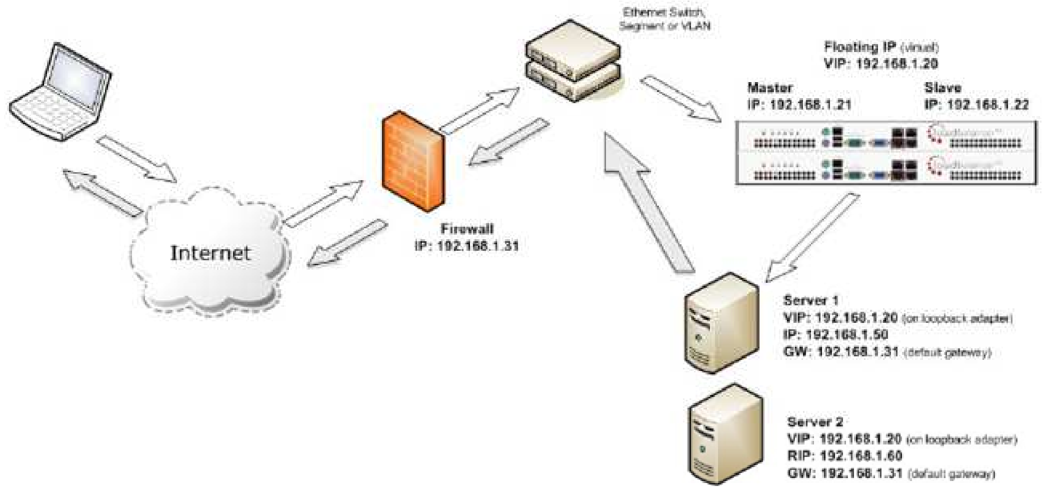
The primary configurations (though not the only ones) supported by loadbalancer.org include Direct Routing (DR), Network Address Translation (NAT), and Source Network Translation (SNAT). Each has its distinctive use.

Direct Routing DR)

The Direct Routing mode is a high-performance solution that requires little change to existing infrastructure. It is a "one-armed" infrastructure in that it need only connect to one subnet. It works at IP Level 4 (the Transport Layer) by changing the MAC (Media Access Control) address of the incoming packet to that of the real server to which the packet is to be routed.

The real server must own both the VIP and its own RIP. However, it must not respond to ARP requests (Address Resolution Protocol – what is the MAC address for this IP address?) for the VIP.

The real servers must be on the same logical network. However, they can be on different subnets so long as there are no router hops between the subnets. If multiple subnets are used, an IP address in each subnet must be defined in the load balancer.

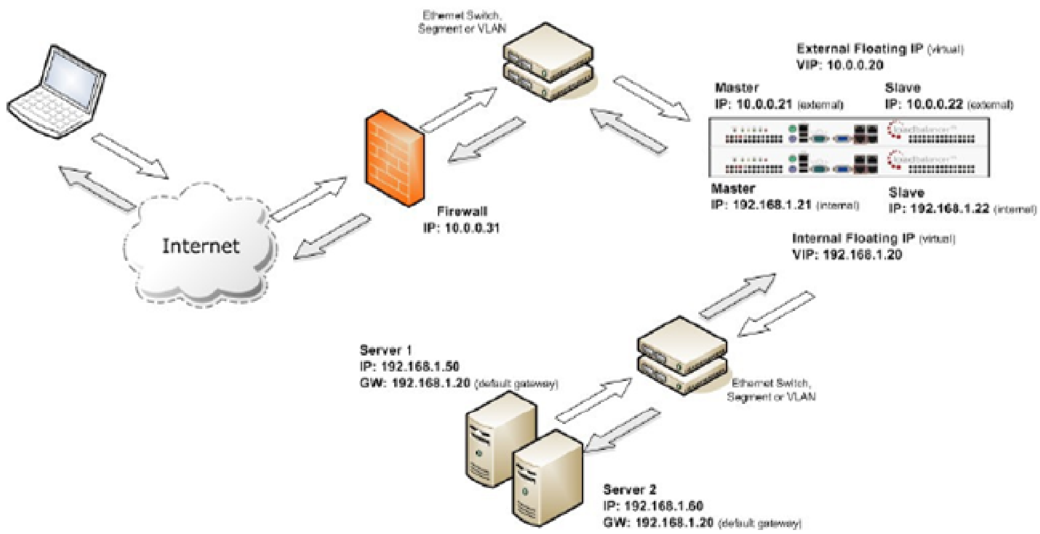


Direct Routing (DR) Configuration
Loadbalancer Quick Start Guide v7.5

Port translation (having a different RIP port than the VIP port) is not possible in DR mode. DR mode is transparent in that the real server will see the source IP address of the client. Administration of the load balancer is via any active IP address on port 9080 via HTTP or on port 9443 via HPPTS.

Network Address Translation (NAT)

DR mode cannot be used if the application cannot bind to the RIP and VIP at the same time or if the operating system cannot be modified to handle the ARP problem. In this case, the Network Address Translation mode can be configured. This is also a high-performance solution that operates at IP Level 4 but requires a “two-armed” infrastructure with an internal and external subnet to carry out the translation.



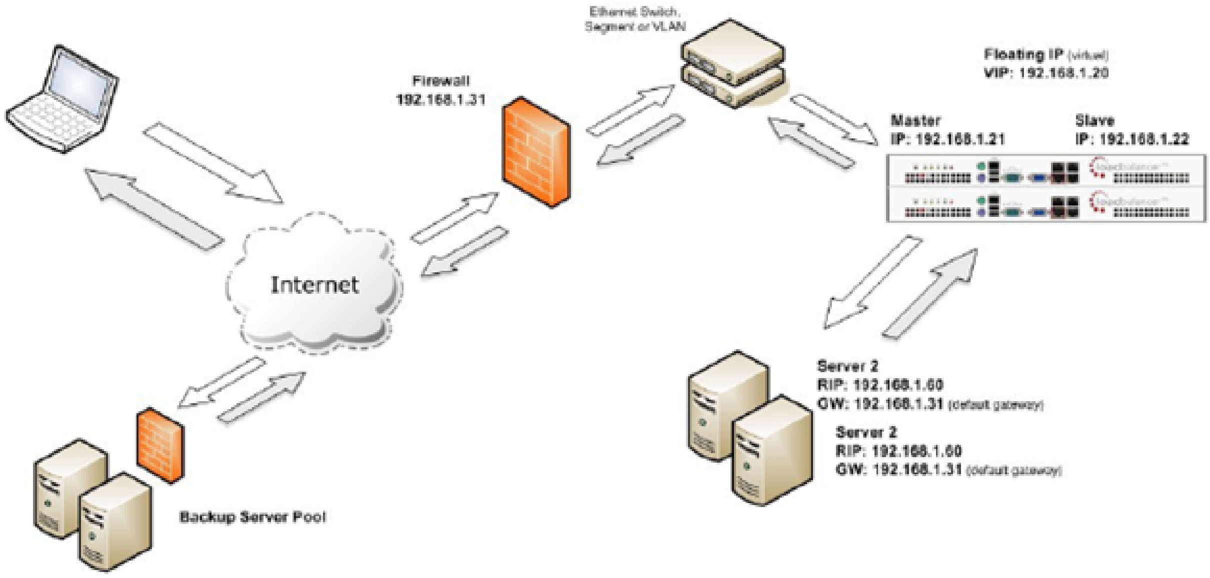
Network Address Translation (NAT) Configuration
Loadbalancer Quick Start Guide v7.5

In the two-armed mode, the load balancer translates all requests from the external VIP address to the internal RIP addresses. The virtual services and the real servers should be configured on different subnets within the same logical network (i.e., no router hops); and the load balancer should have an IP address in each subnet.

The NAT mode is transparent in that the real server will see the source IP address of the client. Port translation is allowed in NAT mode (i.e., VIP:80 → RIP:8080). Administration of the load balancer can be done with any IP address on port 9080 via HTTP or on port 9443 via HPPTS.

Source Network Address Translation (SNAT)

If an application requires that the load balancer handle cookie insertion, then the SNAT configuration must be used. This mode works at Level 7 of the IP stack and is used with applications such as Microsoft's Exchange, SharePoint, and Lync.



Source Network Address Translation (SNAT) Configuration

Loadbalancer Quick Start Guide v7.5

This mode requires only a one-arm configuration and does not require any changes to the application servers. The load balancer proxies the application traffic to the servers so that the source of all traffic becomes the load balancer. However, since the load balancer is acting as a full proxy, this method does not have the same throughput as the Layer 4 methods.

Because SNAT is a full proxy, any server in the application cluster can be on any accessible subnet including across the Internet or a WAN.

SNAT is not transparent. The real servers will not see the source IP address of the client. Rather, they will see the load balancer's IP address.

Load Balancing Algorithms

The Loadbalancer.org load balancers provide several load distribution options:

- <u>Round Robin</u> – packets are distributed to servers one after another.

- Weighted Round Robin – packets are distributed to servers based on their weight. Typically, servers with greater capacity carry a greater weight and will receive more packets to process.

- Least Connection – the servers with the least connections will get the next packets.

- Weighted Least Connection (the default method) – packets are distributed to the servers with the least connections and the greatest weights.

- Destination Hashing – Jobs are assigned to servers by looking up a statically assigned hash table by their IP addresses.

- Real Server Agent – Permits the routing algorithm to be modified by the state of the real servers.

Supported Protocols and Applications

Except for their virtual load balancers (see below), the Loadbalancer.org's load balancers are Intel servers running the Linux operating system with a custom kernel modified for load balancing. Full root access is provided to allow control of all settings.

The load balancers support both IPv4 and IPv6.

The Layer 4 configurations (DR, NAT) make routing decisions based only upon the UDP and TCP protocols, using port numbers and IP addresses. The Layer 7 configuration (SNAT) can make distribution decisions based on upper level protocols such as FTP, HTTP, HTTPS, DNS, RTP, etc.

Loadbalancer.org's load balancers support many applications including:

Microsoft Exchange	Microsoft Office Communications	Streaming Media
Apache/ILS Web Servers	Web Proxies/Filters	Voice over IP (VoIP)
Windows Terminal Services	VMware	Oracle
SharePoint	Hyper-V	Microsoft Dynamics
Lync	Amazon Web Services	

The load balancers do not support:

- Link balancing
- WAN balancing
- Firewall balancing
- Global Server Load Balancing (GSLB)

Persistence

Persistence is the ability to ensure that a specific client connects back to the same server within a given period of time. It is normally required when session state is stored in the real server rather than in a shared database.

In the Level 4 configurations (DR, NAT), Source Persistence is the only option. However, the client will lose its connection in the event of a real server failure.

At Level 7 (SNAT), cookies may be used to preserve persistence across real server failures. Microsoft Connection Broker may also be used.

Loadbalancer.org Load Balancers

Loadbalancer.org provides a series of load balancers offering different characteristics. All may be used as a single load balancer or configured as a high-availability, clustered pair:

- Enterprise – 1.5 gbps throughput with 3 million concurrent sessions.

- Enterprise R16 – Same as Enterprise but licensed only for four application clusters, each with four backend servers.

- Enterprise MAX – 3 gbps throughput with 7 million concurrent sessions.

- Enterprise 10G – More than 10 gbps throughput with 7 million concurrent sessions.

- Enterprise VA – Run on any hardware under VMware or Microsoft Hyper-V hypervisors.

- Enterprise VA R16 – Same as Enterprise VA but licensed only for four application clusters, each with four backend servers.

- Enterprise EC2 - Distribute traffic across Amazon EC2 instances in a single Availability Zone or across multiple Availability Zones

Summary

The Loadbalancer.org load balancer configurations provide a wide range of load balancing options to suit most networks. They are certified for many applications and support several load balancing algorithms. They are offered as appliances with a wide range of performance capabilities from 1.5 gbps to 10 gbps throughput and up to 7 million concurrent connections. They are also offered as virtual load balancers running under VMware or Hyper-V.

Most importantly, they can be configured as master/slave clusters with automatic failover for high availability. They will not be a single point of failure in a network.

Loadbalancer.org offers worldwide services through its offices in the U.S., Canada, the U.K., and Germany.

Excellent descriptions of the Loadbalancer.org load balancers may be found in its Quick Start Guide[2] and in its Administration Manual.[3]

[2] Loadbalancer.org Appliance Quick Start Guide v7.5
http://pdfs.loadbalancer.org/quickstartguideLBv7.pdf

[3] Loadbalancer.org Appliance Administration Manual v7.5
http://pdfs.loadbalancer.org/loadbalanceradministrationv7.pdf