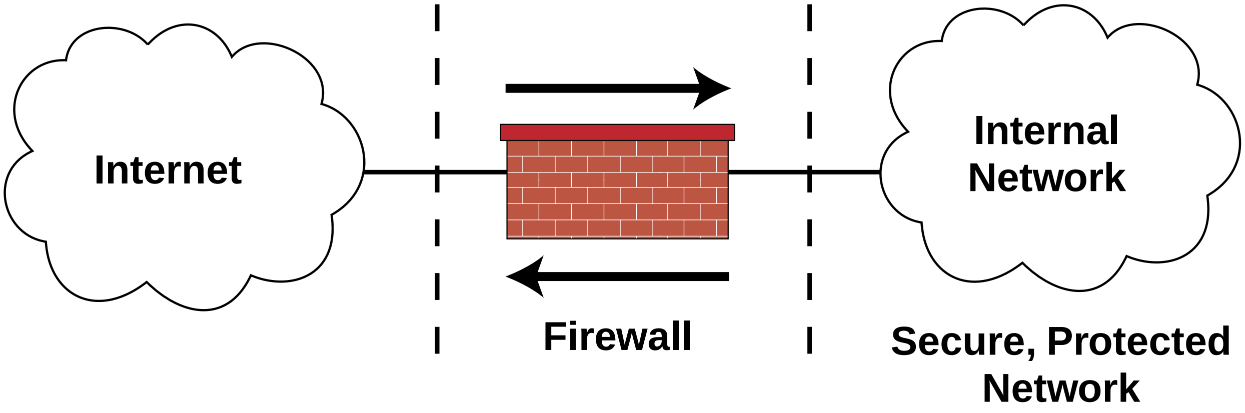
APPLICATION NOTE

FIREWALL LOAD BALANCING WITH SERVERIRON

Today the Internet means connectivity to anyone, anywhere, internal or external to a corporate network. Yet, with all the advantages of such connectivity come challenges to network security. Fortunately, firewalls protect eBusinesses from fraud and unauthorized access and facilitate secure e-commerce over the Internet. But a firewall can be a single point of failure that brings down a network, or can limit network scalability, especially with Internet traffic exploding. Scaling a secure network at such an explosive rate makes it necessary to deploy multiple firewalls for scalability and availability. Foundry's ServerIron firewall load balancing capability simplifies the deployment and management of multiple firewalls for scalability and eliminates the firewall as a single point of failure.

Firewall deployment increases network security at the sacrifice of network performance on two fronts - the number of connections, and the amount of throughput. When a customer accesses a web site or a corporate network through the firewall, the customer's browser establishes TCP connections through the firewall. Because firewalls support a limited number of TCP connections, they limit the number of customers an eBusiness can serve at any given time. Once the browser opens TCP connections, it obtains web pages from the web servers representing the throughput that a firewall must be able to handle.

Multiple firewalls can be deployed for scalability and reliability - but the trick is in balancing the traffic on inbound and outbound paths through these firewalls, and ensuring that users are not affected when a firewall fails or is down for maintenance. Firewalls may or may not be *synchronous:* Synchronous firewalls share connection information among each other, so the data for a given connection can go through any firewall. Non-synchronous firewalls do not



share connection information, and traffic must be revalidated each time it goes through a new firewall. Configuring and managing the different types of firewalls can pose significant challenges to network administrators. Fortunately, Foundry's firewall load balancing implementation ensures that the firewall does not become a brick wall between you and your customers.

In the simplest scenario, shown in the diagram below, a ServerIron switch is connected to the firewalls both on the Internet and the Intranet side for load balancing. The traffic inbound from the Internet is load balanced by ServerIron-A across the firewalls. The traffic outbound to the Internet is load balanced by ServerIron-B across all the firewalls. ServerIron uses a configurable hashing algorithm to select a firewall based on the source IP address, source port number, destination IP address and destination port number. ServerIron-A and ServerIron-B use the same algorithm to ensure that all packets for a given TCP connection flow through the same firewall, thus avoiding any re-authentication overhead on other firewalls. (This is also referred to as "sticky" connections.) Alternately, ServerIron can also ensure load balancing at a session level for specified traffic, ensuring equal distribution of sessions across all the available firewalls.

Extensive Health Checks

Each ServerIron monitors the links to each firewall, the firewall itself, and the ServerIron on the other side. If any component fails in any of the paths between ServerIron-A and ServerIron-B, the ServerIrons transparently redirect the traffic through alternate paths.

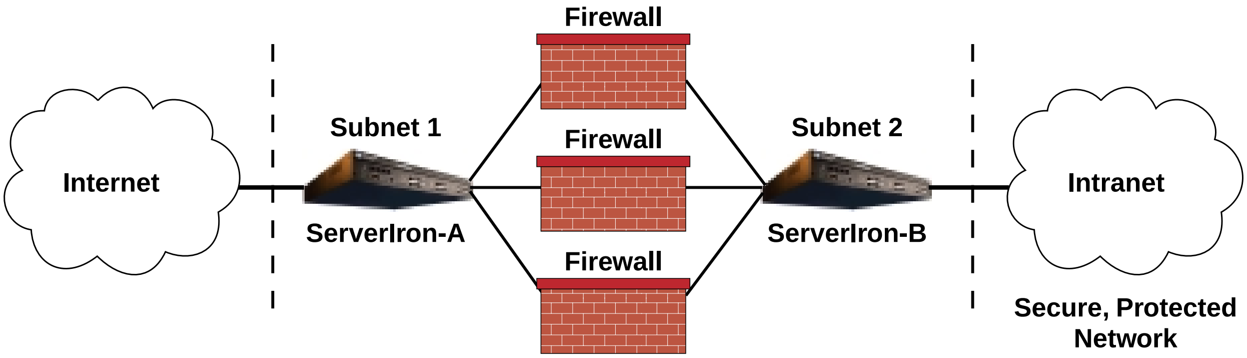
Transparent Deployment

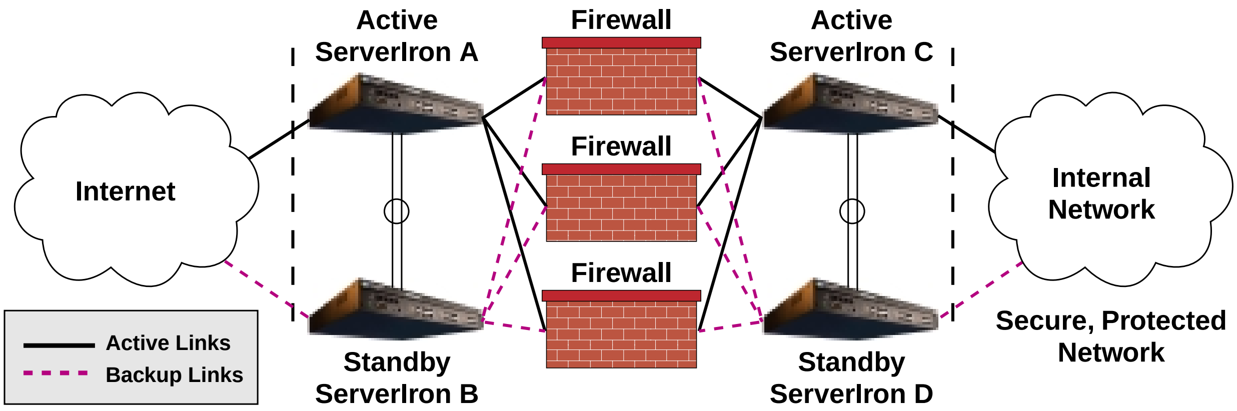
ServerIron can be transparently placed into an existing setup that has multiple firewalls. No matter which firewall the users point to as their gateway, ServerIron intercepts traffic addressed to any firewall, and distributes the load across the available firewalls. Network administrators need not change any network configuration on users' computers.

Support for Network Address Translation

Firewalls may be configured to perform network address translation (NAT) to hide the internal network addresses from the external network. NAT may involve translating all internal IP addresses to one external IP address or mapping each internal IP address to one external IP address. ServerIron sup-

Solution 1: Firewall Sandwich – Designed for Ironclad performance





Solution 2: Firewall Sandwich with fault-tolerance – Designed for Ironclad Performance and Reliability

ports both types of NAT, and ensures sticky connections - data for a given connection always flows through the same firewall.

Two ServerIron switches can be used on both sides of the firewalls in an active-standby configuration. In this configuration, shown in the diagram above, ServerIron-A and ServerIron-C function as the active switch, while the ServerIron-B and ServerIron-D take the role of standby. The active ServerIron switches all traffic, while the stand-by ServerIron monitors the health of the active ServerIron. When using dynamic routing, the standby blocks any layer 3 routing protocol packets, such as OSPF, IGRP and RIP, forcing the adjacent Routers to access the firewalls only through the active ServerIron. When the standby takes on the active role, it immediately begins to forward the OSPF and RIP packets to let adjacent routers discover the new route to access the firewalls. By supporting OSPF, IGRP, and RIP protocols, ServerIron allows faster convergence time compared to products that support only RIP, thereby assuring minimal disruption to network traffic. When static routing is used, both ServerIrons act as high-speed Layer 2 switches, but only the active ServerIron makes the load balancing decisions to ensure sticky connections to firewalls.

Multiple Levels of Fault-Tolerance

When deployed in Active-Standby mode, ServerIrons do periodic health checks to monitor each other's status. The active and standby ServerIrons can be connected over a trunk group consisting of two physical links for fault-tolerance. As shown in the diagram above, ServerIrons provide multiple levels of redundancy:

- Firewall level: If a firewall fails or is taken down for maintenance, traffic is transparently switched to the remaining firewalls.
- Switch level: If a ServerIron fails or is taken down for maintenance, the other ServerIron transparently takes over.
- Link Level: If one of the physical links connecting to the firewall or the ServerIrons fails, an alternate link or firewall is used.

ServerIron Overview

Foundry Networks' ServerIron family of Internet traffic management system switches provide high performance, Layer 4 through 7 switching, enabling network managers to control and manage today's exploding web transaction, web application and eCommerce traffic flows. Internet IronWare - Foundry Networks' unique software suite of Internet traffic management capabilities, powers ServerIronXL, ServerIronXL/G, and BigServerIron (a simple software upgrade to the BigIron chassis) to direct requests to the right server and application based on the information that resides beyond the traditional Layer 2 and 3 packet headers. ServerIron delivers industry leading performance for Internet traffic management functions including local and global server load balancing, firewall load balancing, transparent cache switching, application redirection, packet filtering and prioritization, and support for content-intelligent switching such as cookie-, URL-, and SSL Session ID-based redirection and load balancing.

Foundry's IronCore architecture, combined with custom packet processing ASICs, offers flexible deployment and support for extensive network topologies. ServerIron's shared memory architecture ensures exceptional concurrent connection capacity whether you use 2 ports or 24 ports. With an optional redundant power supply and a rack-optimized form factor, ServerIron provides the performance, port density, reliability, and flexibility required by every network manager and administrator.

FOUNDRY NETWORKS

2100 Gold Street
P.O. Box 649100
San Jose, CA 95164-9100
Tel 408.586.1700
Fax 408.586.1900
www.foundrynet.com