

## White Paper for Load Balancing

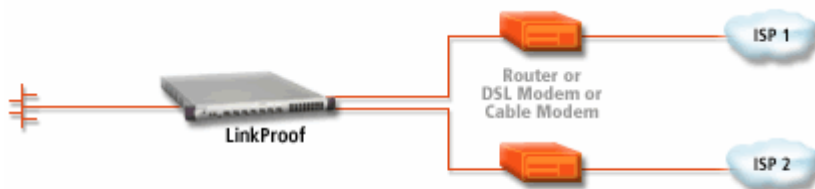
Load balancing allows incoming HTTP requests to a company's website to be distributed across several physical servers. If one server should fail, other servers can pick up the workload. Load balancing can be performed using application software. However, our dedicated switch can perform this function much faster as the redirection is handled by specialized hardware.

With load balancing, clients access servers through a virtual IP address. The load balancer transparently redirects the requests with no change required on the clients or servers; all configuration and redirection is done on the load balancer.

It is possible to load balance using TCP or UDP providing high availability for HTTP, FTP, DNS, SMTP, etc.

Load balancing clients connect to a virtual IP address, which in reality is redirected to one of several physical servers in a load balancing group. In many web page display applications, a client may have its requests redirected to and serviced by different servers in the group. In certain situations, however, it may be critical that all traffic for the client be directed to the same physical server for the duration of the session; this is the concept of session persistence.

When the load balancer receives a new session request from a client for a specific virtual address, the load balancer creates a binding between the client (source) IP address/port socket and the (destination) IP address/port socket of the load balancing server selected for this client. Subsequent packets from clients are compared to the list of bindings: if there is a match, the packet is sent to the same server previously selected for this client; if there is not a match, a new binding is created. Various session persistence settings can be used which include TCP, SSL, Sticky, VPN and IP.



**Multihoming - Load balance multiple ISP links**