



---

## A

**active standby** Redundancy scheme in which two IOS SLB devices can load-balance the same virtual IP address while at the same time acting as backups for each other. See also *active standby*, *stateful backup*.

---

## C

**CASA** Content Aware Services Architecture. CASA is a protocol designed to allow network appliances to selectively control the flow of IP packets through a router, switch, or other network device.

**cluster** Set of computer systems that are connected together through multisystem hardware or software to provide services traditionally provided by a single system. This arrangement provides higher availability and better scalability of the system.

**content-aware networking** Networking strategy that enables content to be distributed dynamically. Because content can be dynamically cached, it can be located at any given place at any given time and distributed between the servers and the location of the Web cache. Cisco has developed the ContentFlow architecture and the DFP to enable networks to provide content-aware networking services.

**ContentFlow architecture** Cisco's content-aware networking architecture that describes message flows and actions in a distributed environment.

---

## D

**DFP** Dynamic Feedback Protocol. Allows host agents to report the change in status of the host systems dynamically. The status reported is a relative weight that specifies a host server's capacity to perform work.

**Dynamic Feedback Protocol** See *DFP*.

---

## L

**load balancing** Spreading user requests among available servers within a cluster of servers, based on a variety of algorithms.

---

## M

**MD5** Message Digest Algorithm Version 5. Neighbor router authentication scheme used to ensure reliability and security when routing updates are exchanged between neighbor routers.

**Message Digest Algorithm Version 5** See *MD5*.

---

## N

**NAT** Network Address Translation. Modification of one or more of the following fields in an IP packet: source IP address, destination IP address, source TCP/UDP port, destination TCP/UDP port.

**Network Address Translation** *See NAT.*

**NetFlow switching** High-performance network-layer switching path that captures as part of its switching function a rich set of traffic statistics including user, protocol, port, and type of service information.

---

## R

**real server** The specification of a physical server associated with a virtual server. The specification includes the real server's IP address and an optional weight to be used by the virtual server predictor.

**round robin** *See weighted round-robin.*

---

## S

**Secure Socket Layer** *See SSL.*

**server cluster** *See server farm.*

**server farm** Also called a server cluster. Group of real servers that provide various applications and services.

**server load balancing** *See SLB.*

**services manager** Functionality built into IOS SLB that makes load-balancing decisions based on application availability, server capacity, and load distribution algorithms such as weighted round-robin or weighted least connections, or the DFP. The services manager determines a real server for the packet flow using load balancing and server/application feedback.

**SLB** *See SLB.*

**SSL** Secure Socket Layer. Encryption technology for the Web used to provide secure transactions such as the transmission of credit card numbers for e-commerce.

**stateful backup** Redundancy scheme that enables IOS SLB to incrementally back up its load-balancing decisions, or "keep state," between primary and backup switches. *See also active standby, stateless backup.*

**stateless backup** Redundancy scheme that provides high network availability by routing IP flows from hosts on Ethernet networks without relying on the availability of a single SLB switch. *See also active standby, stateful backup.*

---

## V

**virtual server** Presents a single address that represents an application server farm to clients.

---

**W**

- weighted least connection** Load-balancing algorithm in which the next real server chosen for a new connection to the virtual server is the server with the fewest active connections. Each real server is assigned a weight,  $n$ , that represents its capacity to handle connections, as compared to the other real servers associated with the virtual server. The server with the fewest connections is based on the number of active connections on each server, and on the relative capacity of each server. The capacity of a given real server is calculated as the assigned weight of that server divided by the sum of the assigned weights of all of the real servers associated with that virtual server, or  $n_1/(n_1+n_2+n_3\dots)$ .
- weighted round-robin** Load balancing algorithm in which the real server used for a new connection to the virtual server is chosen in a circular fashion. Each real server is assigned a weight,  $n$ , that represents its capacity to handle connections, as compared to the other real servers associated with the virtual server. New connections are assigned to a given real server  $n$  times before the next real server in the list is chosen.
- workload agents** Value-added software components developed for specific platforms by third-party developers. Workload agents run on server platforms or on platforms that manage server farms. Workload agents deliver server and application information to the services manager. This information enables the services manager to make optimum server selection.

