Table of Contents

How Does Load Balancing Work?

Introduction

Load−balancing is a concept that allows a router to take advantage of multiple best paths to a given destination. The paths are derived either statically or with dynamic protocols, such as RIP, EIGRP, OSPF, and IGRP.

Before You Begin

Conventions

For more information on document conventions, see the Cisco Technical Tips Conventions.

Prerequisites

There are no specific prerequisites for this document.

Components Used

This document is not restricted to specific software and hardware versions.

The information presented in this document was created from devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. If you are working in a live network, ensure that you understand the potential impact of any command before using it.

Load−Balancing

When a router learns multiple routes to a specific network via multiple routing processes (or routing protocols), it installs the route with the lowest administrative distance in the routing table. For more detailed information about this process, see Route Selection in Cisco Routers.

Sometimes the router must select a route from among many learned via the same routing process with the same administrative distance. In this case, the router chooses the path with the lowest cost (or metric) to the destination. Each routing process calculates its cost differently and the costs may need to be manipulated in order to achieve load−balancing.

If the router receives and installs multiple paths with the same administrative distance and cost to a destination, load−balancing can occur. The IGRP and EIGRP routing processes also support unequal cost load−balancing. You can use the variance command with IGRP and EIGRP to accomplish unequal cost load−balancing. For more information about variance, see How Does Unequal Cost Path Load−Balancing

(Variance) Work in IGRP and EIGRP?.

Equal cost routes can usually be found by using the show ip route command. For example, below is the output for show ip route to a particular subnet that has multiple routes. Notice there are two routing descriptor blocks. Each block is one route. There is also an asterisk (*) next to one of the block entries. This corresponds to the active route that is used for new traffic. The term 'new traffic' corresponds to a single packet or an entire flow to a destination, depending on the type of switching configured.

- For process−switching: the load balancing is on a per−packet basis and * points to the interface over which the next packet is sent.
- For fast−switching: load balancing is on a per−destination basis and * points to the interface over which the next destination−based flow is sent.

The position of the * keeps rotating among the equal cost paths each time a packet/flow is served.

```
M2515−B#show ip route 1.0.0.0
Routing entry for 1.0.0.0/8
  Known via "rip", distance 120, metric 1
  Redistributing via rip
  Advertised by rip (self originated)
  Last update from 192.168.75.7 on Serial1, 00:00:00 ago
  Routing Descriptor Blocks:
  * 192.168.57.7, from 192.168.57.7, 00:00:18 ago, via Serial0
      Route metric is 1, traffic share count is 1
    192.168.75.7, from 192.168.75.7, 00:00:00 ago, via Serial1
      Route metric is 1, traffic share count is 1
```

As of the writing of this document, you can have up to six equal cost routes (a limit imposed by Cisco IOS on the routing table), but some Interior Gateway Protocols (IGPs) have their own limitation. EIGRP, for example, allows up to four equal cost routes.

The following documents provide more information about how various protocols select a best path and calculate their costs to specific destinations.

- Introduction to RIP
- IGRP Metric − Example and Explanation
- Setting a Preferred Route by Influencing EIGRP Metrics
- OSPF Cost
- BGP Path Selection Criteria

You can also set load−balancing to work per destination or per packet. Per−destination load balancing means the router distributes the packets based on the destination address. Given two paths to the same network, all packets for destination1 on that network go over the first path, all packets for destination2 on that network go over the second path, and so on. Per−packet load−balancing means that the router sends one packet for destination1 over the first path, the second packet for (the same) destination1 over the second path, and so on.

Per−destination or per−packet load−balancing depends on the type of switching scheme used for IP packets. By default, on most Cisco routers, fast switching is enabled under interfaces. This is a demand caching scheme that does per−destination load−balancing. To set per−packet load−balancing, enable process switching (or disable fast switching) using the following commands:

```
Router# config t
Router(config)# interface Ethernet 0
Router(config−if)# no ip route−cache
Router(config−if)# ^Z
```

Now the router's CPU looks at every single packet and load balances on the number of routes in the routing table for the destination. This can crash a low−end router because the CPU must do all the processing. To re−enable fast switching, use these commands:

```
Router# config t
Router(config)# interface Ethernet 0
Router(config−if)# ip route−cache
Router(config−if)# ^Z
```

Newer switching schemes such as Cisco Express Forwarding allow you to do per−packet and per−destination load−balancing more quickly. However, it does imply that you have the extra resources to deal with maintaining Cisco Express Forwarding entries and adjacencies. Refer to this Application Note for more information on load−sharing with Cisco Express Forwarding.

For more details on Cisco Express Forwarding load balancing, see Troubleshooting Load Balancing Over Parallel Links Using Cisco Express Forwarding.

Related Information

- **TCP/IP Routing and Routed Protocols Support Page**
- **Technical Support − Cisco Systems**