CISCO SYSTEMS

CCIE Professional Development

Routing TCP/IP

Volume I, Second Edition

A detailed examination of interior routing protocols

No. 1
Bestseller
Fully Updated

Jeff Doyle, CCIE® No. 1919
Jennifer Carroll, CCIE No. 1402

ciscopress.com

CCIE Professional Development Routing TCP/IP Volume I

Second Edition

Jeff Doyle, CCIE No. 1919
Jennifer Carroll, CCIE No. 1402

Cisco Press

800 East 96th Street
Indianapolis, Indiana 46240 USA

Example 7-1 *The output of every show interface command includes metric statistics for the interface. This Ethernet interface shows MTU = 1500 bytes, bandwidth = 10 megabits per second, delay = 1000 microseconds, reliability = 100 percent, and load = 39 percent (minimum load). (Continued)*

```
  Output queue 0/40, 0 drops; input queue 0/75, 0 drops
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
     601753 packets input, 113607697 bytes, 0 no buffer
     Received 601753 broadcasts, 0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
     0 input packets with dribble condition detected
     693494 packets output, 557731861 bytes, 0 underruns
     0 output errors, 5 collisions, 13 interface resets
     0 babbles, 0 late collision, 48 deferred
     0 lost carrier, 0 no carrier
     0 output buffer failures, 0 output buffers swapped out
Newfoundland#
```

Bandwidth is expressed in units of kilobits per second. It is a static number used for metric calculation only and does not necessarily reflect the actual bandwidth of the link—that is, bandwidth is not measured dynamically. For example, the default bandwidth of a serial interface is 1544, whether the interface is attached to a T1 or a 56K line. This bandwidth number may be changed from the default with the **bandwidth** command.

IGRP updates use a three-octet number, referred to in this book as BW_{IGRP}, which is the inverse of the bandwidth scaled by a factor of 10^7. So if the bandwidth of an interface is 1544, then

$$BW_{IGRP} = 10^7/1544 = 6476, \text{ or } 0x00194C.$$

Delay, like bandwidth, is a static figure and is not measured dynamically. It is displayed by the **show interface** command as DLY, in units of microseconds. The default delay of an interface may be changed with the **delay** command, which specifies the delay in tens of microseconds. Example 7-2 shows the **bandwidth** and **delay** commands used to change the defaults of the interface of Example 7-1.

Example 7-2 *The bandwidth and delay commands are used to change the metric defaults of the e0 interface. The new quantities can be seen in the output of the* **show interfaces** *command.*

```
Newfoundland(config)#interface e0
Newfoundland(config-if)#bandwidth 75000
Newfoundland(config-if)#delay 5
Newfoundland(config-if)#^Z
Newfoundland#
%SYS-5-CONFIG_I: Configured from console by console
Newfoundland#show interfaces ethernet0
Ethernet0 is up, line protocol is up
  Hardware is Lance, address is 0000.0c76.5b7c (bia 0000.0c76.5b7c)
  Internet address is 10.2.1.2/24
```

Example 7-2 *The bandwidth and delay commands are used to change the metric defaults of the e0 interface. The new quantities can be seen in the output of the* **show interfaces** *command. (Continued)*

```
  MTU 1500 bytes, BW 75000 Kbit, DLY 50 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive set (10 sec)
  ARP type: ARPA, ARP Timeout 04:00:00
  Last input 00:00:02, output 00:00:02, output hang never
  Last clearing of "show interface" counters never
  Queueing strategy: fifo
  Output queue 0/40, 0 drops; input queue 0/75, 0 drops
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
     601888 packets input, 113637882 bytes, 0 no buffer
     Received 601888 broadcasts, 0 runts, 0 giants, 0 throttles
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored
     0 input packets with dribble condition detected
     693646 packets output, 557884632 bytes, 0 underruns
     0 output errors, 5 collisions, 13 interface resets
     0 babbles, 0 late collision, 48 deferred
     0 lost carrier, 0 no carrier
     0 output buffer failures, 0 output buffers swapped out
Newfoundland#
```

When carried in an IGRP update, delay is a three-octet number expressed in the same 10-microsecond units as specified by the **delay** command. To avoid confusion, this number will be referred to as DLY_{IGRP}, to differentiate it from DLY, in microseconds, observed with **show interface**. For example, if DLY is 50, then

$$DLY_{IGRP} = DLY/10 = 50/10 = 5, \text{ or } 0x000005.$$

IGRP also uses delay to indicate an unreachable route by setting DLY_{IGRP} = 0xFFFFFF. This number translates to approximately 167.8 seconds, so the maximum end-to-end delay of an IGRP route is 167 seconds.

Because IGRP uses bandwidth and delay as its default metrics, these quantities must be configured correctly and consistently on all interfaces of all IGRP routers.

Changing the bandwidth or delay of an interface should be done only for good reasons and only with a full understanding of the results of those changes. In most cases, it is best to leave the default values unchanged. A notable exception is serial interfaces. As noted earlier in this section, serial interfaces on Cisco routers have a default bandwidth of 1544 no matter what the bandwidth is of the connected link. The **bandwidth** command should be used to set the interface to the actual bandwidth of the serial link.

It is important to note that OSPF also uses the bandwidth statement to calculate its metric. Therefore, if IGRP (or EIGRP) metrics are to be manipulated in a network where OSPF is also running, use the **delay** to influence IGRP. Changing the bandwidth will affect both IGRP and OSPF.

Table 7-1 lists the bandwidths and delays for a few common interfaces. (The default bandwidth of a serial interface is always 1544; Table 7-1 shows the figures that would result from using the **bandwidth** command to reflect the actual connected bandwidth.)

Table 7-1 *Common BW_{IGRP} and DLY_{IGRP} quantities.*

Media	Bandwidth	BW_{IGRP}	Delay	DLY_{IGRP}
100M ATM	100000K	100	100 S	10
Fast Ethernet	100000K	100	100 S	10
FDDI	100000K	100	100 S	10
HSSI	45045K	222	20000 S	2000
16M Token Ring	16000K	625	630 S	63
Ethernet	10000K	1000	1000 S	100
T1	1544K	6476	20000 S	2000
DS0	64K	156250	20000 S	2000
56K	56K	178571	20000 S	2000
Tunnel	9K	1111111	500000 S	50000

Reliability is measured dynamically and is expressed as an eight-bit number, where 255 is a 100 percent reliable link and 1 is a minimally reliable link. In the output of **show interface**, reliability is shown as a fraction of 255, for example, 234/255 (see Example 7-3).

Example 7-3 *This interface shows a reliability of 234/255, or 91.8 percent.*

```
Casablanca#show interface ethernet0
Ethernet0 is up, line protocol is up
  Hardware is Lance, address is 0000.0c76.5b7c (bia 0000.0c76.5b7c)
  Internet address is 172.20.1.1 255.255.255.0
  MTU 1500 bytes, BW 10000 Kbit, DLY 1000 usec, rely 234/255, load 1/255
  Encapsulation ARPA, loopback not set, keepalive set (10 sec)
  ARP type: ARPA, ARP Timeout 4:00:00
  Last input 0:00:28, output 0:00:06, output hang never
  Last clearing of "show interface" counters 0:06:05
  Output queue 0/40, 0 drops; input queue 0/75, 0 drops
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
     22 packets input, 3758 bytes, 0 no buffer
     Received 21 broadcasts, 0 runts, 0 giants
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
     0 input packets with dribble condition detected
     125 packets output, 11254 bytes, 0 underruns
     39 output errors, 694 collisions, 0 interface resets, 0 restarts
     0 output buffer failures, 0 output buffers swapped out
Casablanca#
```

Load, in an IGRP update, is an eight-bit number. Load is represented in the output of **show interface** as a fraction of 255, such as 40/255 (Example 7-4); 1 is a minimally loaded link, and 255 is a 100 percent loaded link.

Example 7-4 *This interface shows a load of 40/255, or 15.7 percent.*

```
Yalta#show interface serial 1
Serial1 is up, line protocol is up
  Hardware is HD64570
  Internet address is 172.20.20.2 255.255.255.0
  MTU 1500 bytes, BW 56 Kbit, DLY 20000 usec, rely 255/255, load 40/255
  Encapsulation HDLC, loopback not set, keepalive set (10 sec)
  Last input 0:00:08, output 0:00:00, output hang never
  Last clearing of "show interface" counters 0:05:05
  Output queue 0/40, 0 drops; input queue 0/75, 0 drops
  5 minute input rate 10000 bits/sec, 1 packets/sec
  5 minute output rate 9000 bits/sec, 1 packets/sec
     456 packets input, 397463 bytes, 0 no buffer
     Received 70 broadcasts, 0 runts, 0 giants
     0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored, 0 abort
     428 packets output, 395862 bytes, 0 underruns
     0 output errors, 0 collisions, 0 interface resets, 0 restarts
     0 output buffer failures, 0 output buffers swapped out
     0 carrier transitions
     DCD=up  DSR=up  DTR=up  RTS=up  CTS=up
```

If reliability or load is to be used as a metric or as part of a composite metric, the algorithm for calculating the metric must not allow sudden changes in the error rate or channel occupancy to destabilize the network. As an example, if a "raw," or instantaneous, measure of load is used, a burst of heavy traffic could cause a route to go into holddown and an abrupt drop in traffic could trigger an update. To prevent frequent metric changes, reliability and load are calculated based on an exponentially weighted average with a five-minute time constant, which is updated every five seconds.

The composite metric for each IGRP route is calculated as

$$metric = [k1*BW_{IGRP(min)} + (k2*BW_{IGRP(min)})/(256-LOAD) + k3*DLY_{IGRP(sum)}] [k5/(RELIABILITY + k4)],$$

where $BW_{IGRP(min)}$ is the minimum BW_{IGRP} of all the outgoing interfaces along the route to the destination and $DLY_{IGRP(sum)}$ is the total DLY_{IGRP} of the route.

The values k1 through k5 are configurable weights; their default values are k1=k3=1 and k2=k4=k5=0. These defaults can be changed with the command:[3]

metric weights tos *k1 k2 k3 k4 k5*[3]

[3] *tos* is a relic of the Cisco original intention to have IGRP do type of service routing; this plan was never adopted, and *tos* in this command is always set to zero.

If k5 is set to zero, the [k5/(RELIABILITY+k4)] term is not used.

Given the default values for k1 through k5, the composite metric calculation used by IGRP reduces to the default metric:

$$\text{metric} = BW_{IGRP(min)} + DLY_{IGRP(sum)}$$

The network example in Figure 7-3 shows the bandwidths and delays configured on each interface and a forwarding database from one of the routers with the derived IGRP metrics.[4]

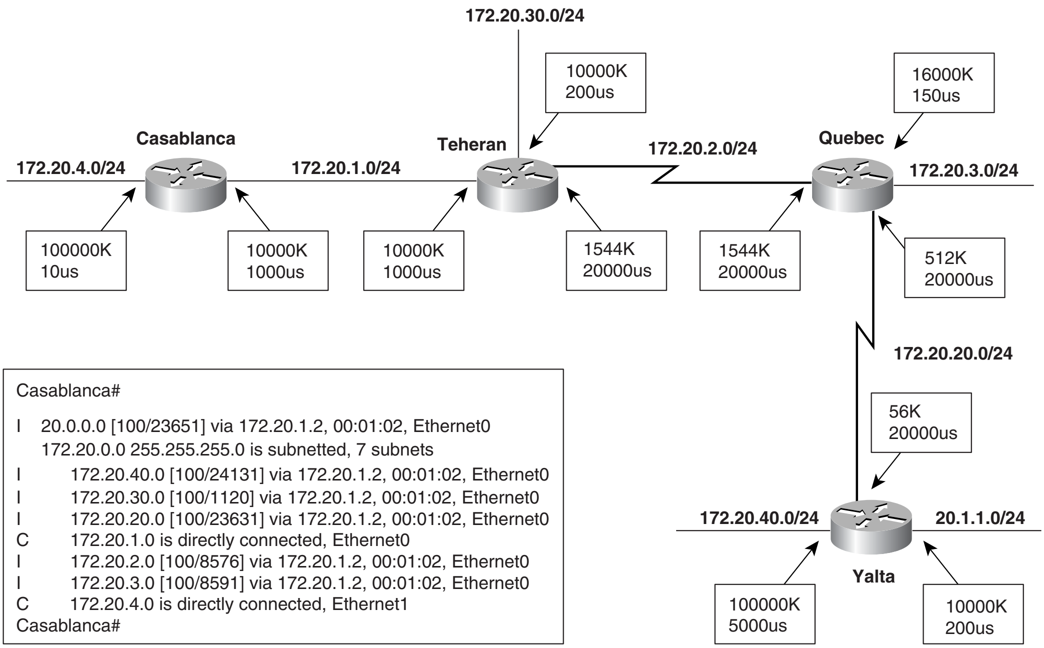

Figure 7-3 *By default, the total delay is added to the minimum bandwidth to derive the IGRP metric.*

The routing table itself shows only the derived metric, but the actual variables recorded by IGRP for each route can be seen by using the command **show ip route** *address*, as in Example 7-5. Here the minimum bandwidth on the route from Casablanca to subnet 172.20.40.0/24 is 512K, at Quebec. The total delay of the route is (1000 + 20000 + 20000 + 5000) = 46000 microseconds:

$$BW_{IGRP(min)} = 10^7/512 = 19531$$
$$DLY_{IGRP(sum)} = 46000/10 = 4600$$
$$\text{metric} = BW_{IGRP(min)} + DLY_{IGRP(sum)} = 19531 + 4600 = 24131$$

[4] Also notice the administrative distance, which is 100 for IGRP.

Example 7-5 *The metric for the route from Casablanca to subnet 172.20.40.0 is calculated from the minimum bandwidth of 512K and the total delay of 46000 microseconds.*

```
Casablanca#show ip route 172.20.40.0
Routing entry for 172.20.40.0 255.255.255.0
  Known via "igrp 1", distance 100, metric 24131
  Redistributing via igrp 1
  Advertised by igrp 1 (self originated)
  Last update from 172.20.1.2 on Ethernet0, 00:00:54 ago
  Routing Descriptor Blocks:
  * 172.20.1.2, from 172.20.1.2, 00:00:54 ago, via Ethernet0
      Route metric is 24131, traffic share count is 1
      Total delay is 46000 microseconds, minimum bandwidth is 512 Kbit
      Reliability 255/255, minimum MTU 1500 bytes
      Loading 1/255, Hops 2
```

Example 7-5 shows that IGRP also records the smallest MTU along the route in addition to the hop count. MTU is not used in the metric calculation. Hop count is the hop count reported by the next-hop router and is used only to limit the diameter of the network. By default, the maximum hop count is 100 and can be configured from 1 to 255 with the command **metric maximum-hops**. If the maximum hop count is exceeded, the route will be marked unreachable by setting the delay to 0xFFFFFF.

Note that all metrics are calculated from the outgoing interfaces along the route. For example, the metric for the route from Yalta to subnet 172.20.4.0/24 is different from the metric for the route from Casablanca to subnet 172.20.40.0/24. This is due to the differences in the configured bandwidth on the link between Yalta and Quebec and to the differences in the delay on the outgoing interfaces to the two destination subnets.

From IGRP to EIGRP

The original motivation for developing EIGRP was simply to make IGRP classless. But early in the development the engineers working on the project recalled some academic proposals for a new kind of convergence algorithm and decided to use that algorithm in their extension of IGRP. The result was a protocol that, while retaining some concepts introduced with IGRP such as multiple metrics, protocol domains, and unequal-cost load balancing, is distinctly different from IGRP.

EIGRP is occasionally described as a distance vector protocol that acts like a link-state protocol. To recap the extensive discussion in Chapter 4, "Dynamic Routing Protocols," a distance vector protocol shares everything it knows, but only with directly connected neighbors. Link-state protocols announce information only about their directly connected links, but they share the information with all routers in their routing domain or area.

All the distance vector protocols discussed so far run some variant of the Bellman-Ford (or Ford-Fulkerson) algorithm. These protocols are prone to routing loops and counting