CHAPTER 24

Interior Gateway Routing Protocol and Enhanced IGRP

Background

The Interior Gateway Routing Protocol (IGRP) is a routing protocol developed in the mid-1980s by Cisco Systems, Inc. Cisco's principal goal in creating IGRP was to provide a robust protocol for routing within an *autonomous system* (AS) having arbitrarily complex topology and consisting of media with diverse bandwidth and delay characteristics. An AS is a collection of networks under common administration that share a common routing strategy. ASs are typically given a unique 16-bit number that is assigned by the *Defense Data Network* (DDN) *Network Information Center* (NIC).

In the mid-1980s, the most popular intra-AS routing protocol was the *Routing Information Protocol* (RIP). Although RIP was quite useful for routing within small- to moderate-sized, relatively homogeneous internetworks, its limits were being pushed by network growth. In particular, RIP's small hop-count limit (16) restricted the size of internetworks, and its single metric (hop count) did not allow for much routing flexibility in complex environments. For more information on RIP, see Chapter 23, "Routing Information Protocol." The popularity of Cisco routers and the robustness of IGRP have encouraged many organizations with large internetworks to replace RIP with IGRP.

Cisco's initial IGRP implementation worked in *Internet Protocol* (IP) networks. IGRP was designed to run in any network environment, however, and Cisco soon ported it to run in *Open System Interconnection* (OSI) *Connectionless Network Protocol* (CLNP) networks.

Cisco developed Enhanced IGRP in the early 1990s to improve the operating efficiency of IGRP. Enhanced IGRP is discussed in detail later in this chapter.

IGRP

IGRP is a *distance vector interior-gateway protocol* (IGP). Distance vector routing protocols call for each router to send all or a portion of its routing table in a routing update message at regular intervals to each of its neighboring routers. As routing information proliferates through the network, routers can calculate distances to all nodes within the internetwork.

Distance vector routing protocols are often contrasted with link state routing protocols, which send local connection information to all nodes in the internetwork. For a discussion of *Open Shortest Path First* (OSPF) and *Intermediate System-to-Intermediate System* (IS-IS), two popular link state routing algorithms, see Chapter 25, "Open Shortest Path First," and Chapter 28, "OSI Routing," respectively.

IGRP uses a combination (vector) of metrics. *Internetwork delay*, *bandwidth*, *reliability*, and *load* are all factored into the routing decision. Network administrators can set the weighting factors for each of these metrics. IGRP uses either the administrator-set or the default weightings to automatically calculate optimal routes.

IGRP provides a wide range for its metrics. For example, reliability and load can take on any value between 1 and 255; bandwidth can take on values reflecting speeds from 1,200 bps to 10 gigabits per second; while delay can take on any value from 1 to 2 to the 24th power. Wide metric ranges allow satisfactory metric setting in internetworks with widely varying performance characteristics. Most importantly, the metric components are combined in a user-definable algorithm. As a result, network administrators can influence route selection in an intuitive fashion.

To provide additional flexibility, IGRP permits multipath routing. Dual equal-bandwidth lines may run a single stream of traffic in round-robin fashion, with automatic switchover to the second line if one line goes down. Also, multiple paths can be used even if the metrics for the paths are different. For example, if one path is three times better than another because its metric is three times lower, the better path will be used three times as often. Only routes with metrics that are within a certain range of the best route are used as multiple paths.

Stability Features

IGRP provides a number of features that are designed to enhance its stability. These include *hold-downs*, *split horizons*, and *poison reverse updates*.

Hold-Downs

Hold-downs are used to prevent regular update messages from inappropriately reinstating a route that may have gone bad. When a router goes down, neighboring routers detect this via the lack of regularly scheduled update messages. These routers then calculate new routes and send routing update messages to inform their neighbors of the route change. This activity begins a wave of triggered updates that filter through the network.

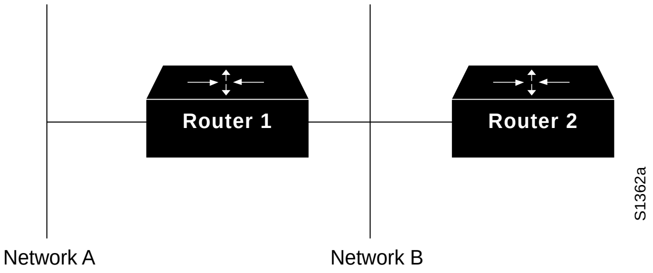
These triggered updates do not instantly arrive at every network device. It is therefore possible for a device that has yet to be informed of a network failure to send a regular update message (indicating that a route that has just gone down is still good) to a device that has just been notified of the network failure. In this case, the latter device would now contain (and potentially advertise) incorrect routing information.

Hold-downs tell routers to hold down any changes that might affect routes for some period of time. The hold-down period is usually calculated to be just greater than the period of time necessary to update the entire network with a routing change.

Split Horizons

Split horizons derive from the fact that it is never useful to send information about a route back in the direction from which it came. For example, consider Figure 24-1.

Figure 24-1 Split Horizons



Router 1 (R1) initially advertises that it has a route to Network A. There is no reason for Router 2 (R2) to include this route in its update back to R1, as R1 is closer to Network A. The split-horizon rule says that R2 should strike this route from any updates it sends to R1.

The split-horizon rule helps prevent routing loops. For example, consider the case where R1's interface to Network A goes down. Without split horizons, R2 continues to inform R1 that it can get to Network A (through R1). If R1 does not have sufficient intelligence, it may actually pick up R2's route as an alternative to its failed direct connection, causing a routing loop. Although hold-downs should prevent this, split horizons are implemented in IGRP because they provide extra algorithm stability.

Poison Reverse Updates

Whereas split horizons should prevent routing loops between adjacent routers, poison reverse updates are intended to defeat larger routing loops. Increases in routing metrics generally indicate routing loops. Poison reverse updates are then sent to remove the route and place it in hold-down. In Cisco's implementation of IGRP, poison reverse updates are sent if a route metric has increased by a factor of 1.1 or greater.

Timers

IGRP maintains a number of timers and variables containing time intervals. These include an update timer, an invalid timer, a hold-time period, and a flush timer. The update timer specifies how frequently routing update messages should be sent. The IGRP default for this variable is 90 seconds. The invalid timer specifies how long a router should wait, in the absence of routing update messages about a specific route, before declaring that route invalid. The IGRP default for this variable is three times the update period. The hold-time variable specifies the hold-down period. The IGRP default for this variable is three times the update timer period plus ten seconds. Finally, the flush timer indicates how much time should pass before a route should be flushed from the routing table. The IGRP default is seven times the routing update period.

Enhanced IGRP

With Software Release 9.21, Cisco introduced an enhanced version of IGRP that combines the advantages of link state protocols with the advantages of distance vector protocols. Enhanced IGRP incorporates the *Diffusing Update Algorithm* (DUAL) developed at SRI International by Dr. J.J. Garcia-Luna-Aceves. Enhanced IGRP includes the following features:

- Fast convergence—Enhanced IGRP uses DUAL to achieve convergence quickly. A router running Enhanced IGRP stores all of its neighbors' routing tables so that it can quickly adapt to alternate routes. If no appropriate route exists, Enhanced IGRP queries its neighbors to discover an alternate route. These queries propagate until an alternate route is found.
- Variable length subnet masks—Enhanced IGRP includes full support for variable length subnet masks. Subnet routes are automatically summarized on a network number boundary. In addition, Enhance IGRP can be configured to summarize on any bit boundary at any interface.
- Partial, bounded updates—Enhanced IGRP does not make periodic updates. Instead, it sends partial updates only when the metric for a route changes. Propagation of partial updates is automatically bounded so that only those routers that need the information are updated. As a result of these two capabilities, Enhanced IGRP consumes significantly less bandwidth than IGRP.
- Multiple network-layer support—Enhanced IGRP includes support for AppleTalk, IP, and Novell NetWare. The AppleTalk implementation redistributes routes learned from the Routing Table Maintenance Protocol (RTMP). The IP implementation redistributes routes learned from OSPF, Routing Information Protocol (RIP), IS-IS, Exterior Gateway Protocol (EGP), or Border Gateway Protocol (BGP). The Novell implementation redistributes routes learned from Novell RIP or Service Advertisement Protocol (SAP).

Enhanced IGRP features four new technologies:

- *Neighbor discovery/recovery*—Used by routers to dynamically learn about other routers on their directly attached networks. Routers must also discover when their neighbors become unreachable or inoperative. This process is achieved with low overhead by periodically sending small hello packets. As long as a router receives hello packets from a neighboring router, it assumes that the neighbor is functioning, and they can exchange routing information.
- *Reliable Transport Protocol* (RTP)—Responsible for guaranteed, ordered delivery of Enhanced IGRP packets to all neighbors. It supports intermixed transmission of multicast or unicast packets. For efficiency, only certain Enhanced IGRP packets are transmitted reliably. For example, on a multiaccess network that has multicast capabilities, such as Ethernet, it is not necessary to send hello packets reliably to all neighbors individually. For that reason, Enhanced IGRP sends a single multicast hello packet containing an indicator that informs the receivers that the packet need not be acknowledged. Other types of packets, such as updates, indicate in the packet that acknowledgment is required. RTP has a provision for sending multicast packets quickly when unacknowledged packets are pending, which helps ensure that convergence time remains low in the presence of varying speed links.
- *DUAL finite state machine*—Embodies the decision process for all route computations. It tracks all routes advertised by all neighbors. DUAL uses distance information to select efficient, loop-free paths and selects routes for insertion in a routing table based on feasible successors. A *feasible successor* is a neighboring router used for packet forwarding that is a least-cost path to a destination that is guaranteed not to be part of a routing loop. When a neighbor changes a metric or when a topology change occurs, DUAL tests for feasible successors. If one is found, DUAL uses it to avoid recomputing the route unnecessarily. When there are no feasible successors but there are neighbors advertising the destination, a recomputation (also known as a diffusing

computation) must occur to determine a new successor. Although recomputation is not processor intensive, it does affect convergence time, so it is advantageous to avoid unnecessary recomputations.

- *Protocol-dependent modules*—Responsible for network-layer protocol-specific requirements. For example, the *IP-Enhanced IGRP module* is responsible for sending and receiving Enhanced IGRP packets that are encapsulated in IP. IP-Enhanced IGRP is also responsible for parsing Enhanced IGRP packets and informing DUAL of the new information that has been received. IP-Enhanced IGRP asks DUAL to make routing decisions, the results of which are stored in the IP routing table. IP-Enhanced IGRP is responsible for redistributing routes learned by other IP routing protocols.

The consistent and superior performance of Enhanced IGRP relies on several new features:

- Packet types
- Neighbor tables
- Topology tables
- Route states
- Route tagging

Packet Types

Enhanced IGRP uses the following packet types:

- *Hello* and *acknowledgment*—Hello packets are multicast for neighbor discovery/recovery and do not require acknowledgment. An *acknowledgment* packet is a hello packet that has no data. Acknowledgment packets contain a nonzero acknowledgment number, and they are always sent using a unicast address.
- *Update*—Update packets are used to convey reachability of destinations. When a new neighbor is discovered, unicast update packets are sent, so the neighbor can build up its topology table. In other cases, such as a link cost change, updates are multicast. Updates are always transmitted reliably.
- *Query* and *reply*—Query and reply packets are sent when a destination has no feasible successors. Query packets are always multicast. Reply packets are sent in response to query packets to indicate to the originator that the originator does not need to recompute the route because there are feasible successors. Reply packets are unicast to the originator of the query. Both query and reply packets are transmitted reliably.
- *Request*—Request packets are used to get specific information from one or more neighbors. Request packets are used in route server applications and can be multicast or unicast. Request packets are transmitted unreliably.

Neighbor Tables

When a router discovers a new neighbor, it records the neighbor’s address and interface as an entry in the *neighbor table*. There is one neighbor table for each protocol-dependent module. When a neighbor sends a hello packet, it advertises a hold time, which is the amount of time a router treats a neighbor as reachable and operational. If a hello packet is not received within the hold time, the hold time expires, and DUAL is informed of the topology change.

The neighbor table entry also includes information required by RTP. Sequence numbers are employed to match acknowledgments with data packets. The last sequence number received from the neighbor is recorded so out-of-order packets can be detected. A transmission list is used to queue packets for possible retransmission on a per-neighbor basis. Round-trip timers are kept in the neighbor table entry to estimate an optimal retransmission interval.

Topology Tables

The *topology table* contains all destinations advertised by neighboring routers. The protocol-dependent modules populate the table, and the table is acted on by the DUAL finite state machine. Each entry in the topology table includes the destination address and a list of neighbors that have advertised the destination. For each neighbor, the entry records the advertised metric, which the neighbor stores in its routing table. An important rule that distance vector protocols must follow is that if the neighbor is advertising this destination, it must be using the route to forward packets.

The metric that the router uses to reach the destination is also associated with the destination. The metric that the router uses in the routing table and to advertise to other routers is the sum of the best advertised metric from all neighbors plus the link cost to the best neighbor.

Route States

A topology table entry for a destination can be in one of two states, active or passive. A destination is in the passive state when the router is not performing a recomputation or in the active state when the router is performing a recomputation. If feasible successors are always available, a destination never has to go into the active state, thereby avoiding a recomputation.

A recomputation occurs when a destination has no feasible successors. The router initiates the recomputation by sending a query packet to each of its neighboring routers. The neighboring router can send a reply packet, indicating it has a feasible successor for the destination, or it can send a query packet, indicating that it is participating in the recomputation. While a destination is in the active state, a router cannot change the destination’s routing table information. Once the router has received a reply from each neighboring router, the topology table entry for the destination returns to the passive state, and the router can select a successor.

Route Tagging

Enhanced IGRP supports internal and external routes. Internal routes originate within an Enhanced IGRP AS. Therefore, a directly attached network that is configured to run Enhanced IGRP is considered an internal route and is propagated with this information throughout the Enhanced IGRP AS. External routes are learned by another routing protocol or reside in the routing table as static routes. These routes are tagged individually with the identity of their origin.

External routes are tagged with the following information:

- Router ID of the Enhanced IGRP router that redistributed the route
- AS number of the destination
- Configurable administrator tag
- ID of the external protocol
- Metric from the external protocol
- Bit flags for default routing

Route tagging allows the network administrator to customize routing and maintain flexible policy controls. Route tagging is particularly useful in transit ASs where Enhanced IGRP typically interacts with an interdomain routing protocol that implements more global policies, resulting in a very scalable, policy-based routing.

Compatibility with IGRP

Enhanced IGRP provides compatibility and seamless interoperation with IGRP routers. An automatic redistribution mechanism allows IGRP routes to be imported into Enhanced IGRP and Enhanced IGRP routes to be imported into IGRP, so it is possible to add Enhanced IGRP gradually into an existing IGRP network. Because the metrics for both protocols are directly translatable, they are as easily comparable as if they were routes that originated in their own ASs. In addition, Enhanced IGRP treats IGRP routes as external routes and provides a way for the network administrator to customize them.