

IGRP

Cisco's Distance-Vector Thing

Interior Gateway Routing Protocol



- **Developed by Cisco in the mid-1980s as a better alternative to RIP**
 - ◆ Before RIPv2 appeared
- **Unequal path load sharing supported**
 - ◆ By default up to 4-way equal cost load sharing only
- **Also used for ISO CLNP**
- **Distance vector, similar to RIP**
 - ◆ On startup a request for immediate update is sent to all neighbors
 - ◆ Then routing table is sent periodically
 - ◆ DA=Broadcast

The Interior Gateway Routing Protocol (IGRP) is a dynamic distance vector routing protocol designed by Cisco in the mid-1980s for routing in an autonomous system that contains large, arbitrarily complex networks with diverse bandwidth and delay characteristics. Historically, IGRP became one of the success factors for the early Cisco IOS software capabilities because of its superiority to RIP version 1.

The important IGRP characteristics are:

- More scalability than RIP
- Fast response to network changes
- Sophisticated metric
- Multiple-path support across unequal path

Timers



- **Every 90 seconds the whole routing table is broadcasted to all neighbors**
 - ◆ **Except those routes suppressed by split horizon**
 - ◆ **Max 20% negative jitter to avoid synchronization**
 - ◆ **Effective update timer: 72-90s**
- **Invalid timer: 270 seconds**
- **Holddown timer: 280 seconds**
- **Flush timer: 630 seconds**

There are 4 critical time constants that control route propagation and expiration. These time constants may be set by the system administrator. However, there are default values. These time constants are:

Broadcast time -- updates are broadcast by all routers on all connected interfaces this often. The default is once every 90 sec.

Invalid time -- if no update has been received for a given path within this amount of time, it is considered to have timed out. It should be \ several times the broadcast time, in order to allow for the possibility that packets containing an update could be dropped by the network. The default is 3 times the broadcast time.

Hold time -- when a destination has become unreachable (or the metric has increased enough to cause poisoning), the destination goes into "holddown". During this state, no new path will be accepted for the same destination for this amount of time. The hold time indicates how long this state should last. It should be several times the broadcast time. The default value is 3 times the broadcast time plus 10 sec.

Flush time -- if no update has been received for a given destination within this amount of time, the entry for it is removed from the routing table. Note the difference between invalid time and flush time: After the invalid time a path is timed out and removed. If there are no remaining paths to a destination, the destination is now unreachable. However, the database entry for the destination remains. It has to remain in order to enforce the holddown. After the flush time the database entry is removed from the table. It should be somewhat longer than the invalid time plus the holddown time. The default is 7 times the broadcast time. Considering a maximum of 20% negative jitter the resulting update timer is something between 72 and 90 seconds.

Stability Measures



- **Count to infinity**
 - ♦ Based on hop count
 - ♦ If hop count increases (assumption: another router poisoned this route) then route is removed from table
- **Split horizon with poison reverse**
 - ♦ Poison reverse update sent if metric increased by **10%**
- **Triggered update with route poisoning**
 - ♦ If a route has failed, the hop count for that route is set to unreachable (255)
- **Holddown**
 - ♦ For unreachable routes
 - ♦ Also if metric increased
 - ♦ Turned off with "no metric holddown"

Normally, new routing tables are sent to neighboring routers on a regular basis (90 seconds). A triggered update is a new routing table that is sent immediately, in response to some change.

Each update triggers a routing table change in the adjacent routers, which, in turn, generate triggered updates notifying their adjacent neighbors of the change. This wave propagates throughout that portion of the network where routes went through the link. Triggered updates would be sufficient if we could guarantee that the wave of updates reached every appropriate router immediately. However, there are two problems:

1. Packets containing the update message can be dropped or corrupted by some link in the network.
2. The triggered updates do not happen instantaneously. It is possible that a router that has not yet received the triggered update will issue a regular update at just the wrong time, causing the bad route to be reinserted in a neighbor that had already received the triggered update.

Coupling triggered updates with holddowns is designed to get around these problems. Because the hold-down rule says that when a route is removed, no new route will be accepted for the same destination for some period of time, the triggered update has time to propagate throughout the network.

Hold-down timers are used to prevent regular update messages from inappropriately reinstating a route that may have gone bad. Hold-downs tell routers to hold 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.

IGRP Characteristics



- **Carried directly in IP**
 - ◆ **Protocol number 9**
- **IGRP "autonomous system" number is only a process number**
 - ◆ **Only routers with same process number can communicate via IGRP updates**
 - ◆ **Traffic between process domains can be exchanged by redistribution**

IGRP updates are encapsulated in ip packet directly, unlike RIP which uses UDP and then IP, means less overhead for IGRP.

AS number is not an "official" number like in BGP, it's used like a process number and it must be the same for all IGRP routers in one administrative domain. Using different numbers one can split network into different domains, by default there is no inter-exchange of routing information between different domains. Administrator can enable manual redistribution of routing information on the router which belongs to more than one domain.

Three Types of Routes (1)



IGRP flags "candidates" for a default route. Candidates are identified as routes to the edge of the network. The candidate with the lowest metric is selected as the default route.

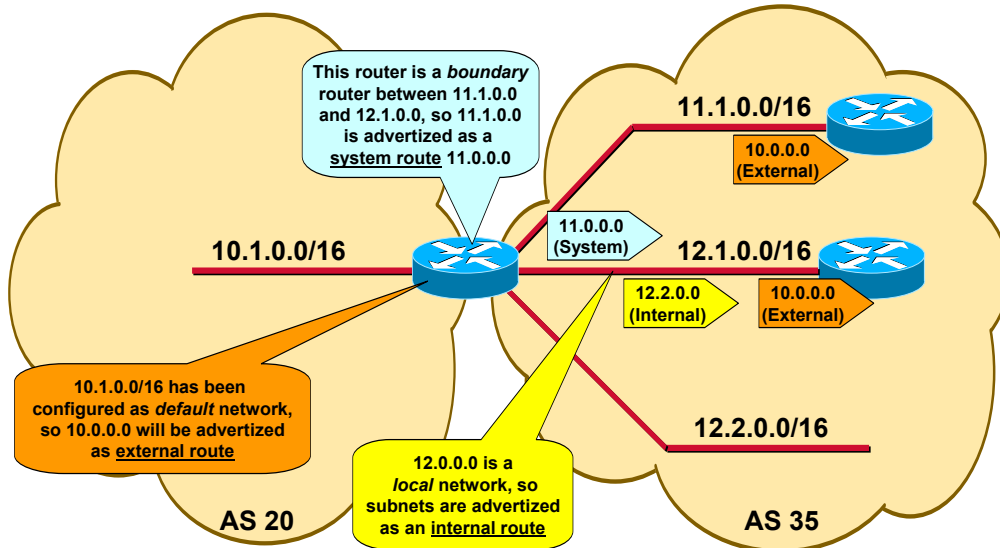
- **Interior routes**
 - ◆ Routes to subnets of local networks
- **System routes**
 - ◆ Summarized routes by network boundary router
- **Exterior routes**
 - ◆ Path to a default network

Note that an IGRP update message has three portions: interior, system (meaning "this autonomous system" but not interior), and exterior. The interior section is for routes to subnets. Not all subnet information is included—only subnets of one network are included. This is the network associated with the address to which the update is being sent.

Normally updates are broadcast on each interface, so this is simply the network on which the broadcast is being sent. (Other cases arise for responses to an IGRP request and point to point IGRP.) Major networks (i.e. non-subnets) are put into the system portion of the update message unless they are specifically flagged as exterior.

A network will be flagged as exterior if it was learned from another gateway and the information arrived in the exterior portion of the update message. Cisco's implementation also allows the system administrator to declare specific networks as exterior. Exterior routes are also referred to as "candidate default". They are routes that go to or through gateways that are considered to be appropriate as defaults, to be used when there is no explicit route to a destination.

Three Types of Routes (2)



IGRP is a distance vector protocol. In such a protocol, gateways exchange routing information only with adjacent gateways. This routing information contains a summary of information about the rest of the network.

IGRP Metrics (1)



- **Composite metric**
 - ♦ **Bandwidth**
 - ♦ **Delay**
 - ♦ Load (opt.)
 - ♦ Reliability (opt.)
 - ♦ MTU (opt.)
- **By default only Bandwidth and Delay is used**
 - ♦ Bandwidth = minimum BW on the route in kbit/s
- **Additionally smallest MTU is tracked**
- **Additionally Hop-Count is tracked**
 - ♦ Only used to limit network diameter
 - ♦ 0..255, default maximum hop-count = 100

Bandwidth and delay are the two metrics that are most commonly used. They also comprise the default metric.

The IGRP 24-bit metric field consists of:

- Bandwidth—Minimum bandwidth on the route in kilobits per second.
- Delay—Route delay in tens of microseconds.
- Reliability—Likelihood of successful packet transmission expressed as an integer from 0 to 255.
- Loading—Effective bandwidth of path.
- MTU—Minimum MTU in path, expressed in bytes.

IGRP Metrics (2)



- **Note: Bandwidth and Delay are static values**
 - ♦ Verified with "show interfaces"
- **IGRP-BW = $10^7/BW$**
 - ♦ Change it with "bandwidth" command
- **Delay is measured in units of 10 μ s**
 - ♦ Change it with "delay" command
- **Reliability is measured dynamically**
 - ♦ 1..255 where 255 means 100% reliable
- **Load is measured dynamically**
 - ♦ 1..255 where 255 means 100% loaded link

Note: Remember to define the bandwidth, especially on serial interfaces.

IGRP uses a 24-bit metric field:

Bandwidth—Minimum bandwidth on the route, in kilobits per second.

Delay—Route delay, in tens of microseconds.

Reliability—Likelihood of successful packet transmission, expressed as an integer from 0 to 255.

Loading—Effective bandwidth of path.

MTU—Minimum MTU in path, expressed in bytes.

Composite Metric



- **metric =**
$$\frac{(k1 * BW_{IGRP} + (k2 * BW_{IGRP}) / (256 - LOAD) + k3 * DLY_{IGRP}) * k5}{(RELIABILITY + k4)}$$
 - ♦ **Minimum BW_{IGRP} of all outgoing interfaces along the route to the destination**
 - ♦ **DLY_{IGRP} is sum of all delays of the route Default: $k1=k3=1, k2=k4=k5=0$**
 - ♦ **Changed with `metric weights` command**
- **Thus default metric = $BW_{IGRP} + DLY_{IGRP}$**
- **Unreachable routes are marked with $DLY_{IGRP} = 0xFFFFFFFF$**

Changing IGRP metrics can have great impact on network performance.

The following equation calculates the metric. It is presented for instructors and is not required to be taught.

$$\text{metric} = [k1 * \text{bandwidth} + (k2 * \text{bandwidth}) / (256 - \text{load}) + k3 * \text{delay}]$$

If $k5$ does not equal 0, an additional operation is done:

$$\text{metric} = \text{metric} * (k5 / (\text{reliability} + k4))$$

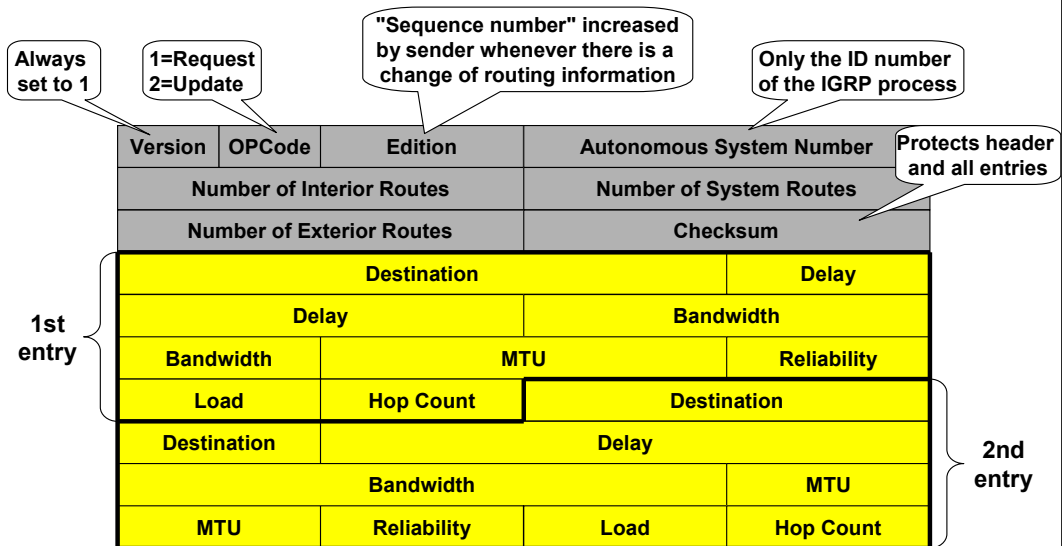
The default constant values are $k1=k3=1$ and $k2=k4=k5=0$.

Again, if default values are set, $\text{metric} = \text{bandwidth} + \text{delay}$.

The constants ($k1, k2, k3$) can be changed using the **metric weights** command. Changes to the IGRP constant values should be made with great care.

The load is a five-minute exponentially weighted average updated every five seconds. It is possible for the load value to rise fast enough to make a route unstable. If the load falls, the metric will decrease, which will cause a flash update.

Packet Format



The picture above shows the IGRP packet format. There is basically one general header (grey) and a number of routing sections following (yellow).

Packet Format Details



- **Three bytes are sufficient !**
 - ◆ Interior routes: byte 1 known, so only bytes 2, 3, and 4 are used
 - ◆ System routes: classful summarization
 - ◆ External routes: classful summarization
- **Up to 104 fourteen-byte entries**
 - ◆ Maximum packet size = 1468 bytes (header included)

Note:



- **variance 5**
 - ♦ Select routes whose metrics have a ratio of 5 at maximum
- **maximum-path 6**
 - ♦ Number of load-sharing paths
- **traffic-share min**
 - ♦ Use only (one) best route—but other routes remain in routing table!
- **INVALID=270 s, HOLDDOWN=280 s**
- **Exterior route**
 - ♦ Redistributed from another IGRP process
- **Interior route**
 - ♦ Directly connected network
- **System route**
 - ♦ Another route in same domain (not directly connected)