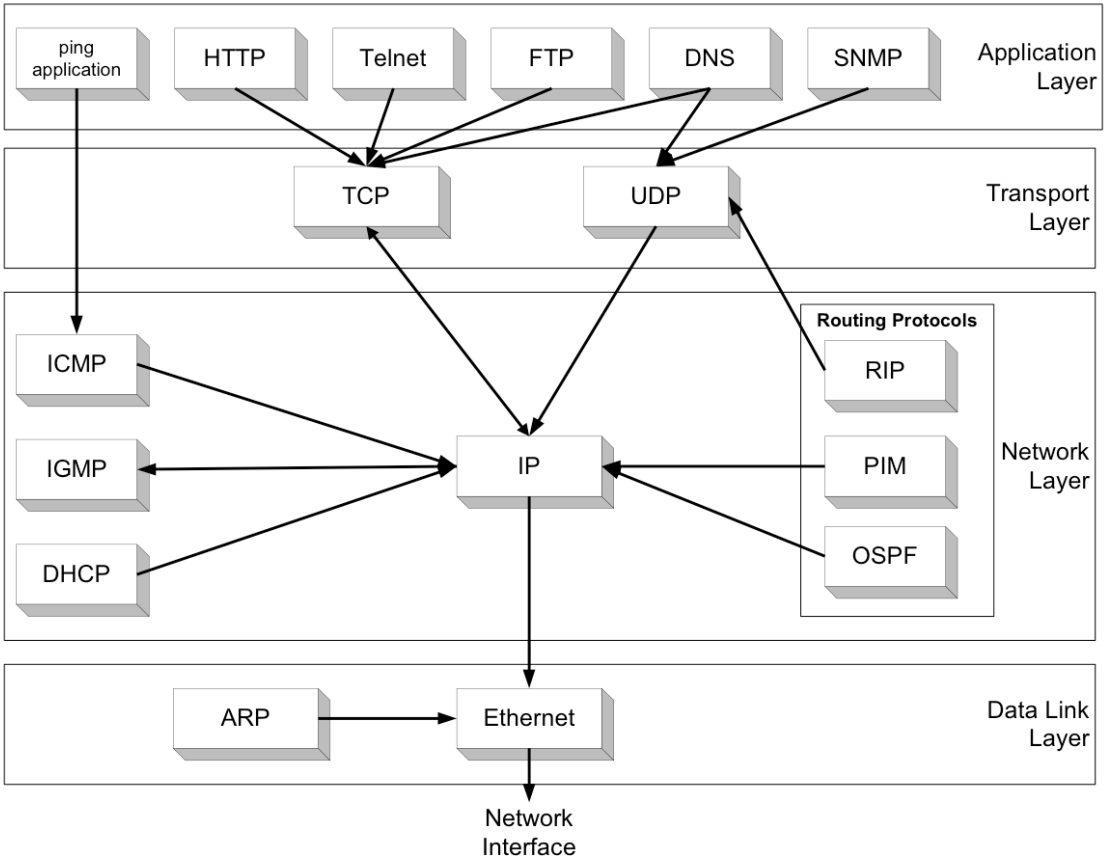
Proto



RIP Protocol(Routing Information Protocol (RIP) is an interior routing protocol (IGP) of the distance vector type (DVA).

Being simple to implement, this protocol is most often used in small networks. For IP, there are two versions of RIP - RIPv1 and RIPv2.

RIPv1 does not support masks. RIPv2 transmits information about network masks, so it is more compliant with today's requirements. Since the construction of routing tables in both versions of the protocol is not fundamentally different, in the future, to simplify the records, the operation of version 1 will be described.

To measure the distance to a network, RIP protocol standards allow various types of metrics: hops, throughput values, introduced delays, network reliability (i.e., corresponding to the D, T, and R attributes in the IP packet quality of service field), and any combination of these metrics. Most RIP implementations use the simplest metric — the number of hops, i.e., the number of intermediate routers that a packet must pass to reach the destination network.

BGP protocol(Border Gateway Protocol) – belongs to the class of external gateway routing protocols (EGP - External Gateway Protocol). The main implementations are Cisco IOS, Juniper JunOS, Bird, OpenBGPD, Quagga, Huawei VRP, Mikrotik RouterOS

Currently, BGP is the main dynamic routing protocol in the Internet between ASs.

1. Building the RIP routing table.

Let us consider the process of constructing a routing table using the RIP protocol using the example of a composite network shown in Fig. 1. We will divide this process into 5 stages.

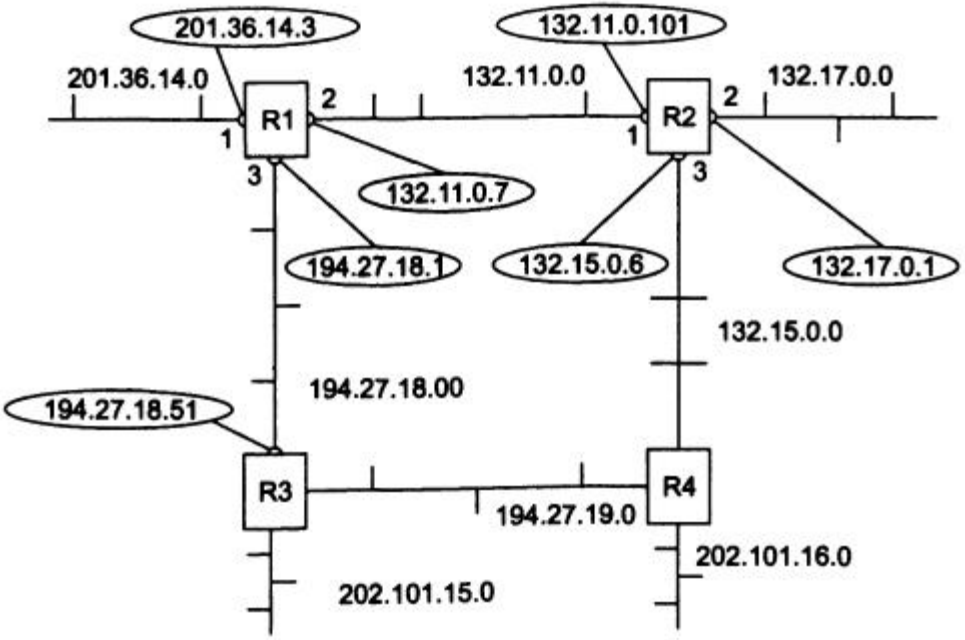


Fig. 1. Network built on RIP routers.

1.1. Creating a minimal table (initialization).

This composite network includes 8 IP networks connected by 4 routers with identifiers: Rl, R2, R3 and R4. In the figure, the router port addresses, unlike the network addresses, are placed in ovals.

In the default state, each router automatically has a minimal routing table created by the TCP/IP stack software that only takes into account directly connected networks. For example,

Table 1. Minimum routing table of router R1.

Network number	Next Router Address	Port	Distance
201.36.14.0	201.36.14.3	1	1
132.11.0.0	132.11.0.7	2	1
194.27.18.0	194.27.18.1	3	1

Table 2. Minimum routing table of router R2.

Network number	Next Router Address	Port	Distance
132.11.0.0	132.11.0.101	1	1
132.17.0.0	132.17.0.1	2	1
132.15.0.0	132.15.0.6	3	1

1.2. Sending the minimum table to neighbors.

After initialization, each router begins sending RIP messages to its neighbors, which contain its minimum table. RIP messages are transmitted in UDP datagrams and include two parameters for each network: its IP address and its distance from the router transmitting the message.

In relation to any router, neighbors are those routers to which this router can transmit an IP packet over any of its networks without using the services of intermediate routers. For example, for router R1, neighbors are routers R2 and R3, and for router R4, neighbors are routers R2 and R3.

Thus, router R1 sends the following messages to routers R2 and R3:

- net201.36.14.0, distance 1;
- net132.11.0.0, distance 1;
- net194.27.18.0, distance 1.

1.3. Receiving RIP messages from neighbors and processing them.

After receiving similar messages from routers R2 and R3, router R1 increments each received metric field by one and remembers through which port and from which router the new information was received. The router then begins to compare the new information with that stored in its TM (see Table 3).

Network number	Next Router Address	Port	Distance
201.36.14.0	201.36.14.3	1	1
132.11.0.0	132.11.0.7	2	1
194.27.18.0	194.27.18.1	3	1
132.17.0.0	132.11.0.101	2	2
132.15.0.0	132.11.0.101	2	2
194.27.19.0	194.27.18.51	3	2
202.101.15.0	194.27.18.51	3	2
~~132.11.0.0~~	~~132.11.0.101~~	~~2~~	~~2~~
~~194.27.18.0~~	~~194.27.18.51~~	~~3~~	~~2~~

Entries 4 through 9 are received from neighboring routers and they are candidates for placement in the TM. However, only entries 4 through 7 are included in the table, while entries 8 and 9 are not. This is because they contain data on networks already in TM R1, and the distance to them is greater (or equal) than in the existing entries.

RIP replaces a network entry only if the new information has a better metric (smaller) than the existing one or the entry that arrived at the router remains.**first in time**. (There is an exception - if the worst information about the network came from the same router, on the basis of whose message the given entry was created, then the worst information replaces the best. Similar operations with the new information are performed by the other routers of the network.

1.4. Sending a new table to neighbors.

Each router sends a new TM as a RIP message to all its neighbors. In this message, it places data on all networks known to it: both directly connected and those the router learned about from RIP messages.

1.5. Receiving RIP messages from neighbors and processing them.

Step 5 is a repeat of step 3 - routers receive RIP messages, process the information they contain, and adjust their routing tables based on it.

Let's see how router R1 does this (Table 4).

At this stage, router R1 receives information about the network 132.15.0.0 from router R3, which it, in turn, received from R3 during the previous work cycle.

router R4. The router already knows about the network 132.15.0.0, and the old information has a better metric than the new one, so the new information about this network is discarded.

At this stage, router R1 learns about network 202.101.16.0 for the first time, and data about it comes from two neighbors - from R3 and R4. Since the metrics in these messages are the same, the data that arrived first is included in the table. In our example, it is considered that router R2 was ahead of router R3 and was the first to forward its RIP message to router R1.

Table 4. Routing table of router R1.

Network number	Next Router Address	Port	Distance
201.36.14.0	201.36.14.3	1	1
132.11.0.0	132.11.0.7	2	1
194.27.18.0	194.27.18.1	3	1
132.17.0.0	132.11.0.101	2	2
132.15.0.0	132.11.0.101	2	2
~~132.15.0.0~~	~~194.27.18.51~~	~~3~~	~~3~~
194.27.19.0	194.27.18.51	3	2
104.27.10.0	132.11.0.101	2	3
202.101.15.0	194.27.18.51	3	2
202.101.16.0	132.11.0.101	2	3
~~202.101.16.0~~	~~104.27.18.51~~	~~3~~	~~3~~

If routers periodically repeat the stages of sending and processing RIP messages, then in a finite time the correct routing mode will be established in the network.

The correct routing mode here means a state of the routing tables when all networks are reachable from any network using some rational route. Packets will reach their destinations and will not get stuck in loops like the one formed in Fig. 1 by routers R1, R2, R3, and R4.

2. Adaptation of RIP routers to changes in network conditions.

Obviously, if all routers, their interfaces and the communication lines connecting them remain operational in the network, then RIP announcements can be made quite rarely, for example, once a day. However, changes occur in networks constantly - the performance of routers and communication lines changes, and routers and communication lines can be added to an existing network or removed from it.

One of the problems with RIP is slow convergence, meaning that changes that occur in one part of the Internet are propagated very slowly through the rest of the Internet. RIP uses a number of mechanisms to adapt to changes in the network.

Towards new routesroutersRIPs adapt simply - they transmit new information in the next message to their neighbors and gradually this information becomes known to all routers in the network.

But as for the changes associated with**loss of any route**, RIP routers are more difficult to adapt to. This is because there is no field in the RIP message format that would indicate that the path to a given network no longer exists.

To notify that a route is invalid, use:

- route lifetime expirationTTL;
- specifying a special (infinite) distance to the network that has become unavailable.

2.1. Route Lifetime Expiration Mechanism.

- Based on the fact that each TM record received according to the protocolRIP or from scanning its own interfaces, has a lifetime (**TTL route**).
- FromTTL is subtracted by one every second. If no new message about this route arrives during the timeout, it is marked as invalid (removed from TM).
- Upon completionTTL of records about own interfaces is their initialization.
- Upon receipt of the nextRIP message that confirms the validity of this record, TTL is set to its initial state.
- The timeout value is set to180 seconds, and the broadcast period is approximately 30 seconds (in reality, a random number between 25 and 35 is used, which is done to prevent routers from synchronizing when sending updates).
- A six-fold timeout reserve is needed to ensure that the network actually became unavailable, and not just because RIP messages were lost (which is possible, since the RIP protocol uses the unreliable UDP:520 transport protocol).
- If any router fails, it stops sending messages to its neighbors about the networks that can be reached through it, and after 180 seconds all records generated by this router will become invalid for its nearest neighbors.
- After this, the process will be repeated for the closest neighbors - they will cross out similar ones records after 360 seconds. For the third neighbor - after 540 seconds (almost 10 minutes), etc.

As you can see, information about networks whose paths can no longer pass through the failed router does not spread across the network very quickly. This is one of the reasons for choosing a small value of 30 seconds as the distribution period.

2.2. Mechanism for sending infinite distance to the network.

The TTL mechanism works in cases where a router cannot send a message to its neighbors about a failed route, because either it itself is inoperative, or the communication line through which the message could be transmitted is inoperative.

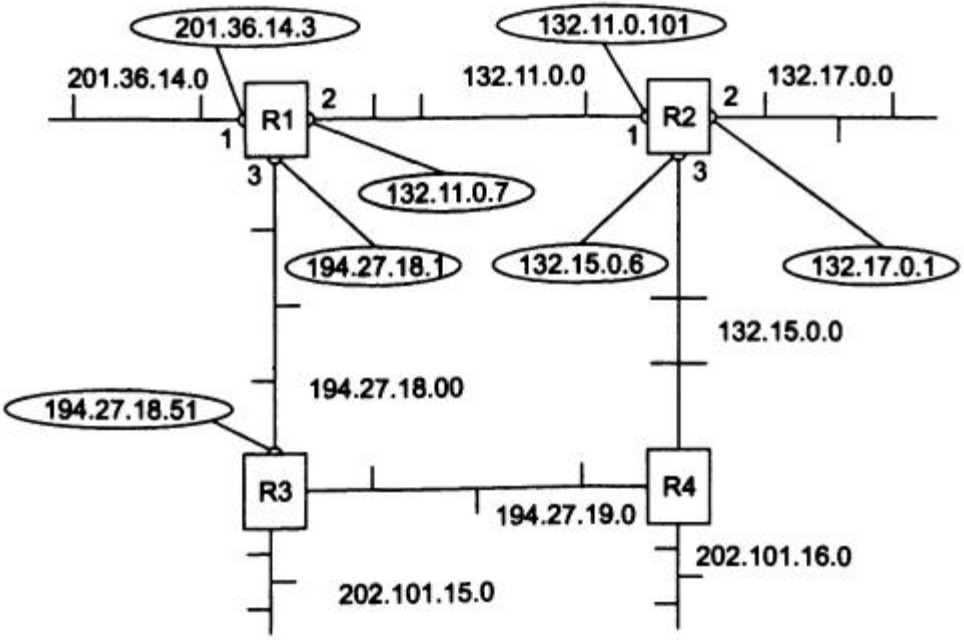
When a message can be sent, RIP routers use a technique that involves specifying an infinite distance to the network that has become unavailable.

- In the protocolRIP infinite is conventionally considered to be a distance of 16 hops. The reason for choosing such a small number as an "infinite" distance is that in some cases, failures of connections in the network cause long periods of incorrect operation of RIP routers, expressed in looping of packets in network loops. And the smaller the distance used as "infinite", the shorter such periods.
- Having received a message in which the distance to a certain network is16 (or 15, which yields the same result, since the router increments the received value by 1), the router must check whether this "bad" information about the network comes from the same router whose message served as the basis for the entry about this network in the routing table. If it is the same router, then the information is considered reliable and the route is marked as unavailable.

3. Route loops in RIP.

3.1. Example of a routing loop in RIP.

Let us consider the case of packet looping using the example of the network shown in Fig.



3.1.1 Let router R1 detect that its connection to the directly connected network 201.36.14.0 has been lost (for example, due to the failure of interface 201.36.14.3).

3.1.2 Router R1 notes in its routing table that**net201.36.14.0 is unavailable.**At best, he will discover this just before sending it.RIP messages, and in the worst case, immediately after sending the next RIP messages, so that on average 15 seconds remain before the start of a new cycle of its announcements, in which it must inform its neighbors that the distance to the network 201.36.14.0 has become equal to 16.

3.1.3. Each router operates on its own timer, without synchronizing its work on sending announcements with other routers. Therefore, it is likely that**50%, router R2 will outpace router R1**and will give him his message (see Table 5) before R1 has time to transmit news about the unreachability of network 201.36.14.0.

Table 5. Routing table of router R2.

Network number	Next Router Address	Port	Distance
201.36.14.0	132.11.0.7	1	2

This entry, received from router R1, was correct before interface 201.36.14.3 failed; now**she's outdated**, but router R2 doesn't know about it.

3.1.4. Router R1 receives new network information**201.36.14.0 - this network is supposedly reachable via router R2 with metric 2**.

Previously, R1 ignored this information because its own metric was better. Now R1 must accept the information about network 201.36.14.0 received from R2 and replace the routing table entry about this network being unreachable (Table 6).

Table 6. Routing table of router R1.

Network number	Next Router Address	Port	Distance
201.36.14.0	132.11.0.101	2	3

3.1.5. As a result, a routing loop is formed in the network:packets sent to nodes network 201.36.14.0 will be forwarded by router R2 to router R1, and router R1 will return them to router R2.

IP packets will continue to circulate around this loop until the lifetime of each IP packet (255 hops) expires.

3.2. Lifetime of a route loop.

Let's consider time periods that are multiples of the lifetime of entries in router tables.

- Time0-180 sec. After the interface fails, R1 and R2 will have invalid entries. R2 still supplies R1 with its entry for network 201.36.14.0 with metric 2, since its time to live has not expired. Packets are in a loop.

- Time180-360 s. At the beginning of this period, router R2's entry about network 201.36.14.0 with metric 2 expires, because router R1 sent it messages about network 201.36.14.0 with metric worse than R2's in the previous period, and they could not acknowledge this entry. Now router R2 receives from router R1 an entry about network 201.36.14.0 with metric 3 and transforms it into an entry with metric 4. Router R1 does not receive any new messages from router R2 about network 201.36.14.0 with metric 2, so the entry's TTL starts decreasing. Packets continue to loop.

- Time360-540 sec. Router R1's entry for network 201.36.14.0 with metric 3 expires. Routers R1 and R2 switch roles again - R2 supplies R1 with outdated information about the path to network 201.36.14.0, now with metric 4, which R1 converts to metric 5. Packets continue to loop.

If RIP had not chosen**distance16 as unattainable**, then the described process would last a very long time (until the field's discharge grid was exhausted

distance, and with the next increase in distance, an overflow would be recorded).

As a result, router R2 at the next stage of the described process receives metric 15 from router R1, which after increasing, turning into metric 16, recording the unreachability of the network. Thus, in our example, the period of unstable network operation lasted 36 minutes!

The 15-hop limitation narrows the scope of the RIP protocolto the networks in which The number of intermediate routers cannot be more than 15. For larger networks, other routing protocols, such as OSPF, must be used, or the network must be divided into autonomous areas.

Given example Fine illustrates home reason instability routers operating under the RIP protocol. This reason is rooted in the very principle of operation of distance vector protocols - the use of information obtained "second-hand". Indeed, router R2 transmits information to router R1 about the reachability of network 201.36.14.0, for the reliability of which it itself is not responsible.

NOTE.

It should not be thought that any failure of interfaces and routers in networks causes routing loops. Routing loops, even without additional methods of combating them, occur on average in no more than half of the potentially possible cases.

4. Methods of combating false routes in the RIP protocol.

Even when combating false routes, the RIP protocol is not able to completely eliminate transient states in the network, but only reduces such problems.

4.1. Split horizon.

With this method, information about reachable network nodes is transmitted to all routers except the one from which this information originated.

If router R2 supported split horizon, it would not have passed outdated information about network 201.36.14.0 to router R1, since it received it from router R1.

Split horizon does not help when loops are formed not by 2, but by a larger number of routers.

For example, if router R1 loses connection to network 201.36.14.0, routers R2 and R3 will not return to R1 information about network 201.36.14.0 with metric 2, since they received this information from R1. However, they will return to R1 information about the reachability of network 201.36.14.0 with metric 4, since they received this information from R4.

To prevent packet loops**by compound loops**When links fail, techniques called trigger updates and freezing changes are used.

4.2 Trigger updates.

The reception of trigger updates consists in the fact that the router, having received data about a change in the metric for any network, does not wait for the expiration of the routing table transmission period, but**transmits data about the changed route immediately.**

This technique can prevent the transmission of outdated information about a failed route in many cases, but it**overloads the network with service messages, therefore Trigger ads are also made with some delay.**For this reason a situation is possible when a regular update in some router slightly precedes the arrival of a trigger update from the previous router in the chain, and this router manages to transmit outdated information about a non-existent route over the network.

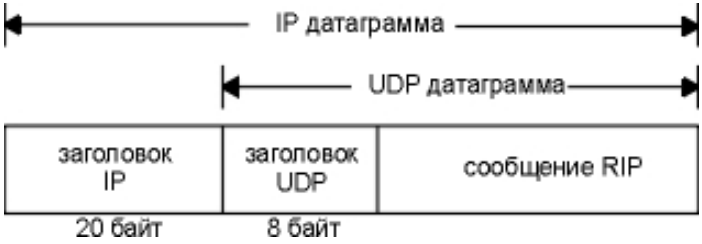
4.3 Freezing changes.

The freezing of changes technique is associated with**introducing a timeout for the adoption of new data about a network that has just become unavailable.**

This timeout protects R1 from accepting stale information from routers some distance away. During the freeze timeout, these routers are expected to remove the route from their tables because they are not receiving new confirmations about it and will not propagate stale information across the network.

5. RIP packet format.

The official specification of the RIPv1 protocol is RFC 1058 [Hedrick 1988]. RIP messages are carried in UDP:520 datagrams, as shown in the figure.



5.1 RIPv1 Request-Response Message Format.

Octet	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
00	**Team (1-6)**								**Version (1)**								**(Must be 0)**															
04	**Address family (0 or 2)**																**(Must be 0)**															
08	**32-bit network (node) address**																															
12	**(Must be 0)**																															
16	**(Must be 0)**																															
20	**Index (1-16)**																															
24-n	**Up to 24 routers with the same format as the previous 20 bytes (04 to 24)**																															

Field commandequals1 is a request, if 2 is a response. There are two more command field values (3 and 4), and two undocumented values: poll (5) and poll item (6). The request contains a request to the other system to send all or part of its routing table. The response contains all or part of the sender's routing table.

Field versionusually installed in1, however for RIP Version 2 (section "RIP Version 2") this value is set to 2.

Address family- always equal2 for TCP/IP addresses.

IP address of the network/node- defines the destination address.RIP allocates 14 bytes for this field when applied to any protocol. However, IP currently uses only 4 bytes. The remainder of the address is filled with zeros.

Indicator- as an indicatorRIP stands for Relay Interchange Propagation.

Up to 25 routes can be advertised in a single RIP message. The limit of 25 is determined by the total size of the RIP message, 20x25+4=504, less than 512 bytes. Because of the limit of 25 routes, one request typically requires multiple responses to be sent in order to transmit the entire routing table.

5.2. Operation of the routed daemon process.

Let's see how routed (the Linux routing daemon) typically works using RIP.

- **Initialization**. When the daemon starts, it detects all active interfaces and sends a request packet to each interface. The request is sent as a broadcast message if the network supports it. The destination port is UDP:520 (the routing daemon on the other router). The characteristics of such a request are: command field = 1, address family field = 0, and index = 16. This format corresponds to**special request**, in response to which it is required to send a full TM.

- **Request accepted**. In the case of a special request, which we have just described, the full routing table is sent to the requestor. A response is returned.
- **Answer accepted**. If the answer is found to be correct, the routing table can be updated. New entries can be added, existing entries can be modified or deleted.

- **Regular update**routing. Every30 seconds, all or part of the routing table is sent to each neighboring router. The routing table is broadcast (in the case of Ethernet) or sent to the other end of a point-to-point link.

- **Unplanned update**. Occurs when a route indicator changes. In this case, there is no need to send the entire routing table, only the entry that has changed is transmitted.

5.3. RIP Version 2.

RFC 1388 [Malkin 1993] defines extensions to RIP functionality that are commonly referred to collectively as RIPv2.

These extensions do not change the protocol, but add additional information to the fields marked "must be zero" in the figure above.

RIP and RIPv2 can interoperate if RIP ignores fields that should be set to zero.

RIPv2 message format.

Octet	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
00	**Team (1-6)**								**Version (2)**								**Routing domain**															
04	**Address family (0 or 2 or 0xFFFF)**																**Route flag**															
08	**32-bit IP address of the network (node)**																															
12	**32-bit subnet mask**																															
16	**32-bit next hop IP address**																															
20	**Index (1-16)**																															
24-n	**Up to 24 routers with the same format as the previous 20 bytes (04 to 24)**																															

Version field ForRIPv2 is set to 2.

Routing domain- is the identifier of the routing daemon to which this packet belongs to. On Unix implementations, this should be the process ID of the daemon. This field allows an administrator to run RIP on the same router multiple times, each operating with the same routing domain.

RIPv2 support external router protocols.**Route indicator field** is intended to store the autonomous system number for EGP and BGP.

RIPv2 supports**Subnet mask**for each item corresponds to its own IP address.

RIPv2 supports**Next hop IP address**- ThisThe IP address of the destination to which packets should be sent. A value of 0 in this field means that packets should be sent to the system that sent the RIP message.

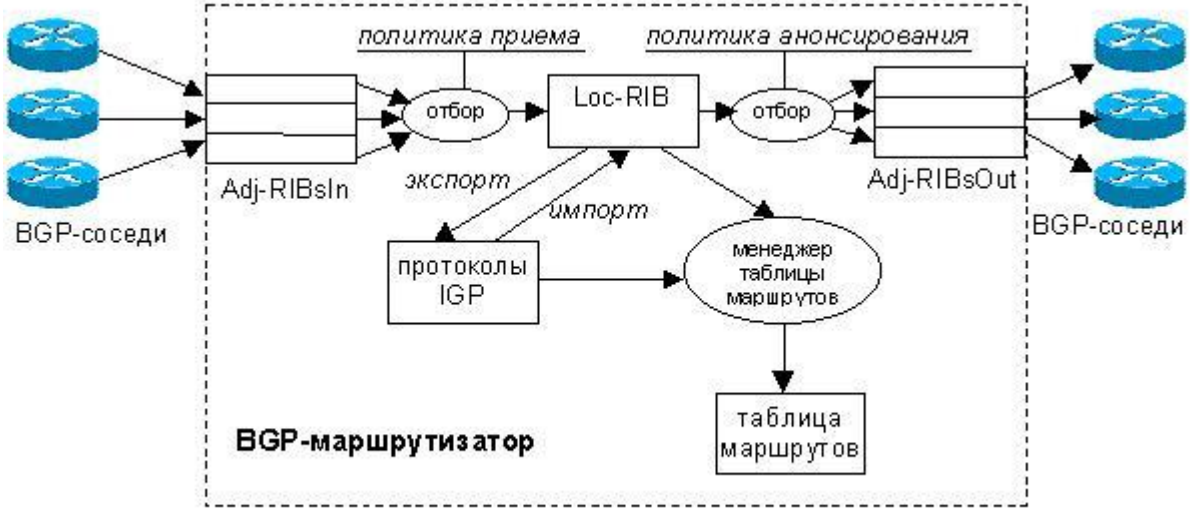
RIPv2 supports**authentication,**which protects messages against unauthorized advertisement. For this purpose, RIPv2 uses the first 20 bytes of the RIP message entry: Address family = 0xffff, Route flag = 2 determines the method used for authentication. The remaining 16 bytes contain the actual authentication data (the password in clear text).

RIPv2 supports**group requests**in addition to broadcast messages. This reduces the load on hosts that do not receive RIP-2 messages.

6. BGP protocol.

6.1. General scheme of work.

Routers of neighboring ASs that decide to exchange routing information establish connections between themselves via the BGP protocol and become neighbors (BGP peers).



BGP uses the path vector algorithm, which is an evolution of DVA. BGP neighbors advertise path vectors to each other. A path vector, unlike

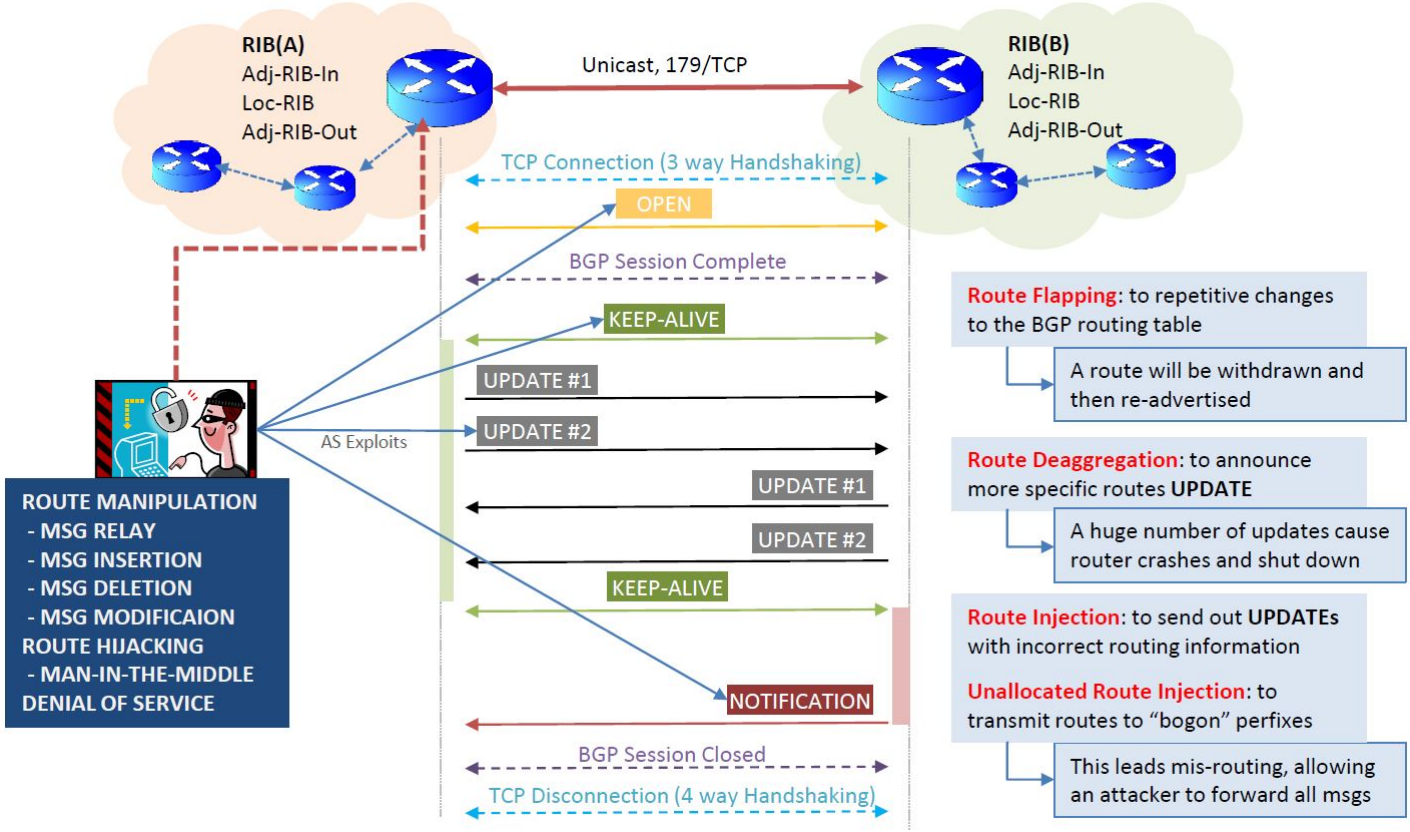
distance vector, contains not just the network address and the distance to it, but the network address and a list of attributes (path attributes) describing various characteristics of the route from the sender router to the specified network.

The data contained in the path attributes must be sufficient for the receiving router, after analyzing them in terms of**politicians**its AS, could make a decision on the acceptability or unacceptability of the received route.

6.2. Implementation of BGP.

A pair of BGP neighbors establishes a connection between themselves via TCP protocol, port 179.

The flow of information exchanged between BGP neighbors over TCP consists of a sequence of BGP messages. The maximum length of a message is 4096 octets, the minimum is 19.



BGP operation diagram.

6.3 BGP Message Types.

There are 4 types of messages.

OPEN - sent after a TCP connection has been established. The response to OPEN is a KEEPALIVE message if the other side agrees to become a BGP neighbor; otherwise, a NOTIFICATION message is sent with a code explaining the reason for the refusal, and the connection is terminated.

KEEPALIVE - the message is intended to confirm the consent to establish neighbor relations, as well as to monitor the activity of an open connection: for this purpose, BGP neighbors exchange KEEPALIVE messages at certain intervals.

UPDATE - the message is intended for announcing and revoking routes. After establishing a connection, all routes that the router wants to announce to the neighbor are sent using UPDATE messages (full update), after which only data on added or deleted routes are sent as they appear (partial update).

NOTIFICATION - this type of message is used to inform the neighbor about the reason for closing the connection. After sending this message, the BGP connection is closed.

6.4 BGP message format.

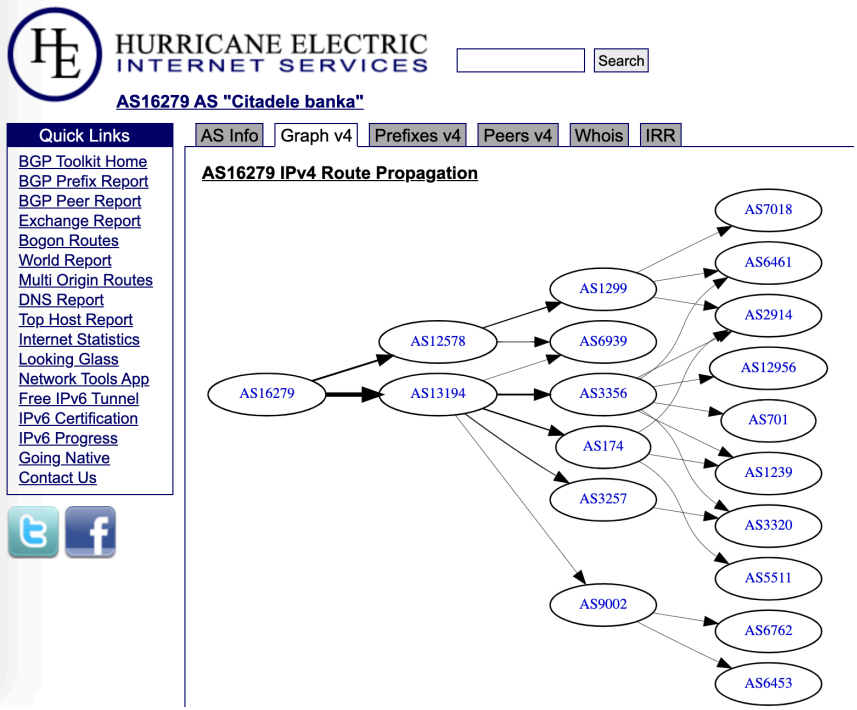
A BGP message consists of a header and a body. The header is 19 octets long and consists of the following fields:

Формат BGP-сообщения

- marker: in messageOPEN is always, and when working without authentication - in other messages, filled with ones. Otherwise, it contains authentication information. An accompanying function of the marker is to increase the reliability of identifying the message boundary in the data stream.
- message length in octets, including header.
- message type:
    - 1 - OPEN
    - 2 - KEEPALIVE
    - 3 - UPDATE
    - 4 - NOTIFICATION

6.5. AS BGP propagation graph example

1. Find Latvian ASNs Report https://bgp.he.net/report/world
2. Find AS Citadele Banka AS16279 https://bgp.he.net/AS16279
3. Look BGP Peers on Graph v4 (“Path to Internet”):
   - AS16279->AS12578->AS6939 (Hurricane Electric)
   - AS16279->AS13194->AS174 (Cogent Communication)
4. Read AS Info about all Ass
   - Company Name & Origin Country
   - Company Website & Network Map
   - Internet Exchanges Nrs
   - Prefixes Originated Nrs
   - Prefixes Announced Nrs
   - AS Paths Observed Nrs
5. Find CAIDA AS Rank on site https://asrank.caida.org/



7. Exercises.

1. What is the size of a RIP message that advertises only one network? What is the size of a RIP message that advertises N networks? Derive a formula that shows the relationship between the number of advertised networks and the size of the RIP message.
2. The router has the following RIP routing table:

Network1	4	B
Network2	2	C
Network3	1	F
Network4	5	G

What will be the contents of the table if the router receives the following RIP message from router C:

Network1	2
Network2	1
Network3	3
Network4	7

3. Using Figure 1, show the router link announcement for router R3.
4. Using Figure 1, show the router link announcement for router R4.