Network Working Group J.L. Lan
Internet-Draft J.H. Zhang
Intended status: Informational B.Wang
Expires: October 12, 2017 W.F. Liu
Y.J. Bu
National Digital Switching System Engineering and Technological Research Center, P.R.China
X. Li
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS
April 12, 2017

Node Potential Oriented Multi-NextHop Routing Protocol
draft-lanzhangwang-rtgwg-npmnrp-03

Abstract

The Node Potential Oriented Multi-Nexthop Routing Protocol (NP-MNRP) bases on the idea of "hop-by-hop routing forwarding, multi-backup next hop" and combines with the phenomena that water flows from higher place to lower. NP-MNRP defines a metric named as node potential, which is based on hop count and the actual link bandwidth, and calculates multiple next-hops through the potential difference between the nodes.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on October 12, 2017.

Table of Contents

1. Introduction 2
2. Terminology 3
3. Protocol operation 5
3.1. LFBs Network prefix information advertisement 5
3.2. Neighbor Node Discovery and Adjacency establishment 7
3.3. Calculation of node potential value 8
3.4. network event sensing 10
3.5. Update of the node potential 12
4. Routing information base 14
5. NP-MNRP packets format 15
5.1. Encapsulation of NP-MNRP packets 15
5.2. Packet format 15
6. Informative References 24
7. Author's Address 24

1. Introduction

The inspiration of this routing protocol comes from the natural water flow, which is a phenomenon of potential energy driven. Water flows through all feasible channels from high potential site to low potential site. If in every router all feasible next-hops for one certain destination can be found, then all the packets to that destination can be transferred in many paths.

This routing protocol obeys the Internet philosophy of "hop-by-hop routing paradigm" and enhances it with "multiple feasible next-hops". That means protocol try to calculate multiple next-hops for each node and not find multi-path end to end. The forwarding direction of packets is constrained by routing metrics. Different packets to one

destination can be distributed into multiple next-hops in parallel, which can make full use of all feasible next hops, so that the link utility ratio is globally even.

Compared with nature water flow phenomenon, the routing protocol defines a routing metric named as node potential in communication networks. Node potential is a parameter to measure the ability to transfer packets from one certain node to its destination. So the routing protocol is named as node potential oriented multiple next-hops routing protocol (NP-MNRP in short).

NP-MNRP is designed to calculate multiple next-hops for one destination in network without knowing the globally topology information. Distance vector routing protocols such as RIP and BGP are used for reference. NP-MNRP protocol is consisted of network reachable information advertisement process and node potential calculation process. But unlike those protocols, NP-MNRP ensures the following properties.

(1) Multiple feasible next-hops for each destination. When node finds the main next-hop unreachable, it chooses one from the backup next-hops to forward the packet.

(2) Accelerate the convergence speed and reduce route flap. Compared with the single shortest-path routing, NP-MNRP can effectively avoids congestion and shorten protection switching delay. The path establishment and maintenance can be accomplished by routing information advertisement mechanism between peers. The control overhead of the protocol is smaller than other multi-path routing protocols. When kinds of concurrent failures occur in network, NP-MNRP can provide more feasible next-hops and high recovery probability.

2. Terminology

Carrier Network: The network in which the nodes run Node Potential Oriented Routing Protocol.

User Network: The network that needs message transfer service provided by Carrier Network.

Boundary Node: According to the location of nodes in the carrier network, the nodes are divided into boundary nodes and intermediate nodes. Boundary nodes are also divided into two categories according to the flow direction. In terms of user network flow, the node through which flow into the carrier network, is called ingress node; the node through which flow out of the carrier network, is called egress node.

Intermediate Node: The other nodes in the carrier network except the Boundary Nodes.

Potential Reference Node: The egress nodes which the user network traffic flows out from the carrier network. In the route calculation process, these nodes will be set its node potential value to zero and be act as reference node in node potential computation procedure. Network Layer Diagram: The diagram that describes the distribution of the layer value of every node relative to a destination node.

Node Potential: A metric that measures the accessibility from a node to zero potential nodes in the carrier network.

Router ID: A 32-bit number assigned to each router running the NP-MNRP protocol. This number uniquely identifies the router in the carrier network. One algorithm for Router ID assignment is to choose the largest or smallest IP address assigned to the router.

Set of Available Next-hops: The set of available next-hops of node i when it sends packets to the destination node j, which is denoted as A(i,j).

Network Layer Reachable Information Flood Advertisement Packet (NLRI-FAP): The packet is used to advertise the information of binding relations between network prefix and egress nodes to its neighbors. IP address and subnet mask of its binding user network are included in this packet.

Network Layer Reachable Information Specific Request Packet (NLRI-SRP): The packet is used to request a particular egress node for certain prefix binding relationship information. IP address and subnet mask of user network whose binding relationship needs to know are included in this packet.

Network Layer Reachable Information Direction Answer Packet (NLRI-DAP): The packet is used to respond to a Network Layer Reachable Information Specific Request Packet. IP address and subnet mask of user network and its bound router ID are included in this packet.

Network Layer Graph Construction Trigger Packet (NLG-CTP): The packet is used to trigger the building of network layer diagram.

Network Layer Graph Information Advertisement Packet (NLG-IAP): The packet is used to advertise its layer information relative to one egress node to its neighbors. The egress node ID and its layer value relative to this egress node is included in this packet.

Network Layer Graph Routing Request Packet (NLG-RRP): The packet is

used to request the neighbors to update it's layer information relative to one egress node and this egress node ID is included in this packet.

Network Layer Graph Request Update Packet (NLG-RUP): The packet is used to inform the nodes which are on a particular layer relative to one destination node, to rebuild their layer information and that layer value is included in this packet.

Link State Detect Packet (Detect): The packet is used to detect the link status between itself and its neighbors. The sequence number of this packet and the period for sending Detect are included in this packet.

REPLY Packet (REPLY): The packet is used to respond to the Detect indicating the good link state. The sequence number of this packet and the period for sending Detect are included in this packet.

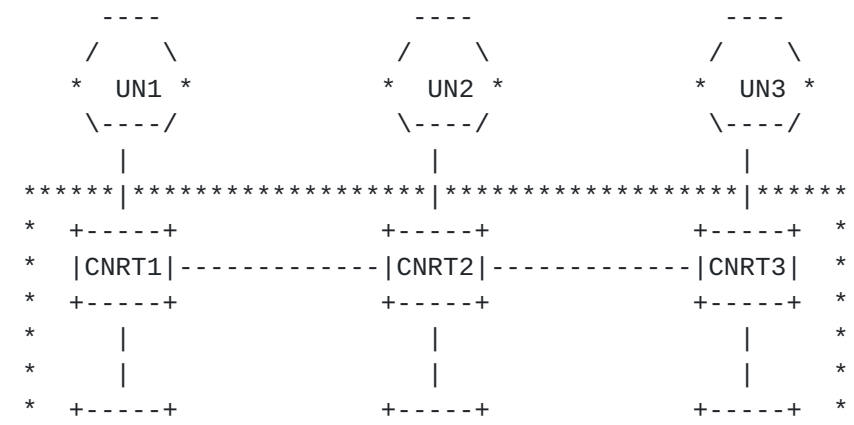
3. Protocol operation

3.1. LFBs Network prefix information advertisement

Because NP-MNRP protocol is designed as an IGP inside the autonomous system, it should route transit network traffic and local network traffic. The network IP prefix is divided into two parts: the IP prefix information of transit network and the IP prefix information of the network running NP-MNRP protocol own.

In view of NP-MNRP protocol, network is divided into user network and carrier network, as shown in the figure 1. User network is abbreviated as UN and carrier network is abbreviated as CN. Each node in CN will run NP-MNRP protocol to establish multiple next-hops routing information base. Otherwise, it is not necessary for nodes in UN to run NP-MNRP protocol.

CNRT: Carrier Network Router



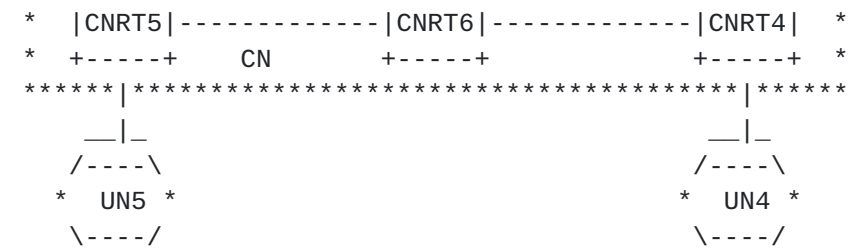


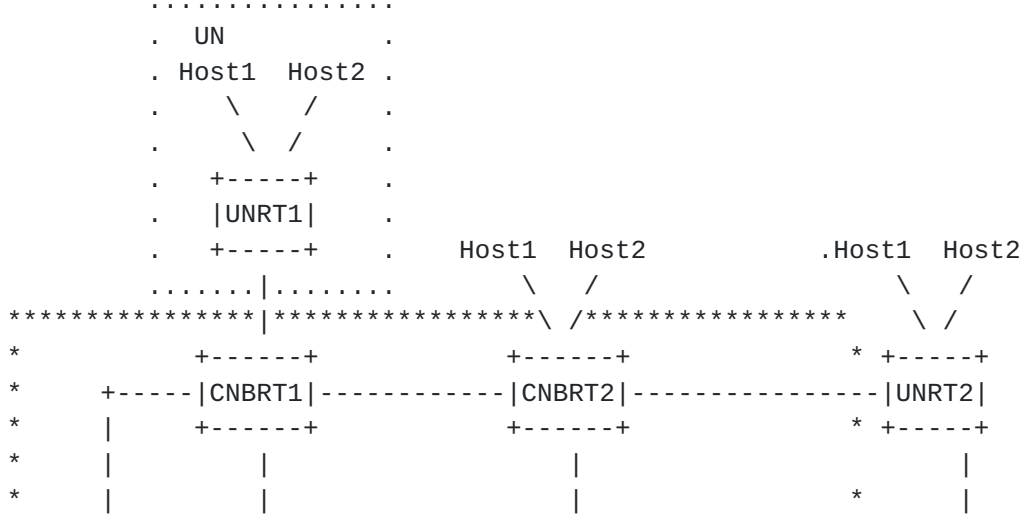
Figure 1: Network Model

The nodes in UN will finish local communication and transmit the network traffic whose destinations are not in the user network itself to the Carrier network. The network prefix of UN is named as Network Layer reachable information. Carrier network mainly transfer transit traffic for UN. The network prefix of CN is treated as network topology information and will be advertised by potential calculation process. The boundary nodes of CN are also named as egress nodes. They advertise network layer reachable information to neighbor nodes periodically and initiate node potential calculation procedure for itself.

This advertisement method separates the information of CN topology from the network layer reachable information and advertises them in different manner. The advertisement of topology information is independent with network layer reachable information. From this way, the flexibility of routing information distribution can be enhanced.

3.1.2 Network layer reachable information advertisement

Network layer reachable information that means the user network prefix is advertised through NLRI-FAP, NLRI-SRP and NLRI-DAP packets. The boundary nodes in CN are responsible for advertising user network reachable information, as shown in the figure 2.



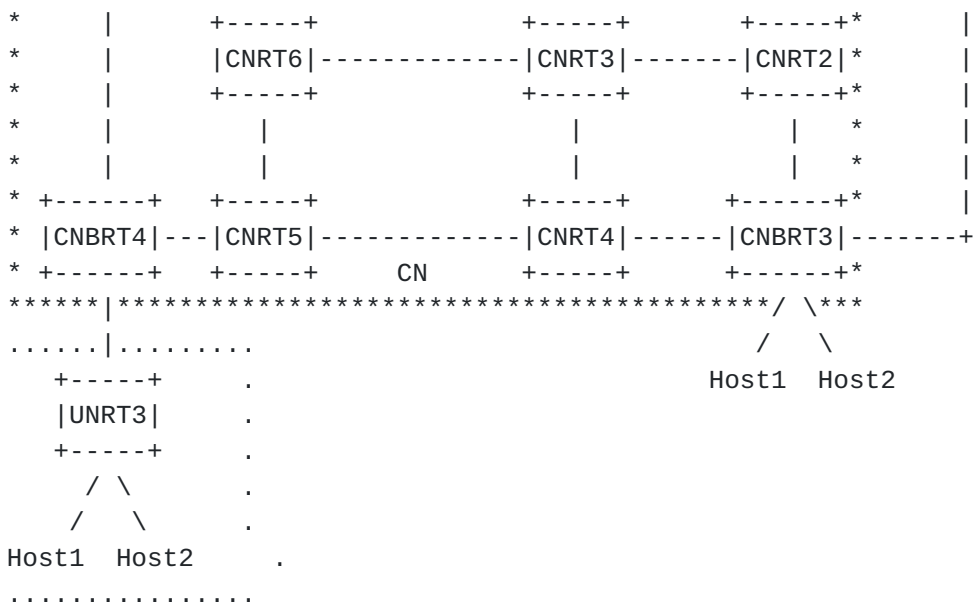


Figure 2: Network layer reachable information advertisements

CNRT: Carrier Network Router.

CNBRT: Carrier Network Boundary Router.

The boundary nodes in CN advertise network layer reachable information to neighbor nodes periodically through NLRI-FAP. Once receiving the NLRI-FAP, the nodes in CN extract network prefix information from packets and add them to the network layer reachable information database of their own.

When the nodes in CN don't know user network prefix, these nodes will act as inquirers and send NLRI-SRP to query other nodes and set a query timer. When a node receives the NLRI-SRP, and the network prefix information or attached router information is in its database, it responses to inquirer with a NLRI-DAP packet. Otherwise, it sends the NLRI-SRP packet to its neighbors except for the enquirers. This process will continue until the inquirer get an response from other nodes or such query timer is timed out.

When bound to the egress node (relative information is absent in network layer reachable information database of the node), the node send NLRI-SRP flooding to inquirer, called enquirers.

3.2. Neighbor Node Discovery and Adjacency establishment

When a new node joins in the network, other nodes discover it and initialize their local routing tables as the following steps.

(1) Broadcast a Detect packet and a NLG-CTP packet. This node sets

the sequence number of Detect with 0.The Detect will be sent periodically by defaulted period set in the protocol;

(2) When other nodes receive a Detect packet or a NLG-CTP packet, they operate as follows.

(a) If the received Detect message is ca from an unknown node and sequence number is equal to 0, the receivers know that a new node add to network and add it to their neighbor node database and record sequence number of the Detect message;

(b) Reply to the unknown node with an REPLY packet; Send a NLG-IAP packet which contains all egress nodes information in the database;

(c) According to the computation rules of node potential, they compute layer value and potential value relative to the egress node.

(3) When received a REPLY packet and a NLG-IAP packet, the new node operates as follows.

(a) Adds the neighbor node to routing information database;

(b) According to NLG-IAP, it calculates its own layer value as following formula.

L(i,i)=0,

L(i,j)=min[L(k,j)]+1,

for each k blongs to Ki, j blongs to N, and k,j are not equal to i.

L(i,j) is the layer value of node i reference to egress node j, N is the set of network nodes, Ki is the set of neighbor nodes about node i.

(c) Based on the layer values of its neighbors, bandwidth and its own layer value, the new node computes node potential value of its neighbors and its own.

3.3. Calculation of node potential value

3.3.1 Definition of node potential value

NP-MNRP defines node potential value as a mixed metric of hop count and bandwidth. Hop count metric records the number of routers which a packet passes through and each router is recorded as one hop. The actual link bandwidth is allocated\configured by router in the network initialization.

3.3.2 Protocol packets sending and processing

3.3.2.1 Network Layer Graph Routing Request Packet processing

Network Layer Graph Routing Request Packet (NLG-RRP) has two different types. One type is all nodes layer information request packet which will request the receiver to send layer information of all nodes. The other is partial nodes layer information request

packet which will request the receiver to send layer information of the specified nodes.

The NLG-RRP will be used in the following conditions.

Condition 1: When a new node is attached to the network, all nodes layer information request packet will be sent to its neighbors.

Condition 2: When the node want to get some certain network prefix information from its neighbors, partial nodes layer information request packet will be sent.

Once receiving a NLG-RRP packet, each node should process as the following steps.

(a) Firstly, the type of the packet will be verified whether or not it is a correct NLG-RRP packet.

(b) When NLG-RRP is the all nodes layer information request packets, it sends its own layer information relative to the other nodes to the request node.

(c) When NLG-RRP is the partial nodes layer information request packet, it sends its own layer information relative to the destination node in the packet to the query node.

3.3.2.2 Network Layer Graph Information Advertisement Packet sending

Network Layer Graph Information Advertisement Packet (NLG-IAP) is broadcast to its neighbors in the following cases.

Case 1: Each layer value of every node is specified after the network layer graph has been divided.

Case 2: When the layer information of any node is changed, which may be caused by the change of layer values of some routers when they received NLG-FAP or some new nodes are added to the network.

Case 3: A NLG-RRP packet is received and the layer value of receiver reference to certain egress node is changed.

The processing rules of NLG-IAP are described in section 3.5.

3.3.3 Calculation of the node potential value

In NP-MNRP, the node potential value is a mixed metric, including hop count and actual link bandwidth. Each egress node activates its neighbors to build Network Layer graph by sending a NLG-CTP packet to

neighbors. Calculated as follows.

(1) In the algorithm initialization, egress node assigns its layer value as 0, and then sends a NLG-CTP packet to its neighbors, which contains the actual link bandwidth.

(2) The neighbor node that received a NLG-CTP packet from the 0-layer node sets itself as 1-layer, and generates a NLG-CTP packet and sends to its neighbors. And so on.

(3) Repeat the above step (2) until the nodes whose layer value are smaller than f-layer which have been determined.

(4) After the f-layer has been constructed completely, each node in f-layer knows that it is in f-layer and knows that which neighbors are in f-1 layer. The nodes in f-1layer know that which ones are in f layer, which ones are in f-1 layer, and which ones are in f-2 layer.

(5) Supposed f-layer has been constructed, and the algorithm has not been terminated, then the construction of the f +1 layer will begins. The nodes in f-layer will send a NLG-CTP packet to its neighbors.

(6) When all the nodes are traveled, algorithm ends.

All nodes in the network not only are aware of their own layer value but also know the layer values of their neighbors and the actual bandwidth between themselves.

Next, each node calculates the potential value as follows.

(1) 0-layer nodes directly define their own potential value as 0;

(2) Each node of the other layers chooses the largest potential value node among its neighbors in the same and lower layers as its reference node for defining potential value, then defines its potential value as potential value of the reference node plus one and potential values of its neighbors which are equal to their layer value;

(3) When each potential value of the neighbors is defined, the node will choose the neighbor nodes whose potential values are less than its own as feasible next-hops of itself.

3.4. network event sensing

NP-MNRP protocol can sense events such as new node add to network or neighbor node is down. The node senses these events through sending out Detect packet and receiving REPLY packet periodically.

3.4.1 Message sending and processing

Each NP-MNRP node broadcasts periodically Detect packet. When a Detect packet is received, the receiver sends a REPLY packet to the sender. The detailed steps are as follows.

Step1: Each node broadcasts periodically a Detect packet. Detect packet carries a sequence number and periodic interval.

Step2: When a Detect packet is received, its sequence number is compared with the expected sequence number for this neighbor. Then the receiver sends a REPLY packet to the sender and evaluates inverse link quality according to the sequence number.

Step3: When a Reply packet is received, its sequence number should be equaled to the sequence number of the Detect packet. Then the receiver evaluates link quality according to the sequence number and the use of sent/receive message time.

3.4.2 Forward direction link sensing

A(Ti) represents the mean interval from node i sending out Detect packet to receiving REPLY packet. FA (Ti) represents the latest mean interval time. Node i adjust the period of sending out Detect packet according to the following rules.

(1) If FA(Ti) < A(Ti), it means that the link quality is stable. At this period, when three continuous Detect packets are all low, then increase the Detect interval;

(2) If FA(Ti) > A(Ti), it means that the link quality is unsteady or become bad. At this period, when three continuous Detect packets are all high, then reduce Detect interval;

(3) When a Detect packet is sent out but a REPLY packet is not received in the 2*FA(Ti) time, the node will reduce the period at half, then continue sending out the next sequence number Detect packet. If a REPLY packet still not received, the node insulates that node and thinks link breakdown or link congestion.

(4) The isolate node will not be used as the next-hop to forwarding packets. The processing node should continue sending Detect packet and adjust strategy according to the following conditions.

(a) If three continuous REPLY packets received, then the node deletes the isolate node and uses nodes which reply the REPLY packet as the next-hops to forwarding packets.

(b) If no packet is received after 6 Detect intervals, so that

node is thought to be breakdown, unable to send Detect packet again. Waiting

(c) If the node can receive a Detect packet or a REPLY packet which is send out from isolate node, and then it will continue sending a Detect packet until the isolate node recovery or judge node breakdown.

3.4.3 Inverse link sensing

Each node broadcasts periodically Detect packet. Every Detect carries a sequence number and the interval.

When a Detect packet is received, compared sequence number with the next expected sequence number for this neighbor, if the sequence number of the received Detect packet is higher than the expected, then one or more Detect packets have been missed. If the sequence number is lower, then this neighbor decrease the Detect interval, and part of the history must be undone.

From the history of received Detect packets, a node computes an estimate of the inverse link quality.

3.5. Update of the node potential

3.5.1. Update of network layer value

As the set of next-hop is empty, the node will trigger update about the layer value through the following operations.

(1) Firstly, the node will check whether the set of its neighbors on the same layer is empty. If is not, it will change itself layer value and generate NLG-IAP packets that are sent to its neighbors.

(2) If the set is empty, then the node will perform the following steps.

(a) Flooding NLG-RUP packets, whose layer value will be revised to the old layer value minus 2, however, if old layer value minus 2 is less than 0, the value is set to 0. What's more, all information about this egress node will be deleted.

(b) The node received a NLG-RUP packet will check if it has ever received this message. If having received such message, the packet will be ignored, otherwise view the layer value in this packet. If the value is greater than the value of its own, the node will discard the packet, otherwise, go on flooding NLG-RUP packets to its neighbors and delete all information about this egress node.

(c) If the layer value in the packet is equal to its own, the node will stop flooding to send NLG-RUP and send a NLG-CTP packet about the egress node.

(d) Neighbor node received a NLG-CTP packet will calculate its new layer value. If new value is same with the original, the node stops sending NLG-CTP packets and the layer value update algorithm will terminate.

3.5.2. Update of node potential value The potential value of the nodes will be updated in the following conditions.

(1) Potential value update caused by NLG-IAP packet. The main content of a NLG-IAP packet are layer information of a node relative to the egress node in the network. The node adjusts its potential value according to the NLG-IAP packets received from its neighbors. Specific updated as follows.

(a) NLG-IAP packet generation When the node finds its layer information relative to one certain egress node changed, the new layer information relative to all egress nodes will be written into its generated NLG-IAP packet.

(b) NLG-IAP to send The node will broadcast its generated NLG-IAP packets to its neighbors, and then wait for confirmation from its neighbors. When the node receives REPLY pockets with sequence number 0XFFFF from its neighbors, this update ends.

(c) When the node received the NLG-IAP packet, it performs the following steps.

Firstly, this node will verify NLG-IAP packet. If the packet is validated, it sends a REPLY packet with sequence number 0XFFFF to source node of NLG-IAP packet.

Secondly, this node will extract layer information from the NLG-IAP packet supposing a received routing entries includes RUIi =< RouterID, EgressnodeID, Layi> and the corresponding routing entry stored in the database is RUI'i =< RouterID, EgressnodeID, Lay'i>, when one of the following three cases are met, the node update its own routing table according to the information received.

Case 1: If the EgressnodeID included in the received routing table entry dose not exist in its own routing table, the router will add it and send a NLG-RRP packet about the destination address of the network to its neighbors. After received NLG-IAP packets, it will calculate its layer value relative to this

egress node according to formula in 3.2, and get layer values of it's neighbors, then define the potential value of neighbor node reference to this destination node.

Case 2: If the EgressnodeID included in the received routing table entry exists in its own routing table and Lay'i < Layi, the router will adjust the layer values of its neighbors and potential value. If the potential value of neighbor is greater than its own and the neighbor node is just in its own next-hop collection, the router will remove the node from next-hop collection, if not, do nothing.

Case 3: If Layi < Lay, then the router will adjust the layer value of neighbor node and potential value. If the potential value of neighbor is less than its own and the neighbor node is not in its own next-hop collection, then the router will add it into the next-hop

(2) Potential value update caused by the change of layer value

When the layer value reference to a particular destination node changes, firstly it will modify layer value of itself and its neighbors. Then the new potential value will be calculated according to update strategy.

4. Routing information base

Each routing node maintains the Routing Information Base to forward IP packets. NP-MNRP protocol can calculate multiple next-hops routing information in the network. The route computation process descried in Section 3 will record all the feasible next-hops for every destination. The multiple next-hops routing entry is composed of three fields.

The first field is named as destination IP address field.

The second field is named as the destination IP address mask filed. The first and second fields are 32-bit numbers which define network destination information of each routing entry. And the route information base is indexed by the first and second field.

The third field is named as feasible next-hop filed, whose form is a couple of next-hop IP address value and node potential difference value. The next-hop IP address is IP address of the neighbor node who will act as a feasible next-hop for this routing node to forward IP packets to the network destination. Compared with the single next-hop routing information, the number of feasible next-hop filed in a multiple next-hops routing entry may be a value bigger than one.

The routing node will gain more agility in IP packet forwarding procedure. It can choose forward all the packets to one best next-hop and use other feasible next-hop entries as backup entries. Through this way, the network availability will be improved in the scene that network failure or mal-function occurs frequently. It can also choose forward all the IP packets to all feasible next-hop entries. In this manner, this routing node can assign each next-hop a traffic ratio for each destination. All the packets will be forwarded to variable next-hop pro rata. This will make the network resource utilization more evenness and avoid network congestion in some ways.

5. NP-MNRP packets format

The NP-MNRP protocol runs directly over the IP network layer. Before any packet format is described, the details of the NP-MNRP encapsulation are explained.

5.1. Encapsulation of NP-MNRP packets

NP-MNRP runs directly over the Internet Protocol's network layer. Therefore, NP-MNRP packets are encapsulated solely by IP and local data-link headers. NP-MNRP does not define a way to fragment its protocol packets, and depends on IP fragmentation when transmitting packets larger than the network MTU.

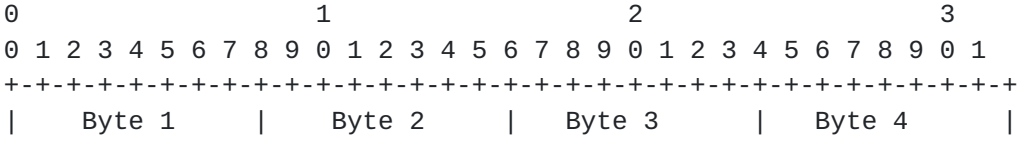
NP-MNRP uses IP protocol number 99. Routing protocol packets are sent with IP precedence set to inter-network Control. NP-MNRP protocol packets should be given precedence over regular IP data traffic, in both sending and receiving.

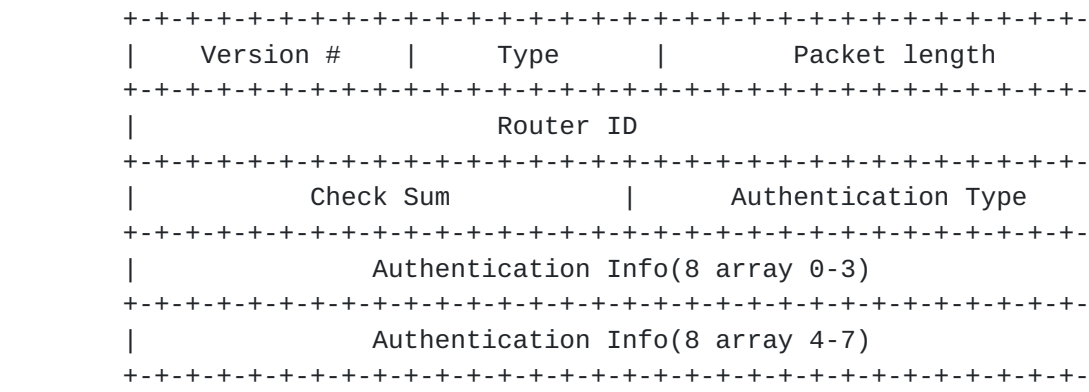
5.2. Packet format

There are nine distinct NP-MNRP packet types. All NP-MNRP packet types begin with a standard 20 byte header. This header is described first. Each packet type is then described in a succeeding section. In these sections each packet's division into fields is displayed, and then the field definitions are enumerated.

5.2.1 The NP-MNRP packet header

Every NP-MNRP packet starts with a standard 20 byte header. This header contains all the information necessary to determine whether the packet should be accepted for further processing. This determination is described in Section 5.2 of the specification.





Version
The NP-MNRP version number. This specification documents version 1 of the protocol.

Type

Type	Description
1	NLRI-FAP
2	NLRI-SRP
3	NLRI-DAP
4	NLG-CTP
5	NLG-IAP
6	NLG-RRP
7	NLG-RUP
8	Detect
9	REPLY

Router ID
The Router ID of the packet's source.

Check Sum
The standard IP checksum of the entire content of the packet. Note that the packet starts with the NP-MNRP header but excluding the 64-bit authentication field. If the length of the package is less than 16-bit, 0 byte WOULD be added before the checksum byte.

Authentication Type Identify the authentication procedure used for the packet. Authentication is discussed in the specification.

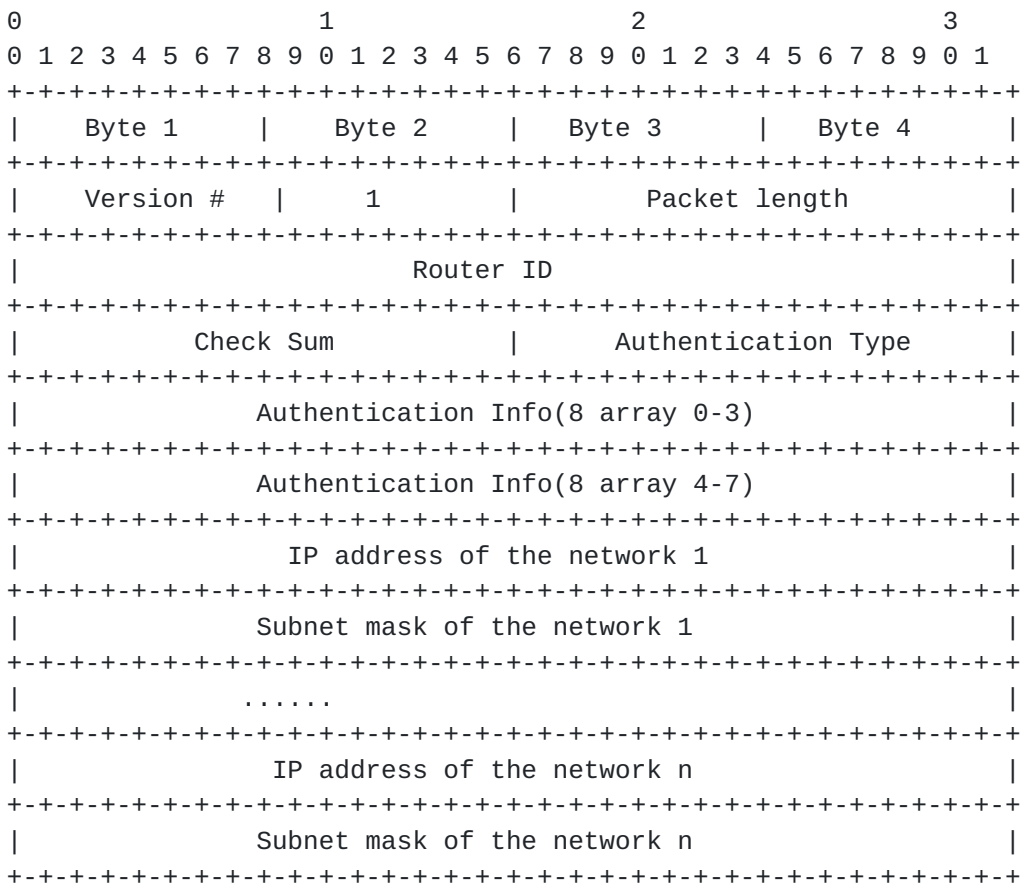
Authentication Info 64-bit authentication information field depends on the chosen authentication type, to carry identification information such as identity authentication.

5.2.2 Various types of protocol packets

NP-MNRP protocol packets have total of 9 species.

5.2.2.1 Network Layer Reachable Information Flood Advertisement Packet, NLRI-FAP

NLRI-FAP packet is NP-MNRP packet type 1.The packet is send by an egress node in the carrier network and is used to advertise the user network prefix information which is bound to this node. The node that received this packet WOULD consider this egress node is reachable and the user network which is bound to this node is reachable too.



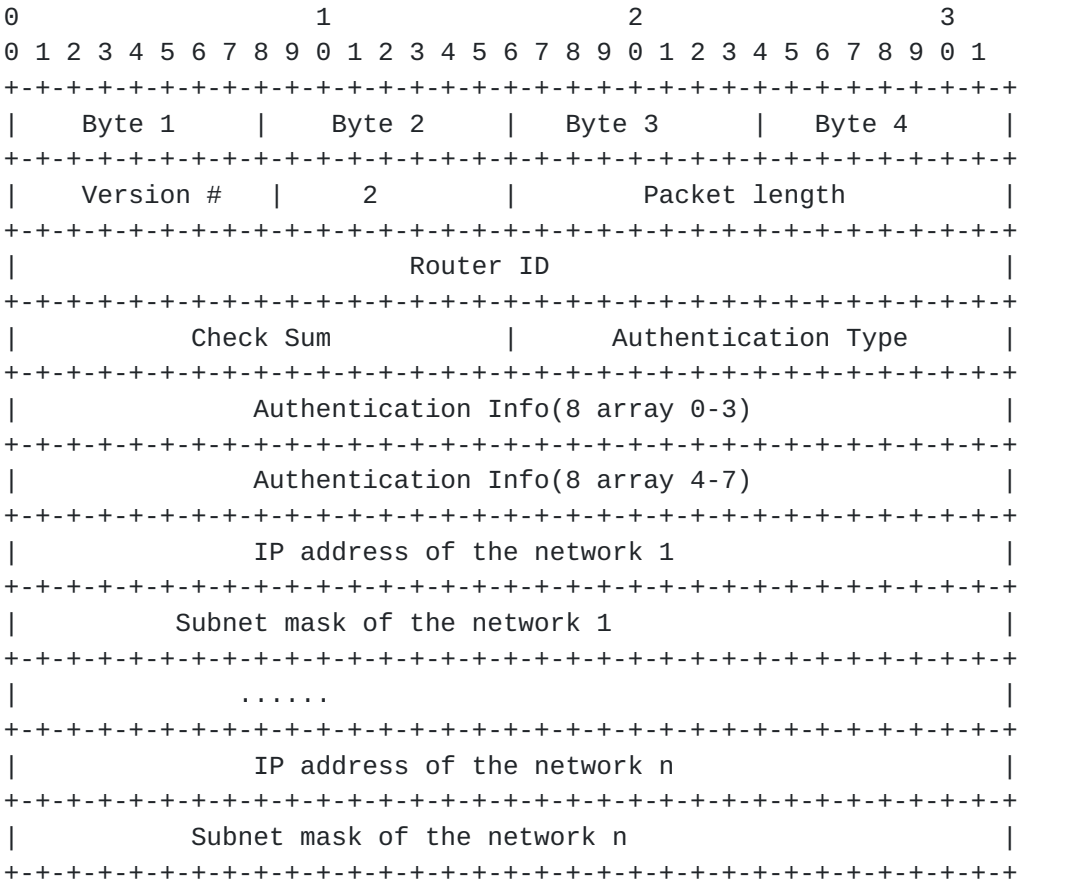
IP address of the network
The IP address of a user network is bound to this egress node.

Subnet mask of the network
Subnet mask of user network is denoted in the above field.

5.2.2.2 Network Layer Reachable Information Specific Request Packet, NLRI-SRP

NLRI-SRP packet is NP-MNRP packet type 2.This packet is used by the node in the carrier network to request the network layer reachable information. When the node doses not know which egress node, one network prefix is bound to, the NLRI-SRP packet is sent.

If the IP address and the subnet mask of the network 1 is all 0, all the network prefixes WOULD be requested, and usually this request is sent by a node which just joins into the carrier network.



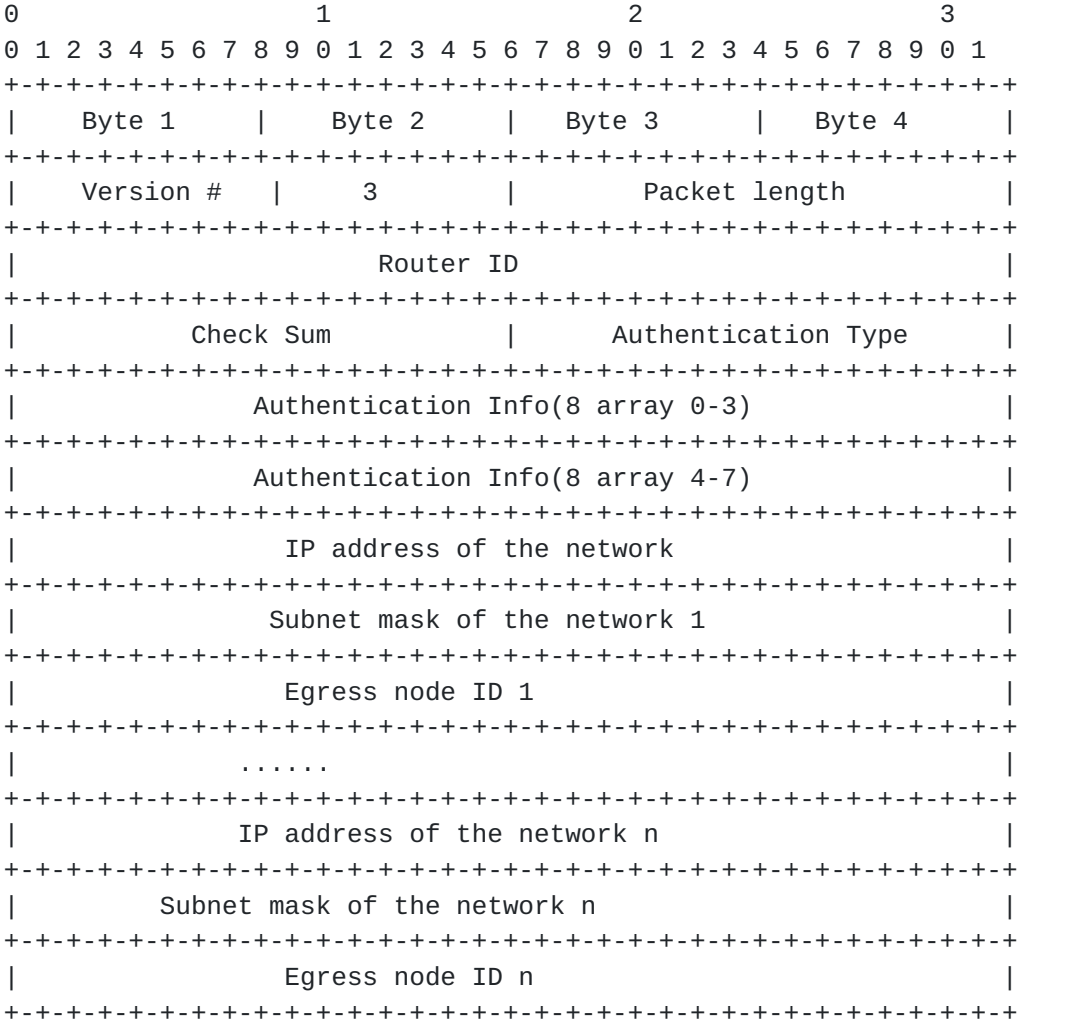
IP address of the network
The IP address of a user network whose binding relationship this node want to request.

Subnet mask of the network
The Subnet mask of a user network denoted in the above field, who binding relationship this node wants to request

5.2.2.3 Network Layer Reachable Information Direction Answer Packet, NLRI-DAP

NLRI-DAP packet is NP-MNRP packet type 3. The node that receives the NLRI-SRP packet WOULD check its information database to seek for the corresponding binding relationship. If the binding relationship exists in its own information database, it WOULD send NLRI-DAP packet to the requesting node, if not, it WOULD send NLRI-SRP packets to its all neighbor nodes except for the requesting node.

If a node who sends the NLRI-SRP packets at the same time receives more than one packet about the same IP prefix binding relationship, and these packets are conflictive, it WOULD go on sending NLRI-SRP



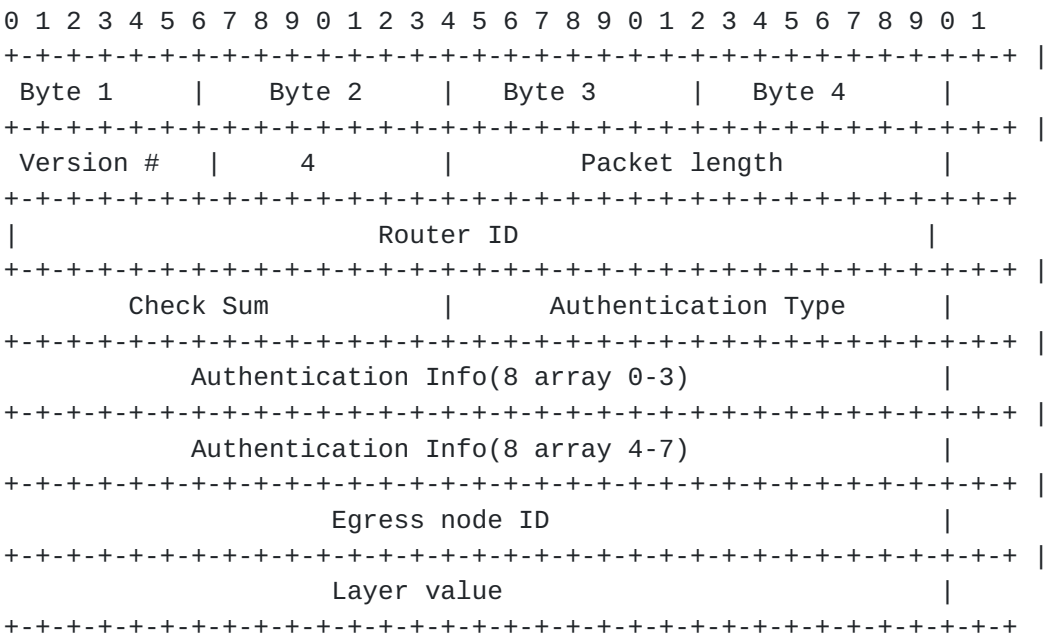
IP address of the network
The IP address of a user network whose binding relationship is wants to know in the receiving NLRI-SRP packet.

Subnet mask of the network
The subnet mask of a user network denoted in the above field, whose binding relationship is wants to know in the receiving NLRI-SRP packet.

Egress node ID
This field denotes the user network, whose IP address and subnet mask emerge in the front two fields, is bound to this egress node.

5.2.2.4 Network Layer Graph Construction Trigger Packet, NLG-CTP

NLG-CTP packet is NP-MNRP packet type 4.When one egress node joins into the carrier network, it WOULD send this packet to urge the other nodes to build network layer graph relative to itself.

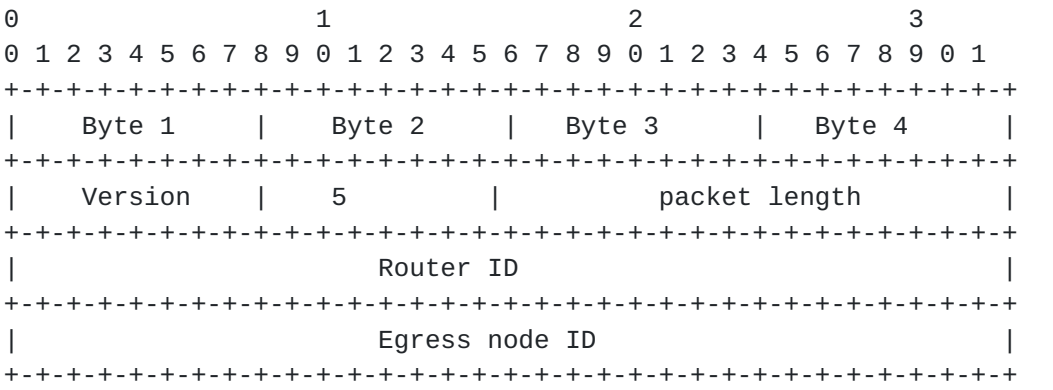


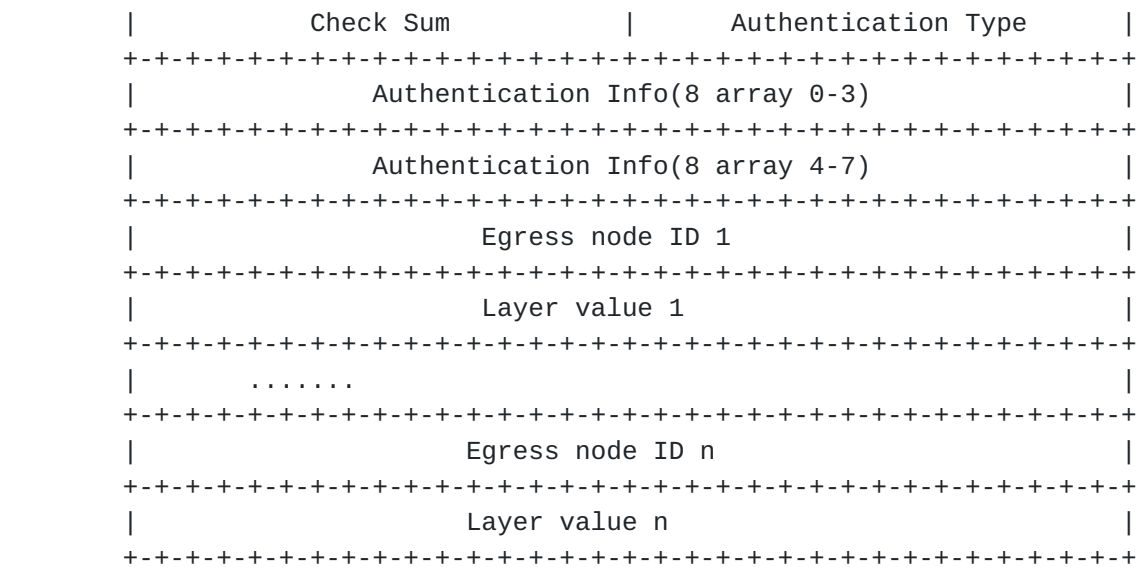
Egress node ID The node that received this packet calculates its layer value relative to this egress node.

Layer value The value of the sender relative to the node denoted in the above field

5.2.2.5 Network Layer Graph Information Advertisement Packet, NLG-IAP

NLG-IAP packet is NP-MNRP packet type 5. This packet is sent to all neighbor nodes using multicast address to advertise its layer value relative to the node whose ID includes in this packet, this packet is usually to start the neighbor nodes to adjust their potential dynamically.



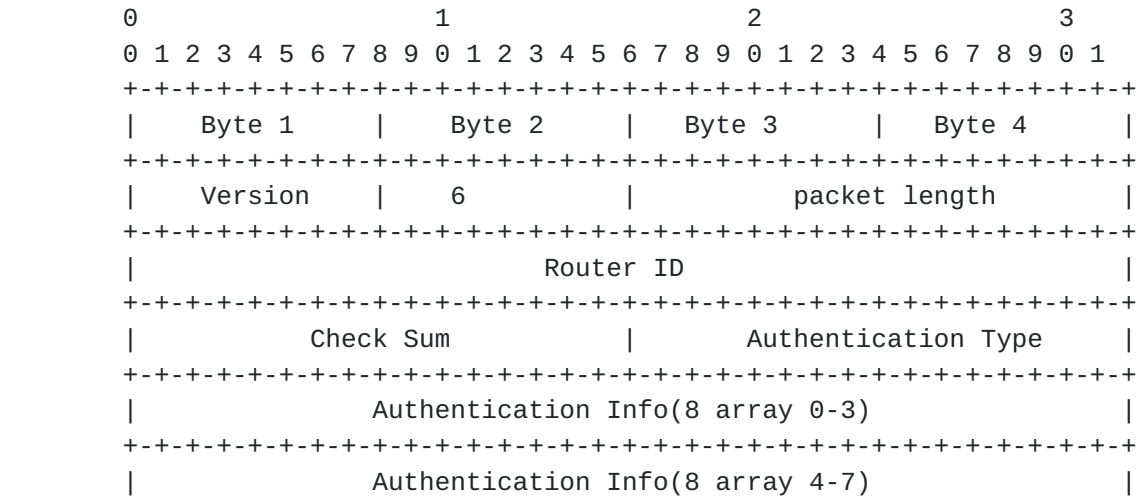


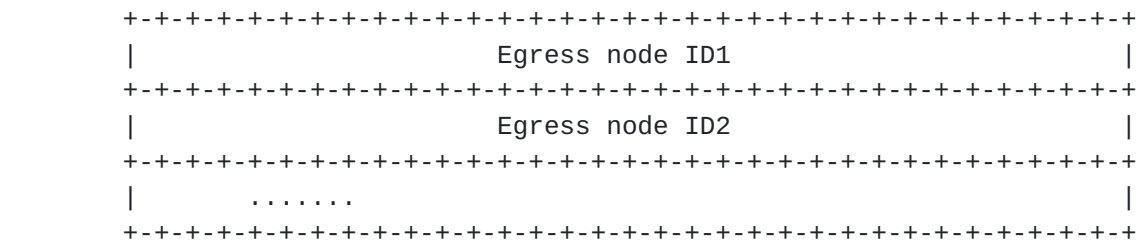
Egress node ID
The receiver need adjust its layer value relative to this node

Layer value
This is the layer value of the sender relative to the egress node denoted in the above field.

5.2.2.6 Network Layer Graph Routing Request Packet, NLG-RRP

NLG-RRP packet is NP-MNRP packet type 6. This packet is used by one node to request the neighbor nodes to send their own layer information relative to the appointed egress node to it. This packet has two types: one is for all nodes and the other is for some nodes. If the egress node ID in the packet is empty, the packet is NLG-RRP packet for all nodes, if not; the packet is NLG-RRP packet for the appointed nodes included in this packet.

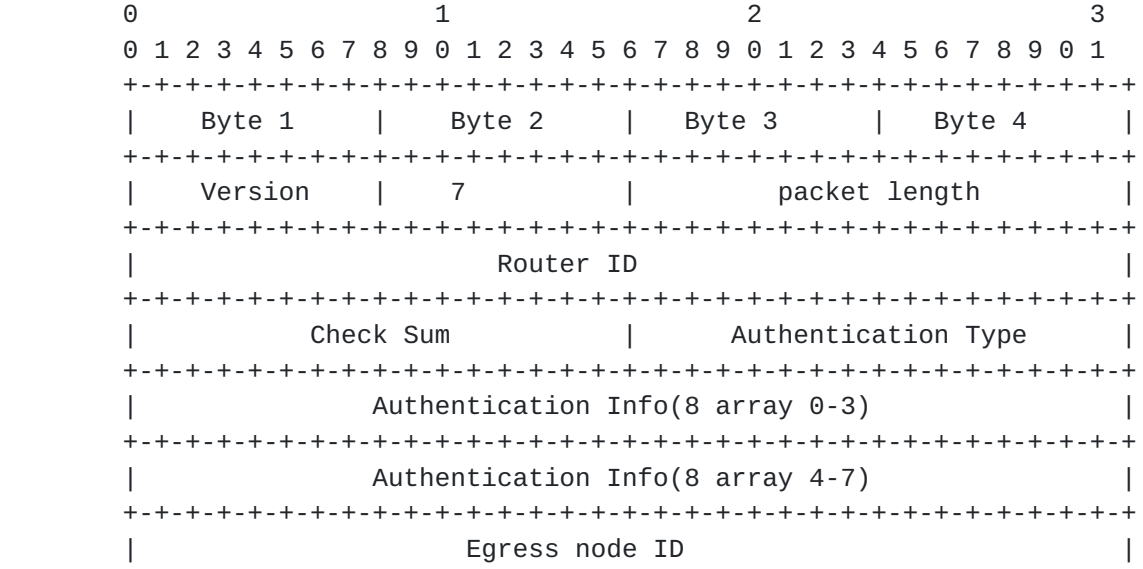




Egress node ID
This field denotes the receiver need to respond its layer value relative to this egress node.

5.2.2.7 Network Layer Graph Request Update Packet, NLG-RUP

NLG-RUP packet is NP-MNRP packet type 7. When a node finds its next hops and its neighbors on the same layer are all unreachable, it sends this packet to neighbors to update its layer value.

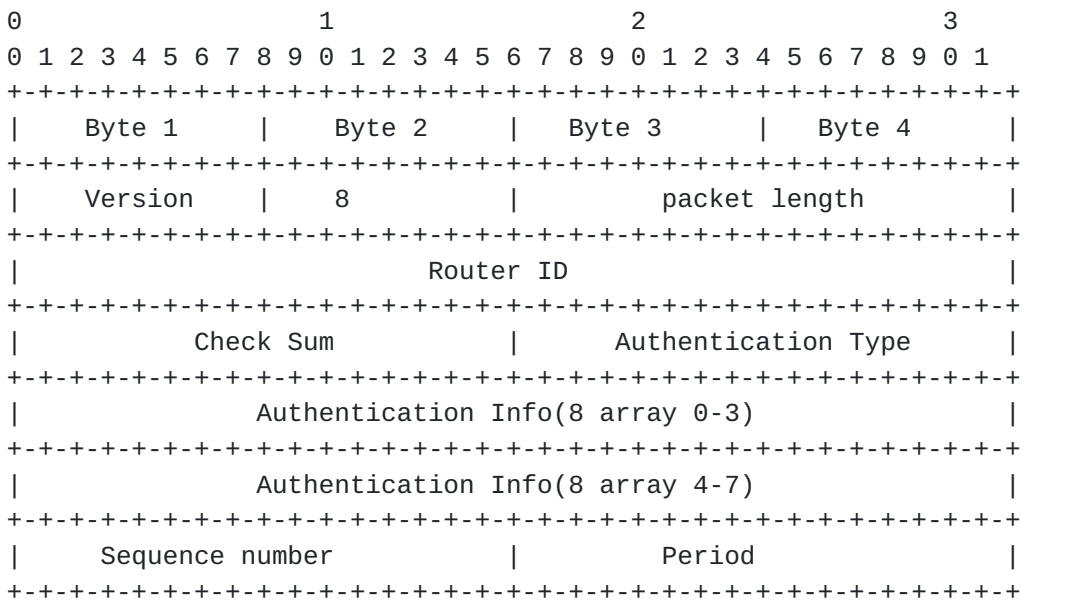


Egress node ID
This field denotes the receiver needs to update its value relative to this node

Layer value
The layer value in NLG-RUP packet is the layer value relative to high egress node minus 2.

5.2.2.8 Link State Detect Packet, Detect

Detect packet is NP-MNRP packet type 8. This packet is periodically sent to neighbor nodes to evaluate the link quality.

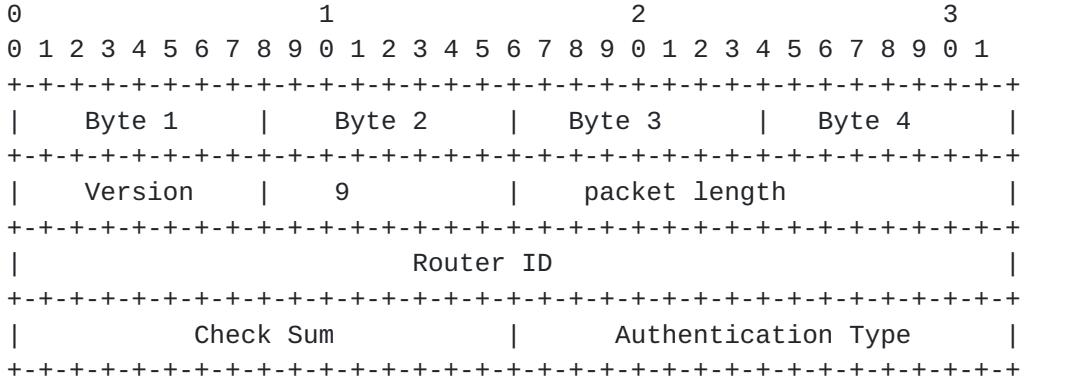


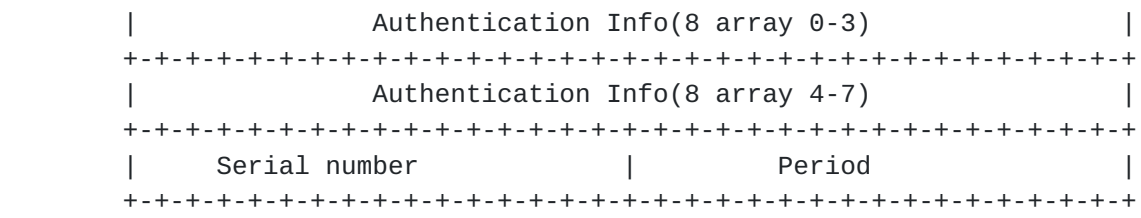
Sequence number
The sequence number in this packet is the sequence number of the Detect message, and it is an increasing positive integer.

Period
The period in this packet is the interval time between this to the last one.

5.2.2.9 REPLY Message Packet, REPLY

REPLY packet is NP-MNRP packet type 9. The node that received a detect packet will send a reply packet to the sender, which can evaluate the reverse link quality according to this packet. The sequence number in this packet denotes it is the reply to the Detect message that has the corresponding sequence number. If the number is full-1, it means the packet is confirmation message to a NLG-IAP message.





Serial number
The sequence number in this packet is the sequence number of the REPLY, and it is an increasing positive integer.

Period
This field denotes that the node sending this packet will sends Detect message at this period.

6. Informative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.

7. Author's Address

Julong Lan
National Digital Switching System Engineering and Technological Research Center
NDSC, No.7 , Jianxue Street,Jinshui District
Zhengzhou, 450002,
P.R.China
Phone: +86-371-8163-2988
Email: ndscljl@163.com
URI: http://www.ndsc.com.cn/

Jianhui Zhang
National Digital Switching System Engineering and Technological Research Center
NDSC, No.7 , Jianxue Street,Jinshui District
Zhengzhou, 450002,
P.R.China
Phone: +86-371-8163-2988
Email: ndsc155@163.com
URI: http://www.ndsc.com.cn/

Bin Wang
National Digital Switching System Engineering and Technological Research Center
NDSC, No.7 , Jianxue Street,Jinshui District
Zhengzhou, 450002,

P.R.China
Phone: +86-371-8163-2909
Email: ndscmt@163.com
URI: http://www.ndsc.com.cn/

Wenfen Liu
National Digital Switching System Engineering and Technological Research Center
NDSC, No.7 , Jianxue Street,Jinshui District
Zhengzhou, 450002,
P.R.China
Phone: +86-371-8163-0340
Email: wenfenliu@sina.com
URI: http://www.ndsc.com.cn/

Youjun Bu
National Digital Switching System Engineering and Technological Research Center
NDSC, No.7 , Jianxue Street,Jinshui District
Zhengzhou, 450002,
P.R.China
Phone: +86-371-8163-2670
Email: buyoujun2009@hotmail.com
URI: http://www.ndsc.com.cn/

Xin Li
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS
No.10 Xi Tu Cheng Road Haidian District
Beijing,100876
P.R.China
Phone: +8613581614576
Email: cplalx@gmail.com
URI: http://www.bupt.edu.cn/