IP Routing Protocols Header Formats

Ali Aydemir
2017

No	INDEX
4	RIPv1
10	RIPv2
18	RIPng
26	OSPFv2
60	OSPFv3
68	BGPv4
93	ISIS
121	EIGRP

Objectives

- Describe common Routing Protocols header design models.
- Describe how to generate a header in different Routing Protocols.
- Review the fundamentals and compare Routing Protocols.

RIPv1

RIPv1 Related RFCs

- **RFC 1058**: Routing Information Protocol

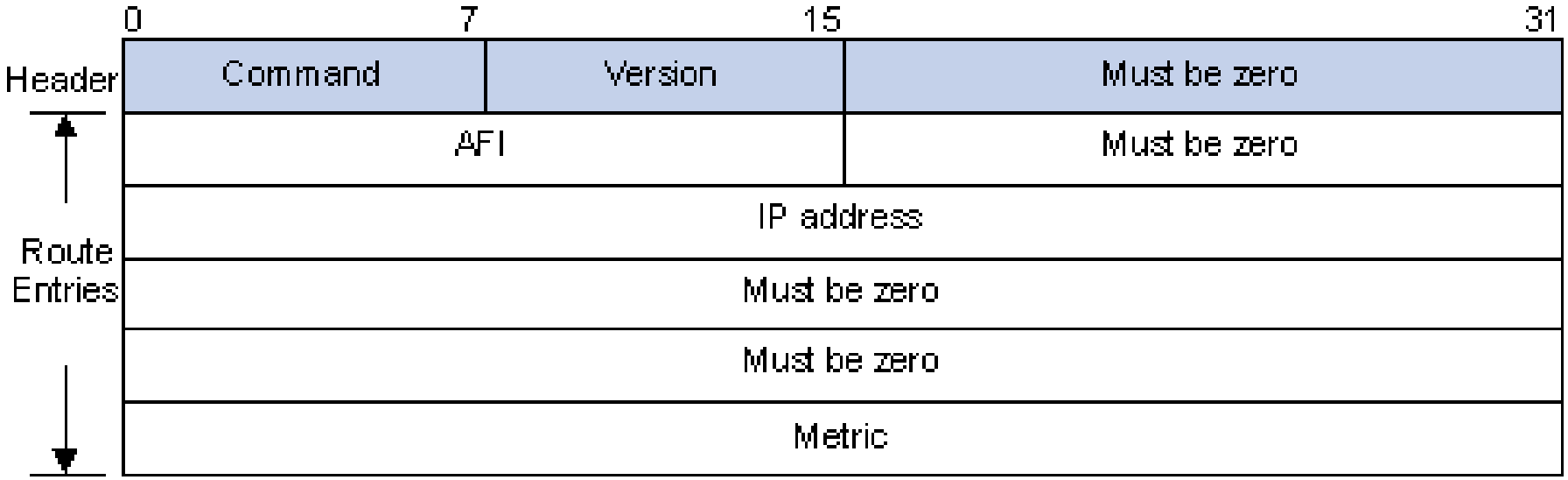
RIPv1 Packet Format

- RIPv1 packets are encapsulated after UDP segments. RIPv1 has the UDP port number **520**.
- The RIPv1 packet format is shown below:

Frame Header	IP Header (Protocol no = 17)	UDP (Port no = 520)	RIP Message	CRC
On a LAN, the RIP packet is encapsulated in an Ethernet frame with a destination broadcast MAC address: • **FF-FF-FF-FF-FF-FF**	The destination broadcast IP address is set to **255.255.255.255** The IP protocol field is **17**.	Used UDP Port Number **520**	The RIP message contains the route information.	

RIPv1 Message Format

- A RIP message consists of the Header and up to 25 route entries.



RIPv1 Message Format

- **Command**: The type of message. 1 indicates Request, 2 indicates Response.
- **Version**: The version of RIP, 0x01 for RIPv1.
- **AFI**: Address Family Identifier, 2 for IP.
- **IP Address**: Destination IP address of the route; can be a natural network, subnet or a host address.
- **Metric**: Cost of the route.

RIPv2

RIPv2 Related RFCs

- **RFC 1721**: RIP Version 2 Protocol Analysis
- **RFC 2082**: RIP-2 MD5 Authentication
- **RFC 2453**: RIP Version 2

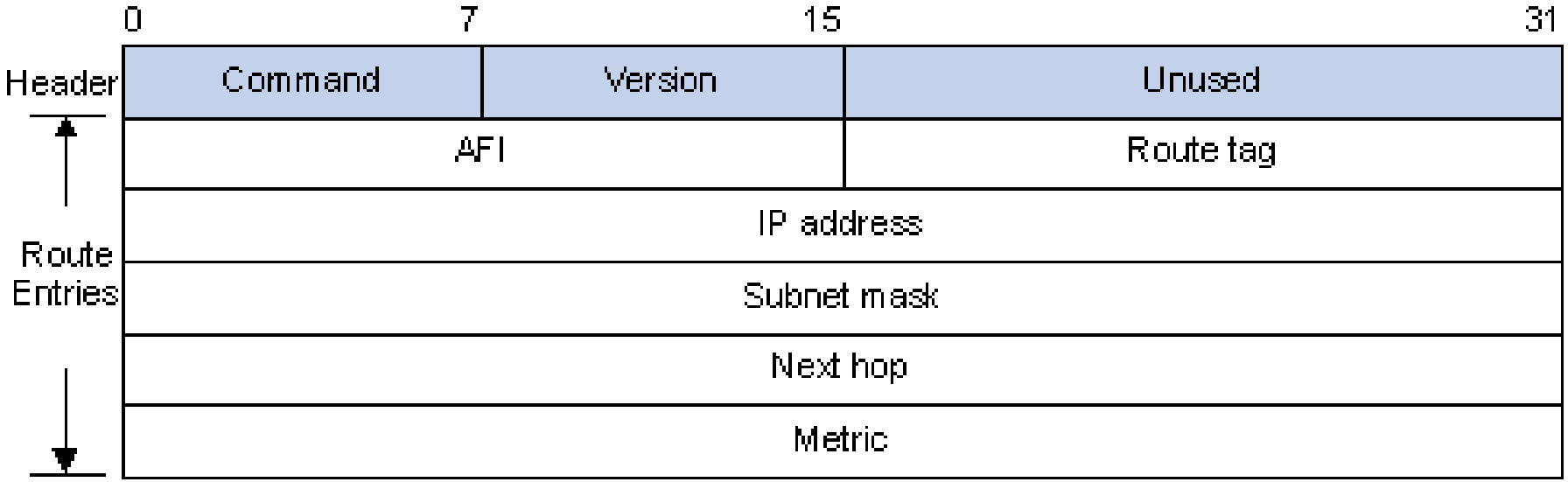
RIPv2 Packet Format

- RIPv2 packets are encapsulated after UDP segments. RIPv2 has the UDP port number **520**.
- The RIPv2 packet format is shown below:

Frame Header	IP Header (Protocol no = 17)	UDP (Port no = 520)	RIP Message	CRC
On a LAN, the RIP packet is encapsulated in an Ethernet frame with a destination multicast MAC address: • **01-00-5E-00-00-09**	The destination Multicast IP address is set to **224.0.0.9** The IP protocol field is **17**.	Used UDP Port Number **520**	The RIP message contains the route information.	

RIPv2 Message Format

- The format of RIPv2 message is similar with RIPv1

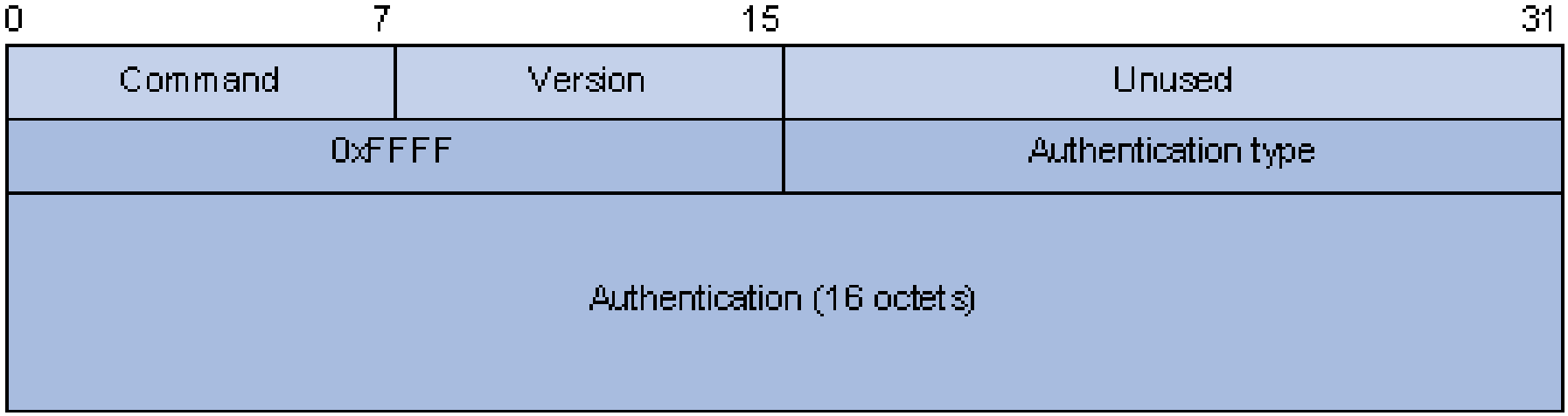


RIPv2 Message Format

- **Version**: Version of RIP. For RIPv2 the value is 0x02.
- **Route Tag**: Route Tag.
- **IP Address**: Destination IP address. It could be a natural network address, subnet address or host address.
- **Subnet Mask**: Mask of the destination address.
- **Next Hop**: If set to 0.0.0.0, it indicates that the originator of the route is the best next hop; Otherwise it indicates a next hop better that the originator of the route.

RIPv2 Authentication

- RIPv2 sets the AFI field of the first route entry to 0xFFFF to identify authentication information



RIPv2 Authentication

- **Authentication Type**: 2 represents plain text authentication, while 3 represents MD5.

- **Authentication**: Authentication data, including password information when plain text authentication is adopted or including key ID, MD5 authentication data length and sequence number when MD5 authentication is adopted.

Note:

- RFC 1723 only defines plain text authentication. For information about MD5 authentication, refer to RFC2082 “RIPv2 MD5 Authentication”.

- With RIPv1, you can configure the authentication mode in interface view. However, the configuration will not take effect because RIPv1 does not support authentication.

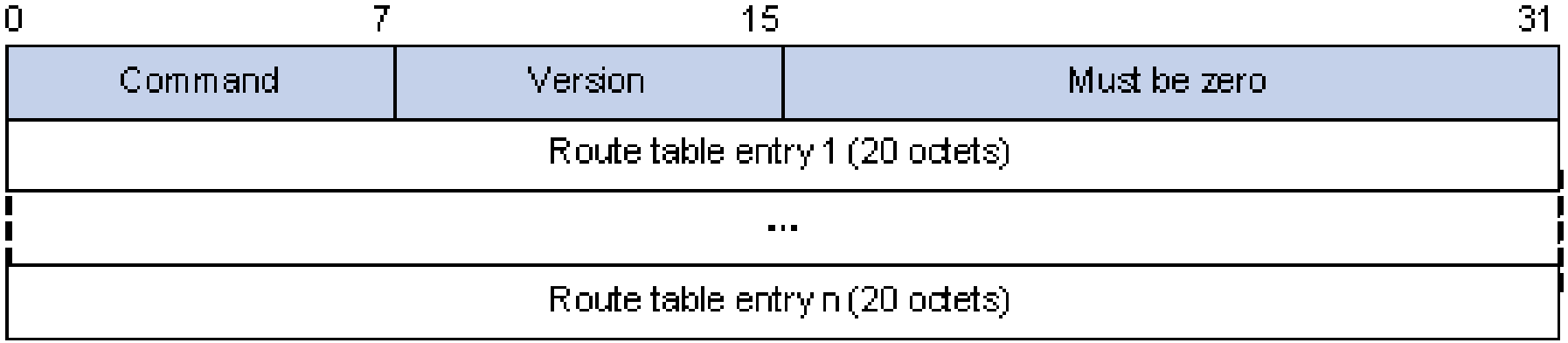
RIPng

RIPng Related RFCs

- **RFC 2080**: RIPng for IPv6
- **RFC 2081**: RIPng Protocol Applicability Statement

RIPng Packet Format

- A RIPng packet consists of a header and multiple route table entries (RTEs). The maximum number of RTEs in a packet depends on the MTU of the sending interface.



RIPng Packet Format

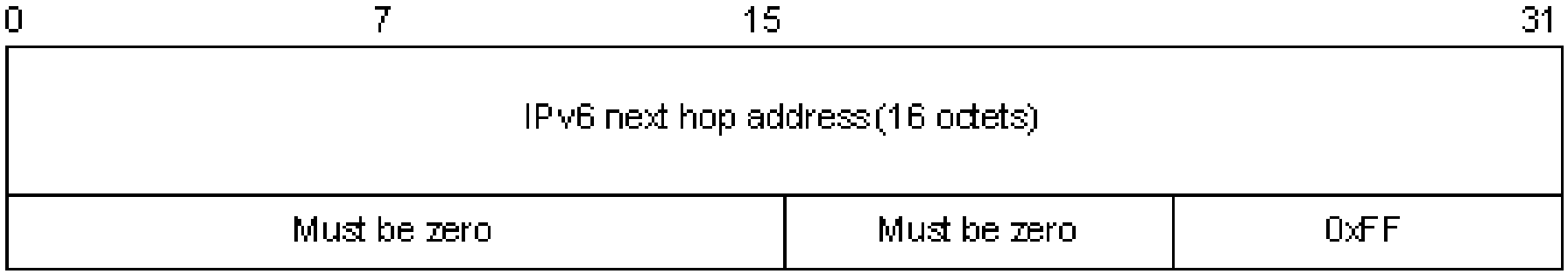
- **Command**: Type of message. 0x01 indicates Request, 0x02 indicates Response.
- **Version**: Version of RIPng. It can only be 0x01 currently.
- **RTE**: Route table entry, 20 bytes for each entry.

RIPng RTE Format

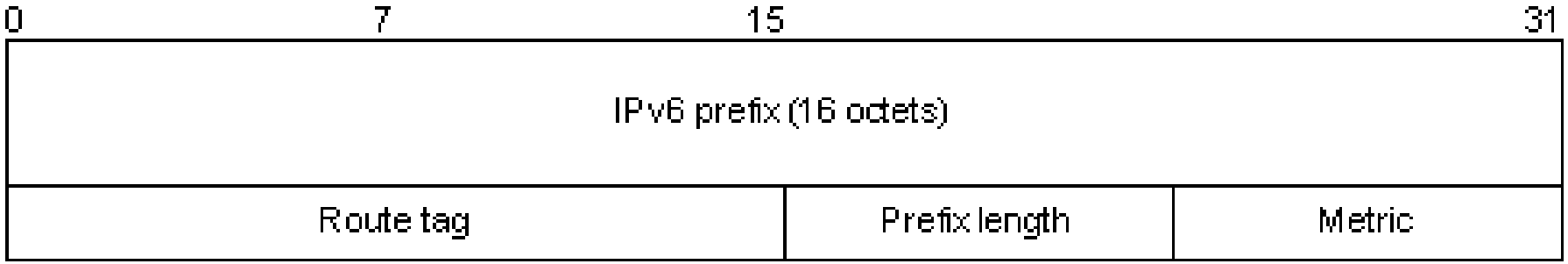
There are two types of RTE in RIPng.

- **Next Hop RTE**: Defines the IPv6 address of a next hop
- **IPv6 Prefix RTE**: Describes the destination IPv6 address, route tag, prefix length and metric in the RIPng routing table.

RIPng Next Hop RTE



RIPng IPv6 Prefix RTE



- **IPv6 prefix**: Destination IPv6 address prefix.
- **Route tag**: Route tag.
- **Prefix len**: Length of the IPv6 address prefix.
- **Metric**: Cost of a route.

OSPFv2

OSPFv2 Related RFCs

- **RFC 1765**: OSPF Database Overflow
- **RFC 2328**: OSPF Version 2
- **RFC 3101**: OSPF Not-So-Stubby Area (NSSA) Option
- **RFC 3137**: OSPF Stub Router Advertisement
- **RFC 3630**: Traffic Engineering Extensions to OSPF Version 2

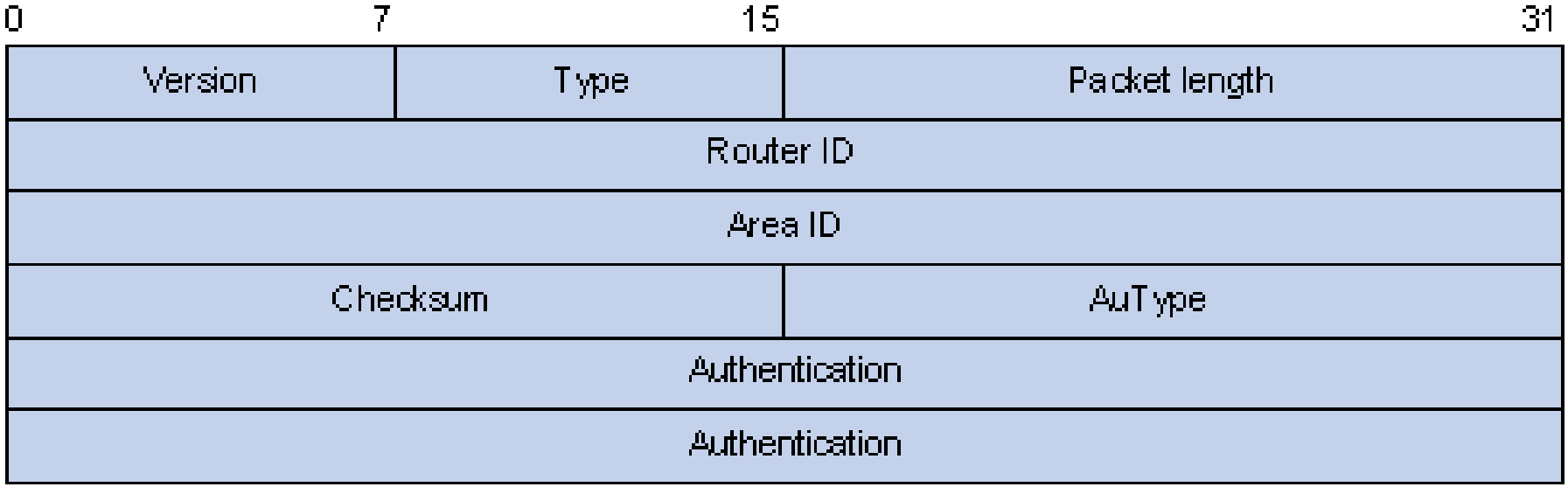
OSPFv2 Packet Format

- OSPF packets are directly encapsulated into IP packets. OSPF has the IP protocol number 89.
- The OSPF packet format is shown below:

Frame Header	IP Header	Protocol Number (OSPF = 89)	OSPF Header	OSPF Message	CRC
On a LAN, the OSPF packet is encapsulated in an Ethernet frame with a destination multicast MAC address of either: • **01-00-5E-00-00-05** • **01-00-5E-00-00-06**	The destination multicast IP address is set to either: • **224.0.0.5** (All OSPF routers listen to this address.) • **224.0.0.6** (All DR and BDR routers listen to this address. The OSPF protocol field is **89**.		The OSPF header identifies the type of OSPF packet, the router ID and the area number.	The OSPF message contains the packet type specific message information.	

OSPFv2 Packet Header

- OSPF packets are classified into five types that have the same packet header, as shown below.



OSPFv2 Packet Header

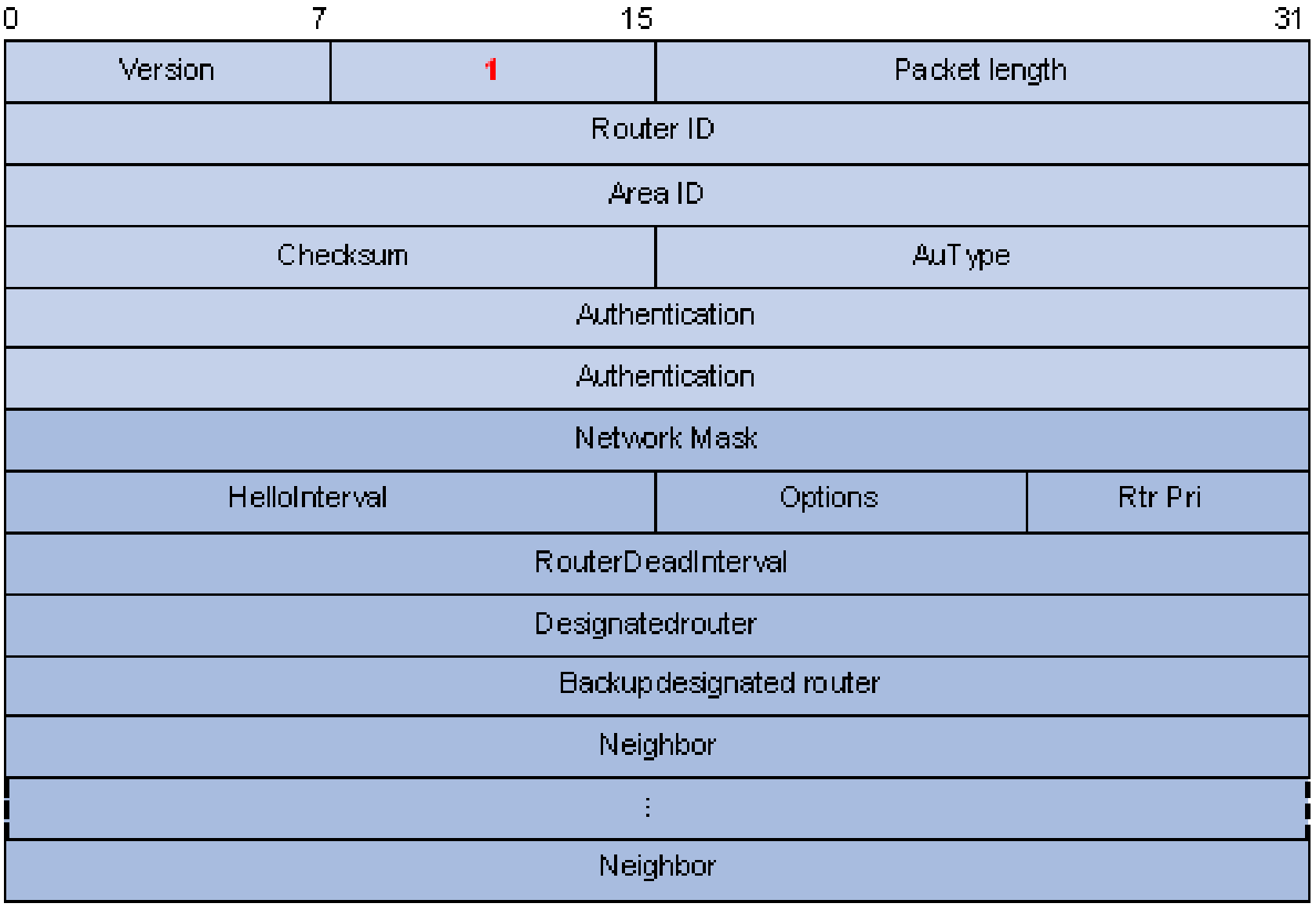
- **Version**: OSPF version number, which is 2 for OSPFv2.
- **Type**: OSPF packet type from 1 to 5, corresponding with hello, DD, LSR, LSU and LSAck respectively.
- **Packet length**: Total length of the OSPF packet in bytes, including the header.
- **Router ID**: ID of the advertising router.
- **Area ID**: ID of the area where the advertising router resides.
- **Checksum**: Checksum of the message.
- **Autype**: Authentication type from 0 to 2, corresponding with non-authentication, simple (plaintext) authentication and MD5 authentication respectively.
- **Authentication**: Information determined by authentication type. It is not defined for authentication type 0. It is defined as password information for authentication type 1, and defined as Key ID, MD5 authentication data length and sequence number for authentication type 2.

OSPFv2 Hello Packet

- A router sends hello packets periodically to neighbors to find and maintain neighbor relationships and to elect the DR/BDR, including information about values of timers, DR, BDR and neighbors already known.

- The format is shown below:

OSPFv2 Hello Packet



OSPFv2 Hello Packet

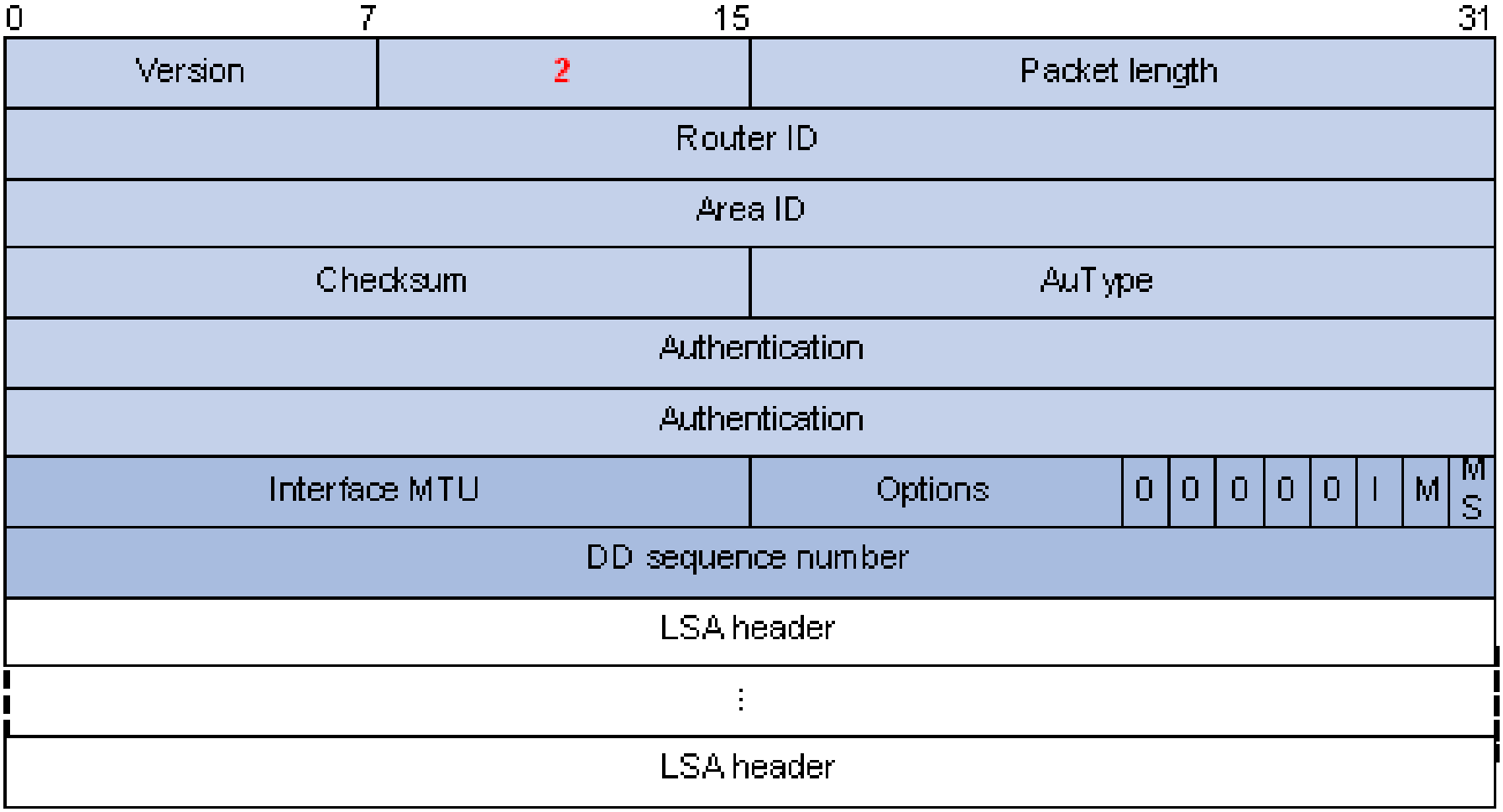
- **Network Mask**: Network mask associated with the router's sending interface. If two routers have different network masks, they cannot become neighbors.
- **HelloInterval**: Interval for sending hello packets. If two routers have different intervals, they cannot become neighbors.
- **Rtr Pri**: Router priority. A value of 0 means the router cannot become the DR/BDR.
- **RouterDeadInterval**: Time before declaring a silent router down. If two routers have different time values, they cannot become neighbors.
- **Designated Router**: IP address of the DR interface.
- **Backup Designated Router**: IP address of the BDR interface
- **Neighbor**: Router ID of the neighbor router.

OSPFv2 Database Description (DD) Packet

- Two routers exchange database description (DD) packets describing their LSDBs for database synchronization, contents in DD packets including the header of each LSA (uniquely representing a LSA). The LSA header occupies small part of an LSA to reduce traffic between routers. The recipient checks whether the LSA is available using the LSA header.

- The DD packet format:

OSPFv2 Database Description (DD) Packet



OSPFv2 Database Description (DD) Packet

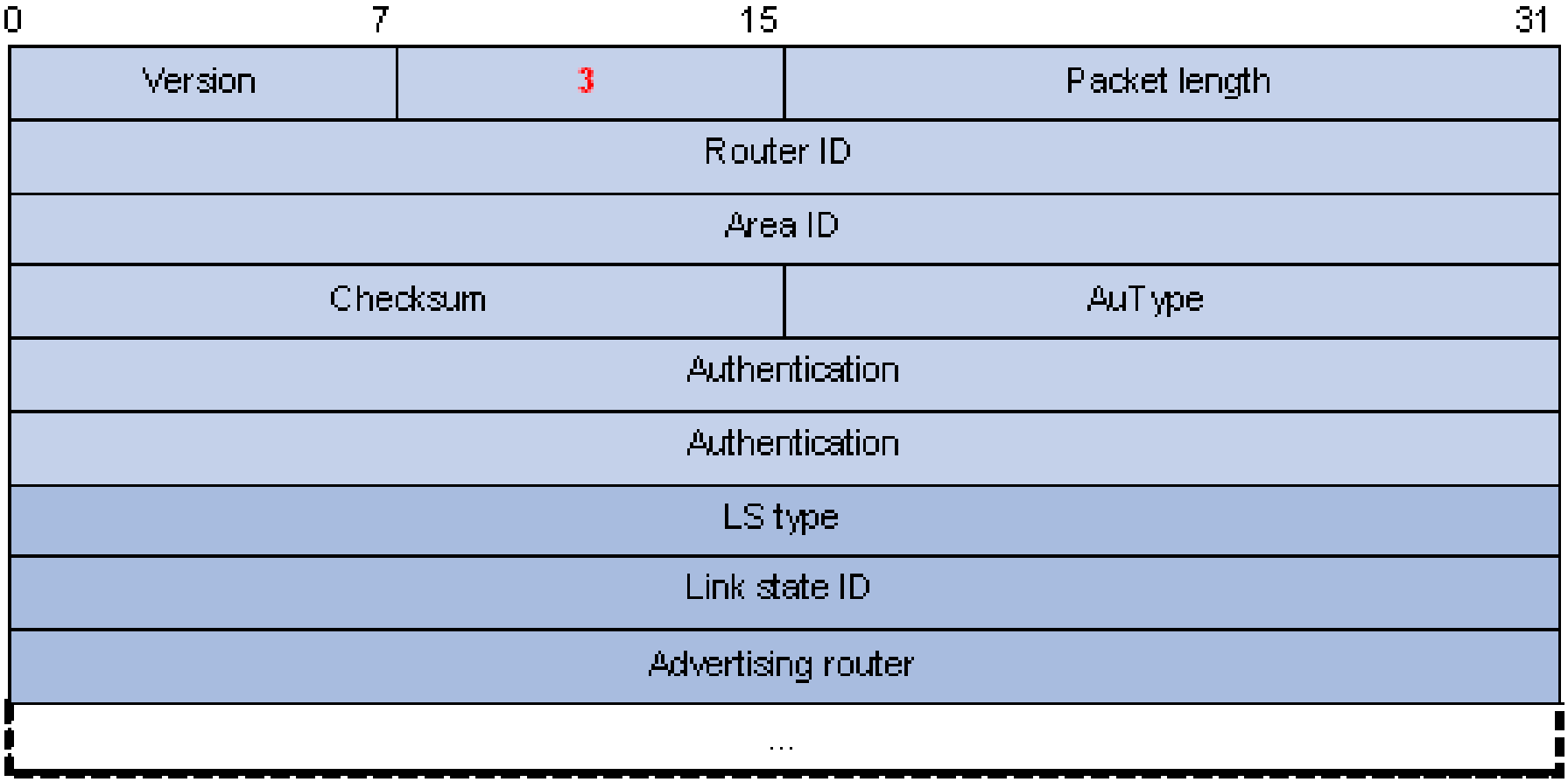
- **Interface MTU**: Size in bytes of the largest IP datagram that can be sent out the associated interface, without fragmentation.
- **I (Initial):** The Init bit, which is set to 1 if the packet is the first packet of database description packets, and set to 0 if not.
- **M (More):** The More bit, which is set to 0 if the packet is the last packet of DD packets, and set to 1 if more DD Packets are to follow.
- **MS (Master/Slave):** The Master/Slave bit. When set to 1, it indicates that the router is the master during the database exchange process. Otherwise, the router is the slave.
- **DD Sequence Number**: Used to sequence the collection of database description packets for ensuring reliability and intactness of DD packets between the master and slave. The initial value is set by the master. The DD sequence number then increments until the complete database description has been sent.

OSPFv2 LSR Packet

- After exchanging DD packets, any two routers know which LSAs of the peer routers are missing from the local LSDBs. In this case, they send LSR (link state request) packets, requesting the missing LSAs. The packets contain the digests of the missing LSAs.

- The following figure shows the LSR packet format.

OSPFv2 LSR Packet



OSPFv2 LSR Packet

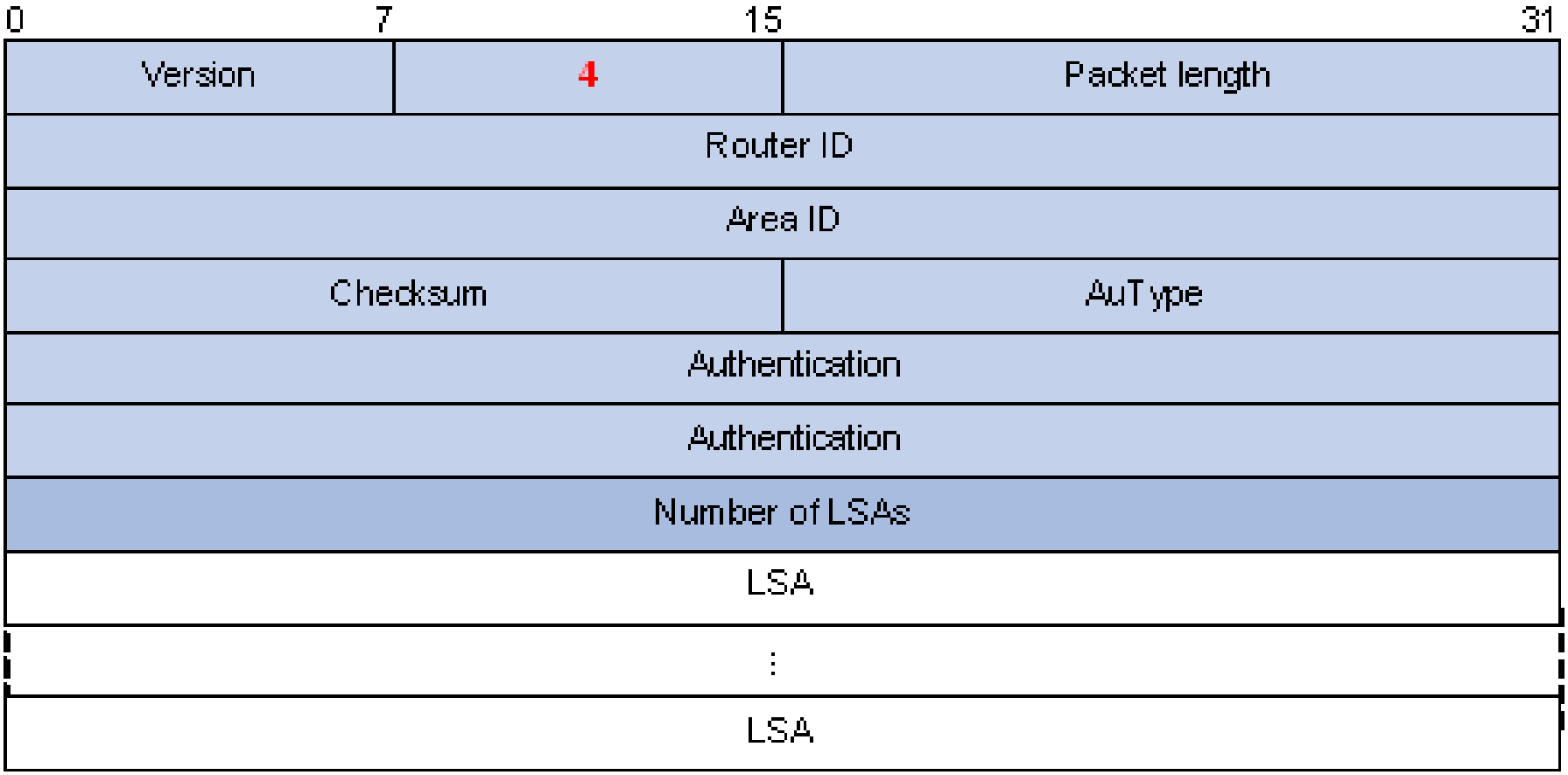
- **LS type**: Type number of the LSA to be requested. Type 1 for example indicates the Router LSA.
- **Link State ID**: Determined by LSA type.
- **Advertising Router**: ID of the router that sent the LSA.

OSPFv2 LSU Packet

- LSU (Link State Update) packets are used to send the requested LSAs to peers, and each packet carries a collection of LSAs.

- The LSU packet format is shown below.

OSPFv2 LSU Packet

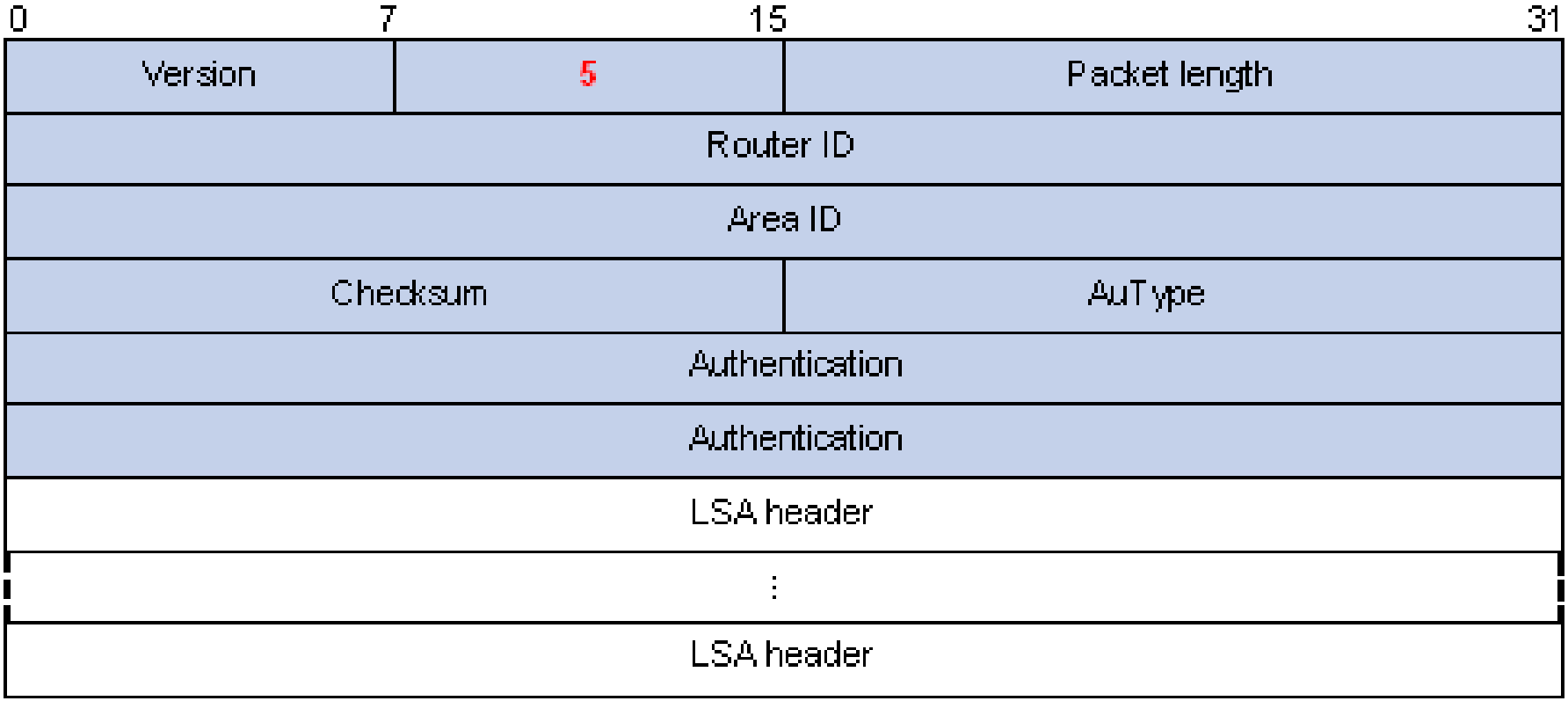


OSPFv2 LSAck Packet

- LSAack (Link State Acknowledgment) packets are used to acknowledge received LSU packets, contents including LSA headers to describe the corresponding LSAs. Multiple LSAs can be acknowledged in a single Link State Acknowledgment packet.

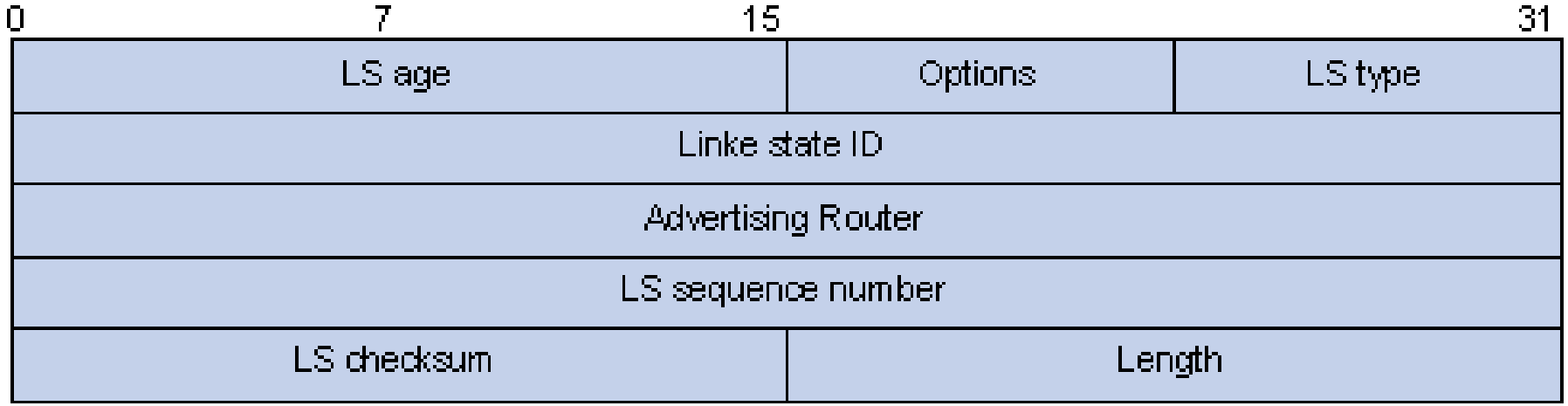
- The following figure gives its format.

OSPFv2 LSAck Packet



OSPFv2 LSA Header Format

- All LSAs have the same header, as shown in the following figure.



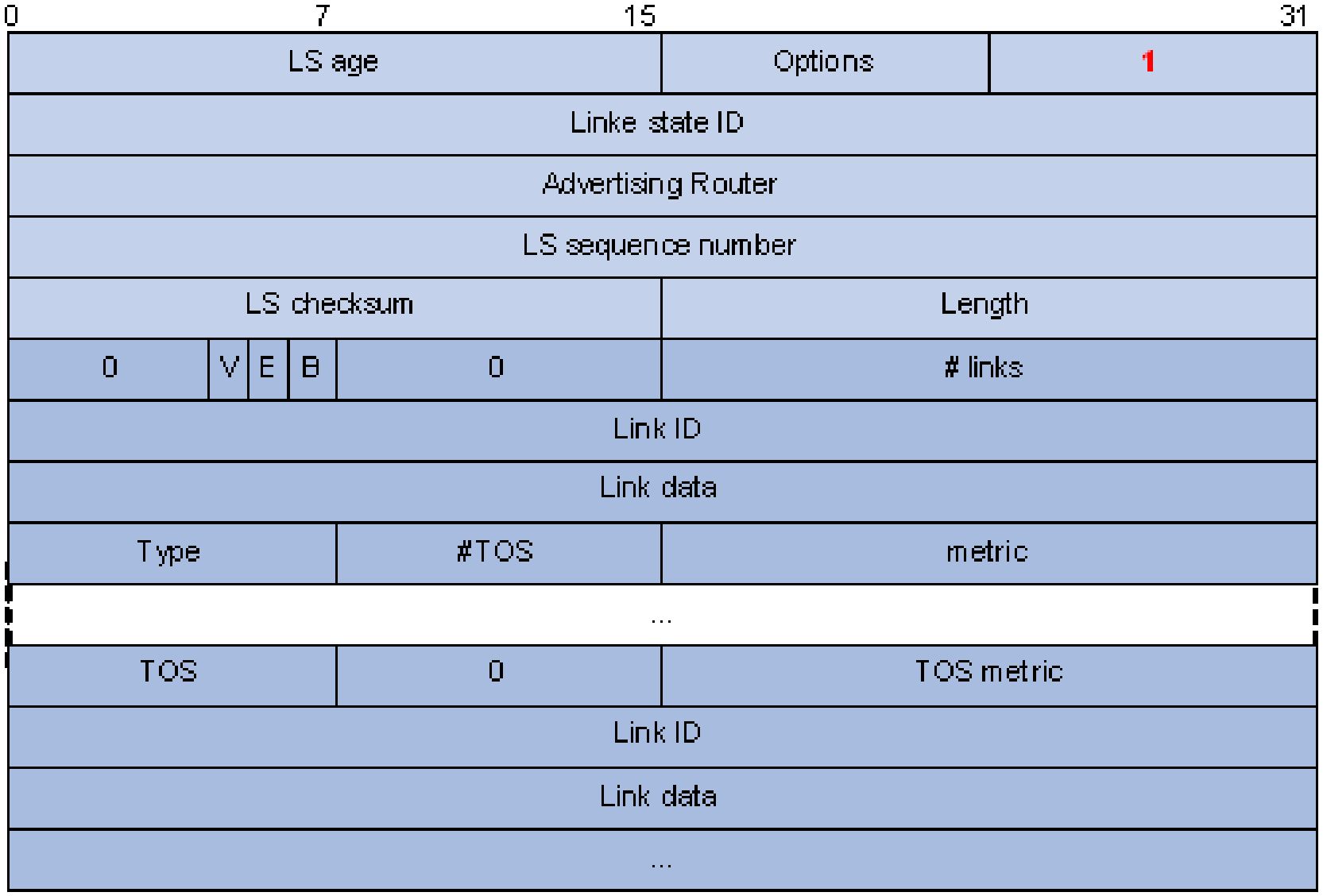
OSPFv2 LSA Header Format

- **LS age**: Time in seconds elapsed since the LSA was originated. A LSA ages in the LSDB (added by 1 per second), but does not in transmission.
- **LS type**: Type of the LSA.
- **Link State ID**: The contents of this field depend on the LSA's type
- **LS sequence number**: Used by other routers to judge new and old LSAs.
- **LS checksum**: Checksum of the LSA except the LS age field.
- **Length**: Length in bytes of the LSA, including the LSA header.

OSPFv2 Type of Common LSAs

- (1)Router LSA
- (2) Network LSA
- (3,4) Summary LSA
- (5) AS External LSA
- (7) NSSA External LSA

OSPFv2 Router LSA

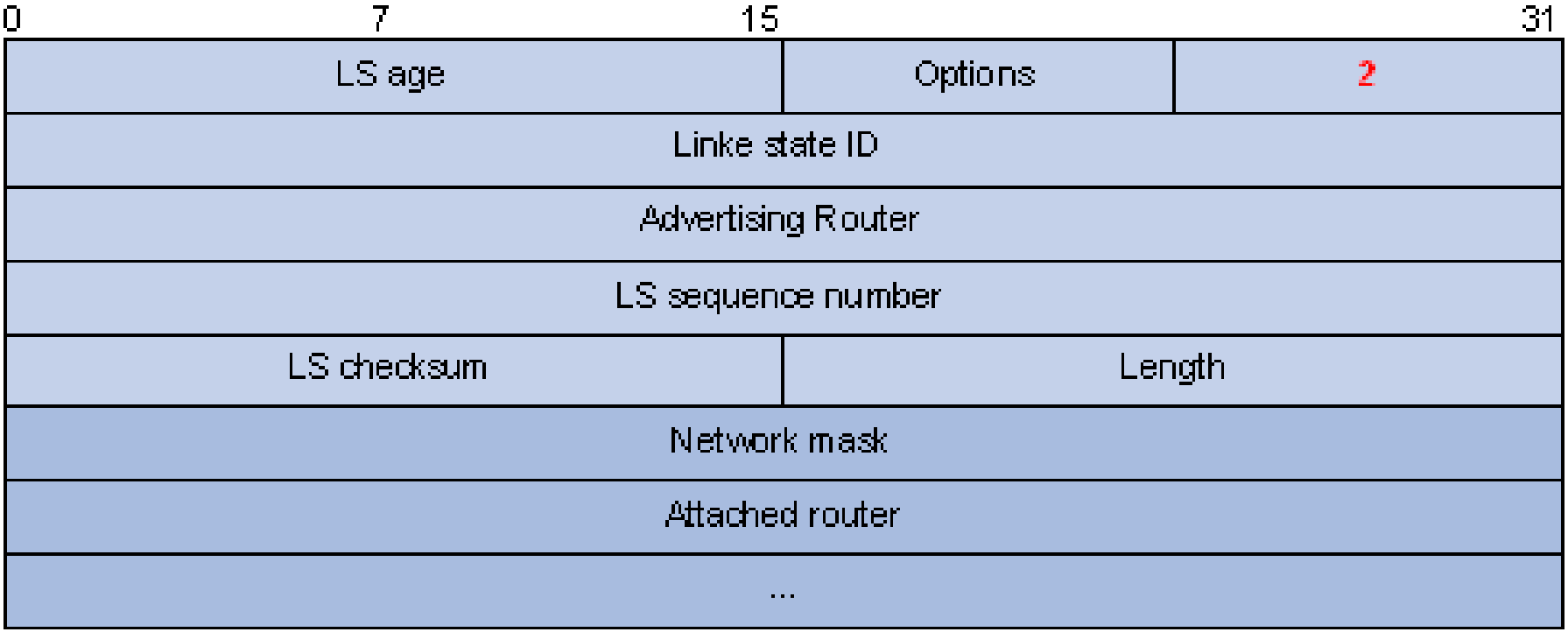


OSPFv2 Router LSA

- **Link State ID**: ID of the router that originated the LSA.
- **V (Virtual Link)**: Set to 1 if the router that originated the LSA is a virtual link endpoint.
- **E (External)**: Set to 1 if the router that originated the LSA is an ASBR.
- **B (Border)**: Set to 1 if the router that originated the LSA is an ABR.
- **# links**: Number of router links (interfaces) to the area, described in the LSA.
- **Link ID**: Determined by Link type.
- **Link Data**: Determined by Link type.
- **Type**: Link type. A value of 1 indicates a point-to-point link to a remote router; a value of 2 indicates a link to a transit network; a value of 3 indicates a link to a stub network; a value of 4 indicates a virtual link.
- **#TOS**: Number of different TOS metrics given for this link.
- metric: Cost of using this router link.
- **TOS**: IP Type of Service that this metric refers to.
- **TOS metric**: TOS-specific metric information.

OSPFv2 Network LSA

- A Network LSA is originated by the DR on a broadcast or NBMA network. The LSA describes all routers attached to the network.



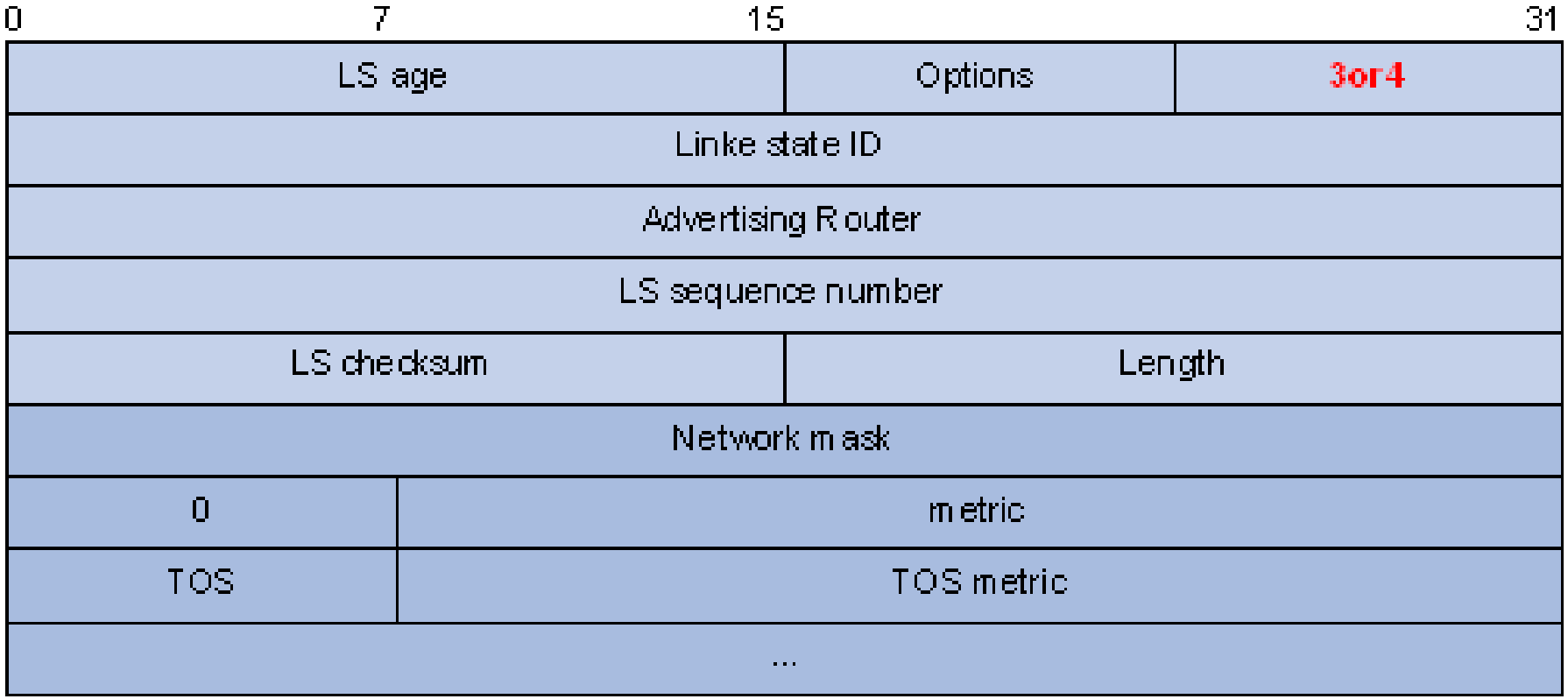
OSPFv2 Network LSA

- **Link State ID**: The interface address of the DR
- **Network Mask**: The mask of the network (a broadcast or NBMA network)
- **Attached Router**: The IDs of the routers, which are adjacent to the DR, including the DR itself

OSPFv2 Summary LSA

- Network summary LSAs (**Type-3 LSAs**) and ASBR summary LSAs (**Type-4 LSAs**) are originated by ABRs. Other than the difference in the Link State ID field, the format of type 3 and 4 summary-LSAs is identical.

OSPFv2 Summary LSA



OSPFv2 Summary LSA

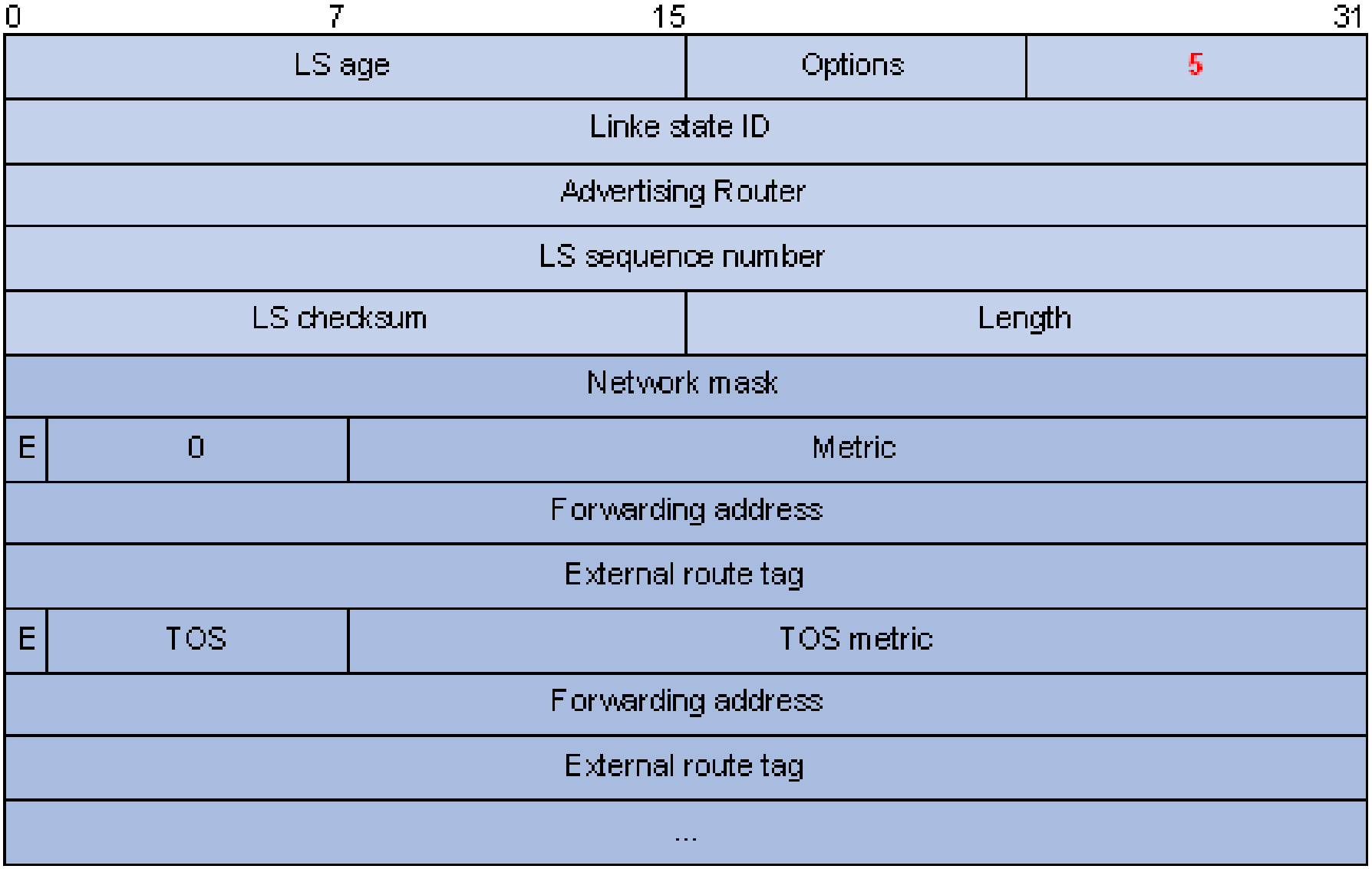
- **Link State ID**: For a Type-3 LSA, it is an IP address outside the area; for a type 4 LSA, it is the router ID of an ASBR outside the area.
- **Network Mask**: The network mask for the type 3 LSA; set to 0.0.0.0 for the Type-4 LSA
- **metric**: The metric to the destination

Note: A Type-3 LSA can be used to advertise a default route, having the Link State ID and Network Mask set to 0.0.0.0

OSPFv2 AS External LSA

- An AS external LSA originates from an ASBR, describing routing information to a destination outside the AS.

OSPFv2 AS External LSA



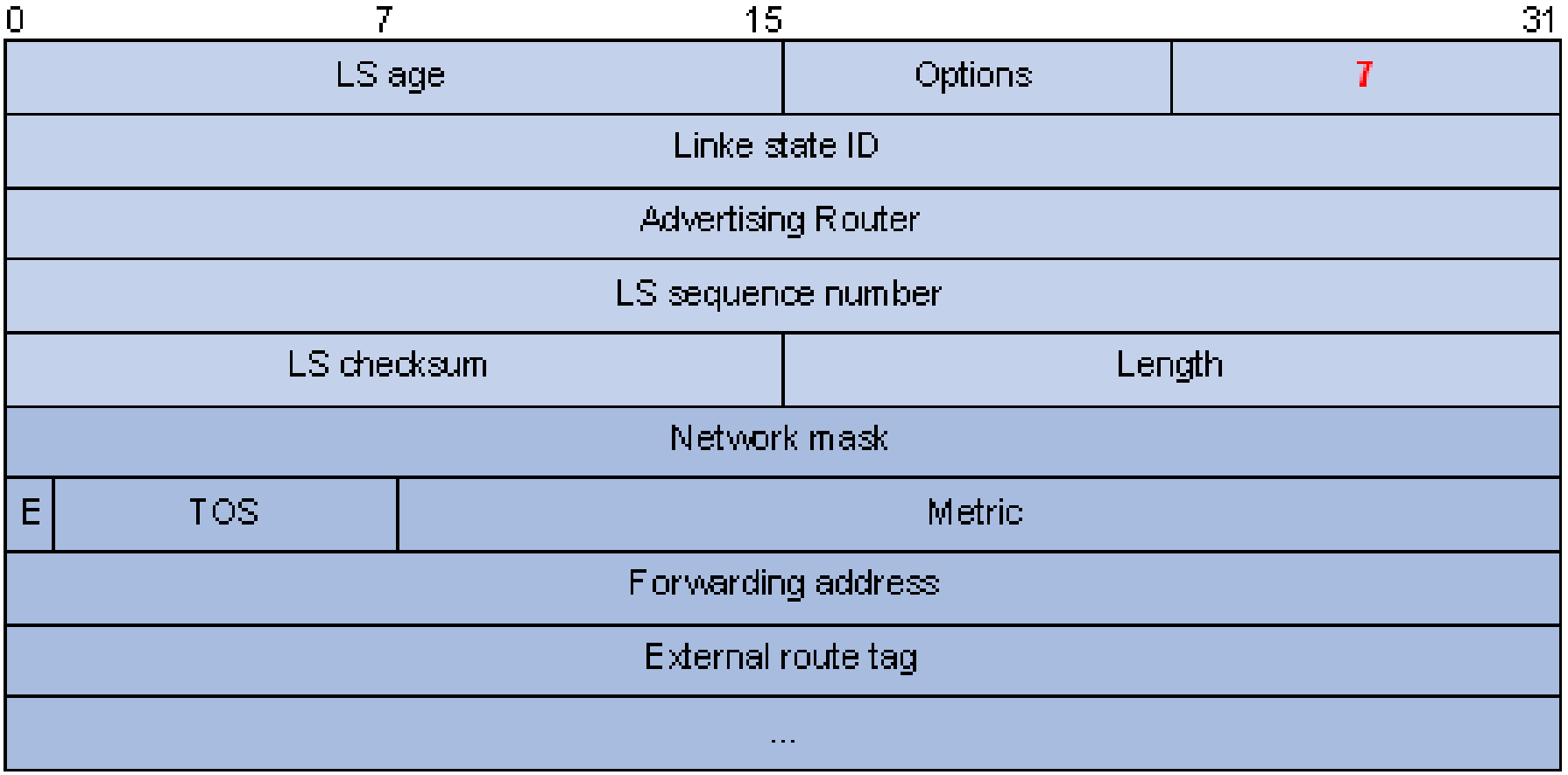
OSPFv2 AS External LSA

- **Link State ID**: The IP address of another AS to be advertised. When describing a default route, the Link State ID is always set to Default Destination (0.0.0.0) and the Network Mask is set to 0.0.0.0
- **Network Mask**: The IP address mask for the advertised destination
- **E (External Metric)**: The type of the external metric value, which is set to 1 for type 2 external routes, and set to 0 for type 1 external routes. Refer to Route types for description about external route types
- **metric**: The metric to the destination
- **Forwarding Address**: Data traffic for the advertised destination will be forwarded to this address
- **External Route Tag**: A tag attached to each external route. This is not used by the OSPF protocol itself. It may be used to manage external routes.

OSPFv2 NSSA external LSA

- An NSSA external LSA originates from the ASBR in a NSSA and is flooded in the NSSA area only.
- It has the same format as the AS external LSA.

OSPFv2 NSSA external LSA



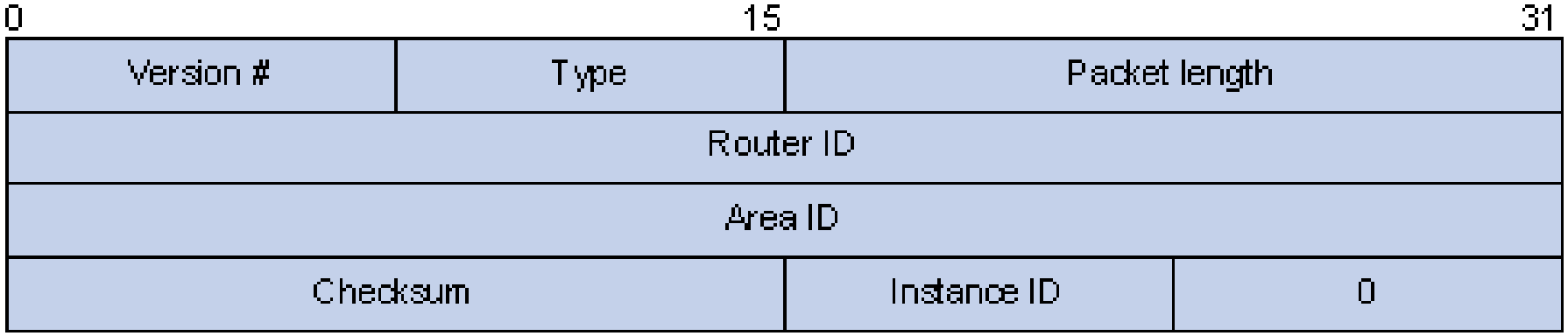
OSPFv3

OSPFv3 Related RFCs

- **RFC 2740**: OSPF for IPv6
- **RFC 2328**: OSPF Version 2

OSPFv3 Packet Header

- OSPFv3 has five types of packets: hello, DD, LSR, LSU, and LSAck.
- The five packets have the same packet header, which different from the OSPFv2 packet header is only 16 bytes in length, has no authentication field, but is added with an Instance ID field to support multi-instance per link.



OSPFv3 Packet Header

- **Version #:** Version of OSPF, which is 3 for OSPFv3.
- **Type**: Type of OSPF packet, from 1 to 5 are hello, DD, LSR, LSU, and LSAck respectively.
- **Packet Length**: Packet length in bytes, including header.
- **Instance ID**: Instance ID for a link.
- **0**: Reserved, which must be 0.

OSPFv3 LSA Types

OSPFv3 sends routing information in LSAs, which as defined in RFC2740 have the following types:

OSPFv3 LSA Types

- **Router-LSAs**: Originated by all routers. This LSA describes the collected states of the router's interfaces to an area. Flooded throughout a single area only.
- **Network-LSAs**: Originated for broadcast and NBMA networks by the Designated Router. This LSA contains the list of routers connected to the network. Flooded throughout a single area only.
- **Inter-Area-Prefix-LSAs**: Similar to Type 3 LSA of OSPFv2, originated by ABRs (Area Border Routers), and flooded throughout the LSA's associated area. Each Inter-Area-Prefix-LSA describes a route with IPv6 address prefix to a destination outside the area, yet still inside the AS (an inter-area route).
- **Inter-Area-Router-LSAs**: Similar to Type 4 LSA of OSPFv2, originated by ABRs and flooded throughout the LSA's associated area. Each Inter-Area-Router-LSA describes a route to ASBR (Autonomous System Boundary Router).
- **AS-external-LSAs**: Originated by ASBRs, and flooded throughout the AS (except Stub and NSSA areas). Each AS-external-LSA describes a route to another Autonomous System. A default route can be described by an AS external LSA.
- **Link-LSAs**: A router originates a separate Link-LSA for each attached link. Link-LSAs have link-local flooding scope. Each Link-LSA describes the IPv6 address prefix of the link and Link-local address of the router.
- **Intra-Area-Prefix-LSAs**: Each Intra-Area-Prefix-LSA contains IPv6 prefix information on a router, stub area or transit area information, and has area flooding scope. It was introduced because Router-LSAs and Network-LSAs contain no address information now.

OSPFv2 and OSPFv3 All LSA Types

LSA Type	OSPFv2 Description	OSPFv3 Description	OSPFv3 Type Code	Scope
LSA type 1	Router LSA	Router LSA	0x2001	Area scope
LSA type 2	Network LSA	Network LSA	0x2002	Area scope
LSA type 3	Network Summary LSA	Inter-Area-Prefix-LSA	0x2003	Area scope
LSA type 4	ASBR summary LSA	Inter-Area-Router-LSA	0x2004	Area scope
LSA type 5	AS external LSA	AS external LSA	0x4005	AS scope
LSA type 6	Group Membership LSA (N/A)	Group Membership LSA (N/A)	N/A	N/A
LSA type 7	NSSA External LSAs	Type 7 LSA	0x2007	Area scope
LSA type 8	External Attribute LSA (N/A)	Link local LSA	0x0008	**Link local scope**
LSA type 9	Opaque LSA (link scope)	Intra-area-Prefix-LSA	0x2009	Area scope
LSA type 10	Opaque LSA (area scope)			
LSA type 11	Opaque LSA (AS scope) Used for grace LSA	Used for Grace LSA	0x000b	NSF graceful restart

BGPv4

BGP Related RFCs

- **RFC 1771**: A Border Gateway Protocol 4 (BGP-4)
- **RFC 2858**: Multiprotocol Extensions for BGP-4
- **RFC 3392**: Capabilities Advertisement with BGP-4
- **RFC 2918**: Route Refresh Capability for BGP-4
- **RFC 2439**: BGP Route Flap Damping
- **RFC 1997**: BGP Communities Attribute
- **RFC 2796**: BGP Route Reflection
- **RFC 3065**: Autonomous System Confederations for BGP
- **RFC 4724**: Graceful Restart Mechanism for BGP

BGPv4 Packet Format

- BGPv4 packets are encapsulated after TCP segments. BGPv4 has the TCP port number **179**.
- The BGPv4 packet format is shown below:

Frame Header	IP Header (Protocol no = 6)	TCP (Port no = 179)	BGP Message	CRC
On a LAN, the BGPv4 packet is encapsulated in an Ethernet frame with a destination unicast MAC address: • **xx-xx-xx-xx-xx-xx**	The destination unicast IP address is set to **X.X.X.X** The IP protocol field is **6**.	Used TCP Port Number **179**	The BGP message contains the route and a lot of informations.	

BGP Messages

- BGP has five types of messages:

- **Open**
- **Update**
- **Notification**
- **Keep-alive**
- **Route-refresh**

BGP Message Types Overview

- BGP Message Types:

Open Message

Octets	16	2	1	1	2	2	4	1	7
	Marker	Length	Type	Version	AS	Hold Time	BGP ID	Optional Length	Optional

Update Message

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes length	Withdrawn Routes	Attribute Length	Attributes	NLRI

Notification Message

Octets	16	2	1	1	1	Variable
	Marker	Length	Type	Error Code	Error Sub-code	Diagnostic Data

Keepalive Message

Octets	16	2	1
	Marker	Length	Type

BGP Message Header

- All messages begin with the same 3 field headers

Open Message

Octets	16	2	1	1	2	2	4	1	7
	Marker	Length	Type	Version	AS	Hold Time	BGP ID	Optional Length	Optional

Update Message

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes length	Withdrawn Routes	Attribute Length	Attributes	NLRI

Notification Message

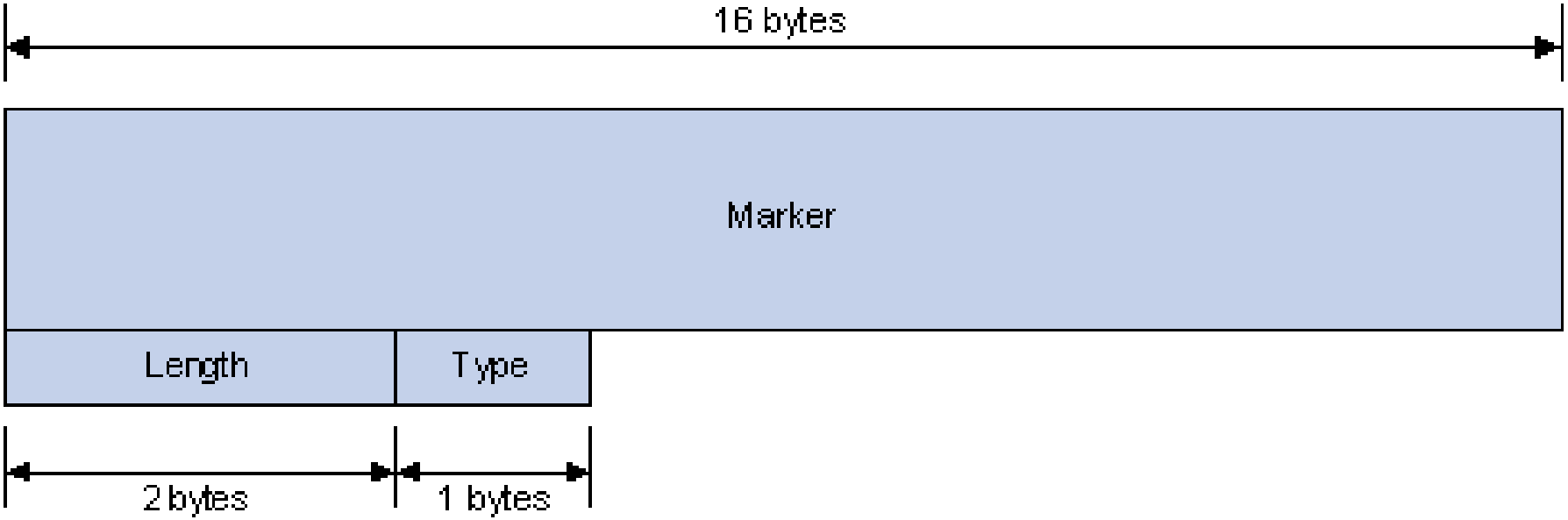
Octets	16	2	1	1	1	Variable
	Marker	Length	Type	Error Code	Error Sub-code	Diagnostic Data

Keepalive Message

Octets	16	2	1
	Marker	Length	Type

BGP Message Header Format

- They have the same header, as shown below:

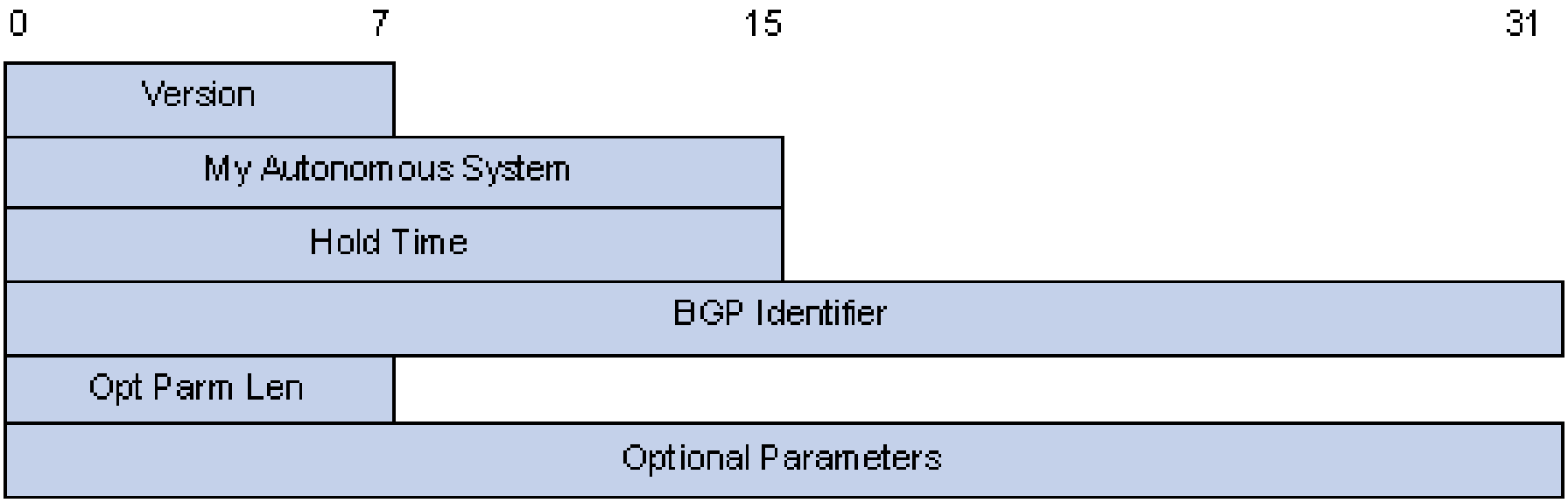


BGP Message Header Format

- **Marker**: The 16-byte field is used for BGP authentication. If no authentication information is available, then the Marker must be all ones.
- **Length**: The 2-byte unsigned integer indicates the total length of the message.
- **Type**: This 1-byte unsigned integer indicates the type code of the message. The following type codes are defined: 1–Open, 2-Update, 3-Notification, 4–Keepalive, and 5–Route-refresh. The former four are defined in RFC1771, the last one defined in RFC2918.

BGP Open Messages Format

- After a TCP connection is established, the first message sent by each side is an Open message for peer relationship establishment. The Open message contains the following fields:



BGP Open Messages Format

- **Version**: This 1-byte unsigned integer indicates the protocol version number of the message. The current BGP version is 4.
- **My Autonomous System**: This 2-byte unsigned integer indicates the Autonomous System number of the sender.
- **Hold Time**: When establishing peer relationship, two parties negotiate an identical hold time. If no Keepalive or Update is received from a peer after the hold time, the BGP connection is considered down.
- **BGP Identifier**: In IP address format, identifying the BGP router
- **Opt Parm Len** (Optional Parameters Length): Length of optional parameters, set to 0 if no optional parameter is available

BGP Open Messages Format

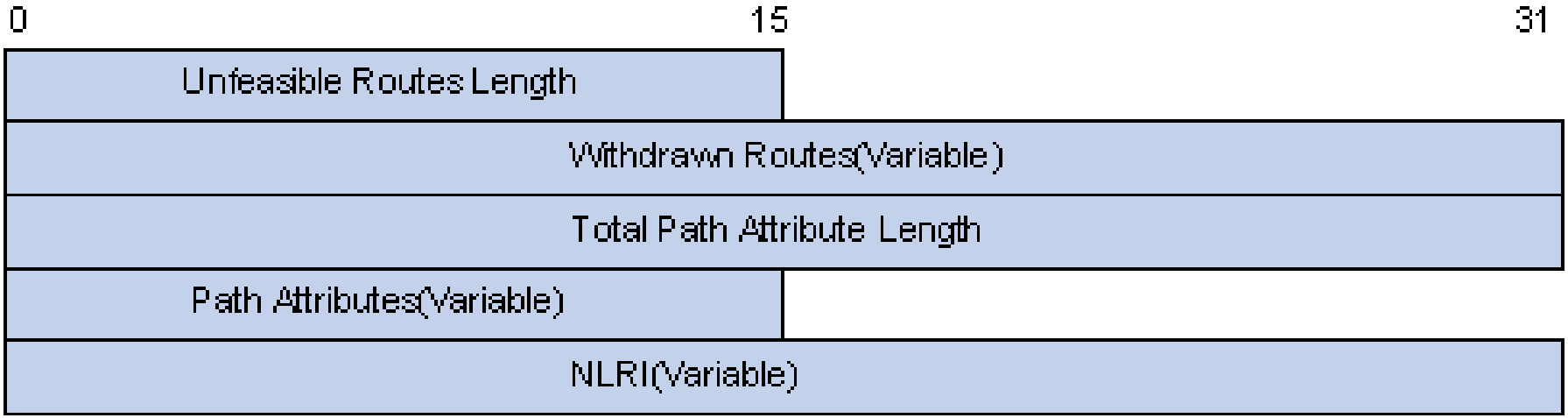
- Once a TCP connection has been established, the Open message is sent and includes a set of parameters that have to be agreed upon before a full BGP adjacency can be established.
- Once both BGP peers have agreed upon mutual capabilities, they can start exchanging routing information by means of BGP Update messages.

Open Message

Octets	16	2	1	1	2	2	4	1	7
	Marker	Length	Type	Version	AS	Hold Time	BGP ID	Optional Length	Optional

BGP Update Message Format

- The Update messages are used to exchange routing information between peers. It can advertise a feasible route or remove multiple unfeasible routes. Its format is shown below:



BGP Update Message Format

Each Update message can advertise a group of feasible routes with similar attributes, which are contained in the network layer reachable information (NLRI) field. The Path Attributes field carries attributes of these routes that are used by BGP for routing. Each message can also carry multiple withdrawn routes in the Withdrawn Routes field.

- **Unfeasible Routes Length**: The total length of the Withdrawn Routes field in bytes. A value of 0 indicates neither any route is being withdrawn from service, nor Withdrawn Routes field is present in this Update message.
- **Withdrawn Routes**: This is a variable length field that contains a list of IP prefixes of routes that are being withdrawn from service.
- **Total Path Attribute Length**: Total length of the Path Attributes field in bytes. A value of 0 indicates that no Network Layer Reachability Information field is present in this Update message.
- **Path Attributes**: List of path attributes related to NLRI. Each path attribute is a triple <attribute type, attribute length, attribute value> of variable length. BGP uses these attributes to avoid routing loops, perform routing and protocol extension.
- **NLRI** (Network Layer Reachability Information): Reachability information is encoded as one or more 2-tuples of the form <length, prefix>

BGP Update Message Format

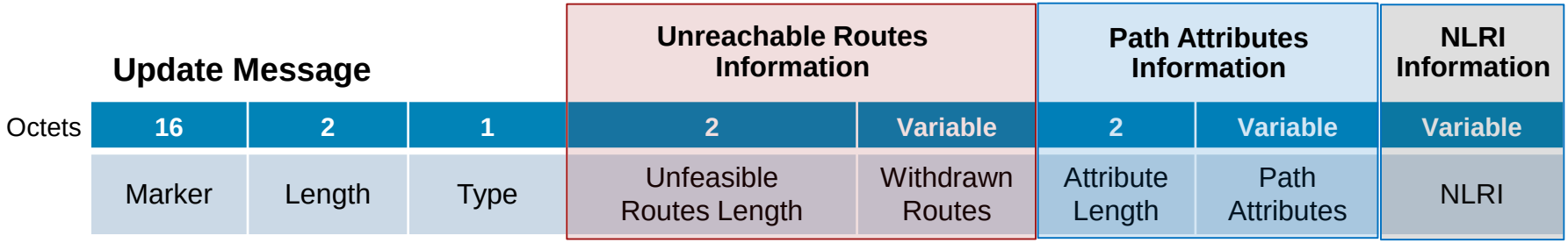
- Update messages contain all the information BGP uses to construct a loop-free picture of the internetwork.
- A BGP update message has information on one path only; multiple paths require multiple update messages.
  - All the attributes in the update message refer to that path, and the networks are those that can be reached through it.

Update Message

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes Length	Withdrawn Routes	Attribute Length	Path Attributes	NLRI

BGP Update Message Format

- An update message includes the following information:
  - Unreachable routes information
  - Path attribute information
  - Network-layer reachability information (NLRI)
    - This field contains a list of IP address prefixes that are reachable by this path.



BGP Update Message NLRI format

- The NLRI is a list of <**length**, **prefix**> tuples.
  - One tuple for each reachable destination.
  - The **prefix** represents the reachable destination
  - The prefix **length** represents the # of bits set in the subnet mask.

IP Address Subnet Mask	NLRI
10.1.1.0 255.255.255.0	**24, 10.1.1.0**
192.24.160.0 255.255.224.0	**19, 192.24.160.0**

BGP Update Message Path Attributes

- A BGP update message includes a variable-length sequence of path attributes describing the route.
- A path attribute consists of three fields:
  - Attribute type
  - Attribute length
  - Attribute value

BGP Attribute Type

- Type code 1 ORIGIN
- Type code 2 AS_PATH
- Type code 3 NEXT_HOP
- Type code 4 MULTI_EXIT_DISC
- Type code 5 LOCAL_PREF
- Type code 6 ATOMIC_AGGREGATE
- Type code 7 AGGREGATOR
- Type code 8 Community (Cisco-defined)
- Type code 9 Originator-ID (Cisco-defined)
- Type code 10 Cluster list (Cisco-defined)

Update Message

Path Attributes Information

Octets	16	2	1	2	Variable	2	Variable	Variable
	Marker	Length	Type	Unfeasible Routes Length	Withdrawn Routes	Attribute Length	Path Attributes	NLRI

BGP Update Message Path Attributes

- Some attributes are mandatory and automatically included in update messages while others are manually configurable.

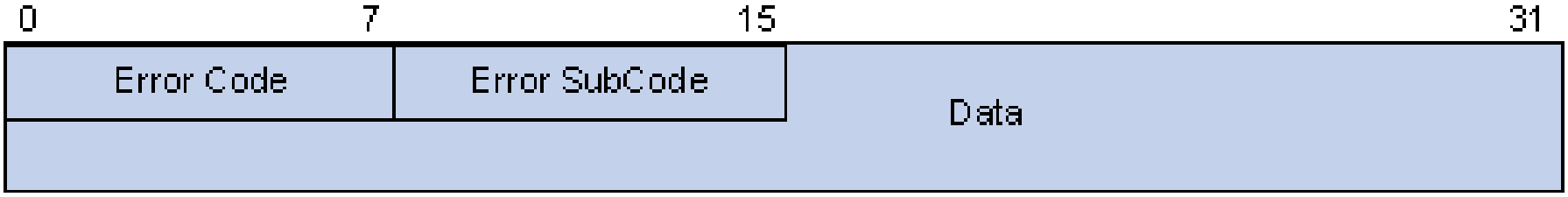
Attribute	EBGP	IBGP
AS_PATH	Well-known Mandatory	Well-known Mandatory
NEXT_HOP	Well-known Mandatory	Well-known Mandatory
ORIGIN	Well-known Mandatory	Well-known Mandatory
LOCAL_PREF	Not allowed	Well-known Discretionary
ATOMIC_AGGREGATE	Well-known Discretionary	Well-known Discretionary
AGGREGATOR	Optional Transitive	Optional Transitive
COMMUNITY	Optional Transitive	Optional Transitive
MULTI_EXIT_DISC	Optional Nontransitive	Optional Nontransitive

Automatically included in update message

Can be configured to help provide path control.

BGP Notification Message Format

- A Notification message is sent when an error is detected. The BGP connection is closed immediately after sending it. Notification message format is shown below:



BGP Notification Message Format

- **Error Code**: Type of Notification.
- **Error Subcode**: Specific information about the nature of the reported error.
- **Data**: Used to diagnose the reason for the Notification. The contents of the Data field depend upon the Error Code and
- **Error Subcode**. Erroneous part of data is recorded. The Data field length is variable.

BGP Notification Message Format

- A BGP notification message is sent when an error condition is detected.
  - The BGP connection is closed immediately after this is sent.
- Notification messages include an error code, an error subcode, and data related to the error.

Notification Message

Octets	16	2	1	1	1	Variable
	Marker	Length	Type	Error Code	Error Sub-code	Diagnostic Data

BGP Notification Message Format

Error Code	Error Subcode
1--Message Header Error	1--Connection Not Synchronized 2--Bad Message Length 3--Bad Message Type
2--OPEN Message Error	1--Unsupported Version Number 2--Bad Peer AS 3--Bad BGP Identifier 4--Unsupported Optional Parameter 5--Authentication Failure 6--Unacceptable Hold Time
3--UPDATE Message Error	1--Malformed Attribute List 2--Unrecognized Well-Known Attribute 3--Missing Well-Known Attribute 4--Attribute Flags Error 5--Attribute Length Error 6--Invalid Origin Attribute 7--AS Routing Loop 8--Invalid NEXT_HOP Attribute 9--Optional Attribute Error 10--Invalid Network Field 11--Malformed AS_path
4--Hold Timer Expired	NOT applicable
5--Finite State Machine Error (for errors detected by the FSM)	NOT applicable
6--Cease (for fatal errors besides the ones already listed)	NOT applicable

BGP Keepalive Message Format

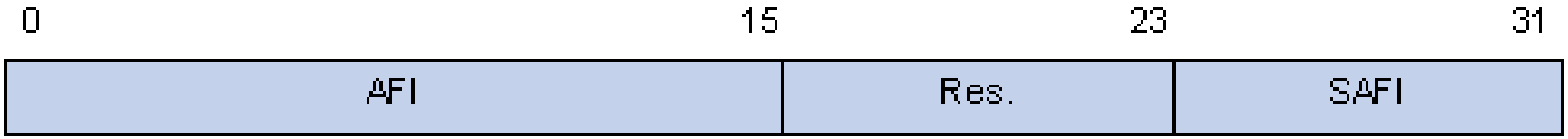
- Keepalive messages are sent between peers to maintain connectivity. Its format contains only the **Message Header**.
- Keepalive messages are sent between peers every 60 seconds (by default) to maintain connections.
- The message consists of only a message header (19 bytes).
  - Hold time is three times the KEEPALIVE timer of 60 seconds.
  - If the periodic timer = 0, no keepalives are sent.
  - Recommended keepalive interval is one-third of the hold time interval.

Keepalive Message

Octets	16	2	1
	Marker	**Length**	**Type**

BGP Route-Refresh Message Format

- A route-refresh message is sent to a peer to request the resending of the specified address family routing information. Its format is shown below:



- **AFI**: Address Family Identifier.
- **Res**: Reserved. Set to 0.
- **SAFI**: Subsequent Address Family Identifier.

IS-IS

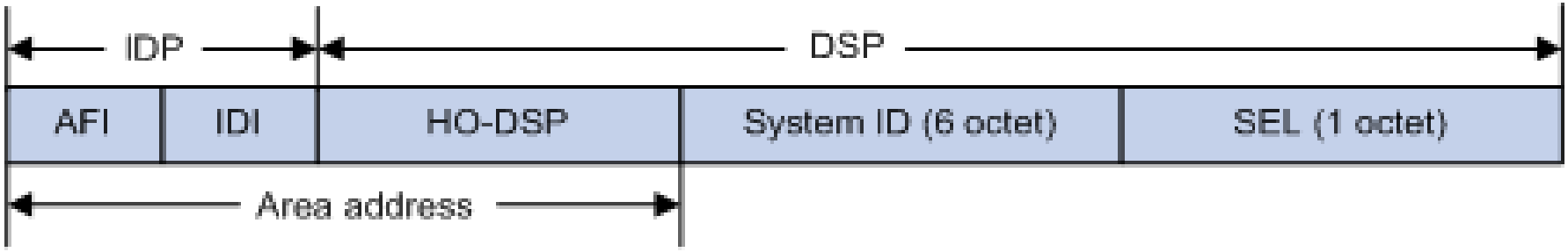
IS-IS Related RFCs

- **RFC 1195**: Use of OSI IS-IS for Routing in TCP/IP and Dual Environments
- **RFC 2763**: Dynamic Hostname Exchange Mechanism for IS-IS
- **RFC 2966**: Domain-wide Prefix Distribution with Two-Level IS-IS
- **RFC 2973**: IS-IS Mesh Groups
- **RFC 3277**: IS-IS Transient Blackhole Avoidance
- **RFC 3358**: Optional Checksums in ISIS
- **RFC 3373**: Three-Way Handshake for IS-IS Point-to-Point Adjacencies
- **RFC 3567**: Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication
- **RFC 3719**: Recommendations for Interoperable Networks using IS-IS
- **RFC 3786**: Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit
- **RFC 3787**: Recommendations for Interoperable IP Networks using IS-IS
- **RFC 3784**: IS-IS extensions for Traffic Engineering
- **RFC 3847**: Restart signaling for IS-IS

IS-IS Related Standards

- **ISO 10589** ISO IS-IS Routing Protocol
- **ISO 9542** ES-IS Routing Protocol
- **ISO 8348/Ad2** Network Services Access Points

IS-IS NSAP Address Format



IS-IS NSAP Address Format

- As shown in figure, an NSAP address consists of the Initial Domain Part (IDP) and the Domain Specific Part (DSP). The IDP is equal to the network ID of an IP address, and the DSP is equal to the subnet and host ID.
- **The IDP** includes the Authority and Format Identifier (AFI) and the Initial Domain Identifier (IDI).
- **The DSP** includes the High Order Part of DSP (HO-DSP), System ID and SEL, where the HO-DSP identifies the area, the System ID identifies the host, and the SEL identifies the type of service.

The IDP and DSP are variable in length. The length of an NSAP address varies from 8 bytes to 20 bytes

IS-IS NSAP Address Format

- **Area Address** The area address comprises the IDP and the HODSP of the DSP, which identify the area and the routing domain. Different routing domains cannot have the same area address.Typically, a router only needs one area address, and all nodes in the same routing domain must share the same area address. However, a router can have three area addresses at most to support smooth area merging, partitioning and switching.
- **System ID** A system ID identifies a host or router uniquely. It has a fixed length of 48 bits (6 bytes). The system ID of a router can be generated from the Router ID. For example, a router uses the IP address 168.10.1.1 of Loopback 0 as the Router ID. The system ID in IS-IS can be obtained in the following ways:
  - Extend each decimal number of the IP address to 3 digits by adding 0s from the left, like 168.010.001.001.
  - Divide the extended IP address into 3 sections with 4 digits in each section to get the system ID 1680.1000.1001.

If you use other methods to define a system ID, always make sure that it can uniquely identify a host or router.

IS-IS NSAP Address Format

- **SEL** The NSAP Selector (SEL), or the N-SEL, is similar to the protocol identifier in IP. Different transport layer protocols correspond to different SELs. All SELs in IP are 00.
- **Routing method** Because the area information is identified in IS-IS addresses, a Level-1 router can easily identify packets destined to other areas.
  - A Level-1 router makes routing decisions based on the system ID. If the destination is not in the area, the packet is forwarded to the nearest Level-1-2 router.
  - A Level-2 router routes packets across areas according to the area address.

IS-IS NET Address Format

A network entity title (NET) indicates the network layer information of an IS and does not include transport layer information. It is a special NSAP address with the SEL being 0. The length of the NET is equal to the NSAP and is in the range 8 bytes to 20 bytes.A NET comprises the following parts:

- **Area ID**—Its length is in the range of 1 to 13 bytes.
- **System ID**—A system ID uniquely identifies a host or router in the area and has a fixed 6-byte length.
- **SEL**—It has a value of 0 and a fixed 1-byte length.

For example, a **NET** is **ab.cdef.1234.5678.9abc.00**

Area ID is ab.cdef,

Aystem ID is 1234.5678.9abc,

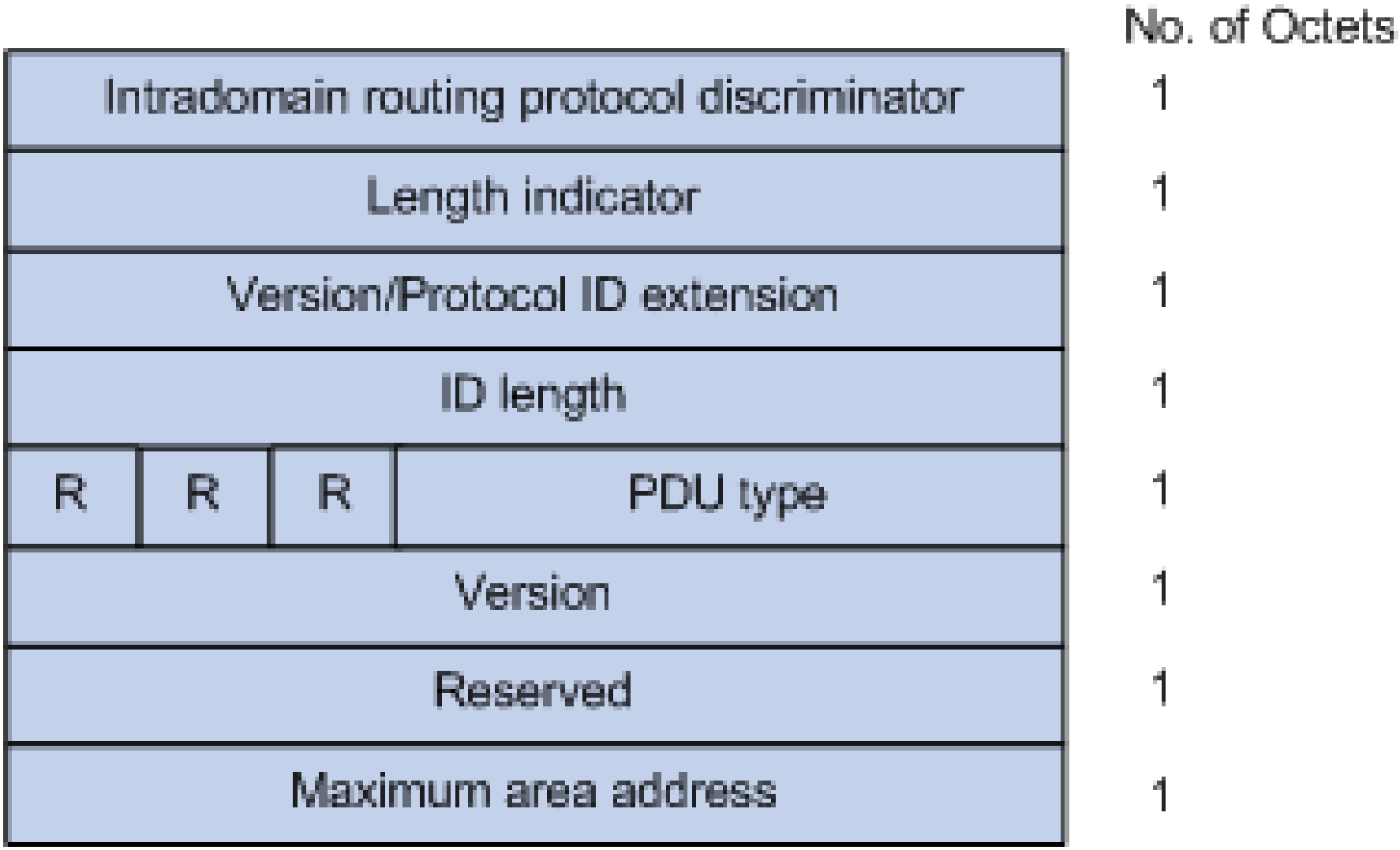
SEL is 00.

Typically, a router only needs one NET, but it can have three NETs at most for smooth area merging and partitioning. When you configure multiple NETs, make sure their system IDs are the same.

IS-IS PDU Header Format

- IS-IS packets are encapsulated into link layer frames. The Protocol Data Unit (PDU) consists of two parts, the headers and the variable length fields. The headers comprise the PDU common header and the PDU specific header. All PDUs have the same PDU common header. The specific headers vary by PDU type.

IS-IS PDU Common Header Format



IS-IS PDU Common Header Format

- **Intradomain Routing Protocol Discriminator—Set** to 0x83.
- **Length Indicator**—Length of the PDU header in bytes, including both common and specific headers.
- **Version/Protocol ID Extension**—Set to 1(0x01).
- **ID Length**—Length of the NSAP address and NET ID.
- **R(Reserved)**—Set to 0.
- **PDU Type**—For detailed information, see next page.
- **Version**—Set to 1(0x01).
- **Maximum Area Address**—Maximum number of area addresses supported.

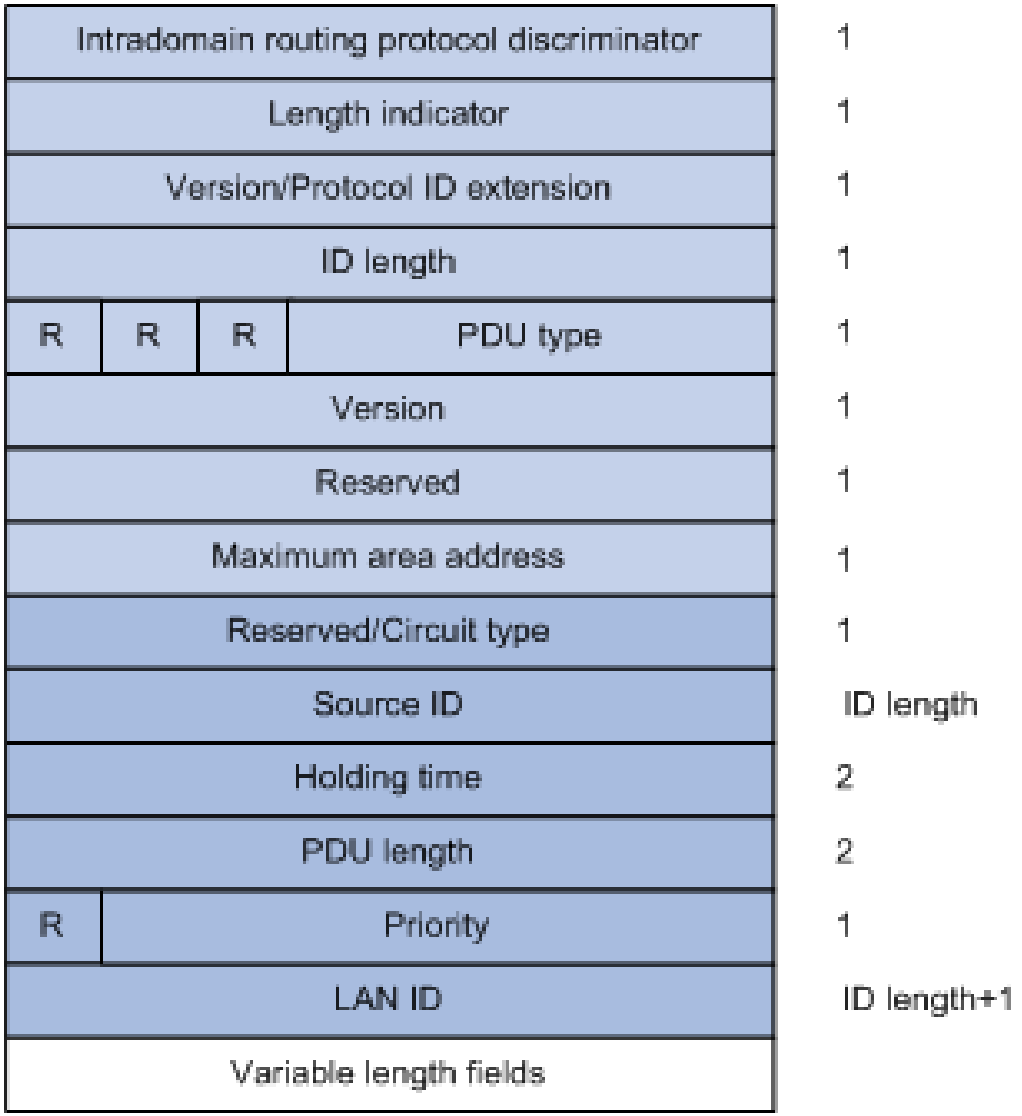
IS-IS PDU Common Header Format

Type	PDU Type	Acronym
15	Level-1 LAN IS-IS hello PDU	L1 LAN IIH
16	Level-2 LAN IS-IS hello PDU	L2 LAN IIH
17	Point-to-Point IS-IS hello PDU	P2P IIH
18	Level-1 Link State PDU	L1 LSP
20	Level-2 Link State PDU	L2 LSP
24	Level-1 Complete Sequence Numbers PDU	L1 CSNP
25	Level-2 Complete Sequence Numbers PDU	L2 CSNP
26	Level-1 Partial Sequence Numbers PDU	L1 PSNP
27	Level-2 Partial Sequence Numbers PDU	L2 PSNP

IS-IS Hello PDU

- Hello packets are used by routers to establish and maintain neighbor relationships. A hello packet is also an IS-to-IS hello PDU (IIH).
- For broadcast networks, the Level-1 routers use the Level-1 LAN IIHs; and the Level-2 routers use the Level-2 LAN IIHs.
- The P2P IIHs are used on point-to-point networks.

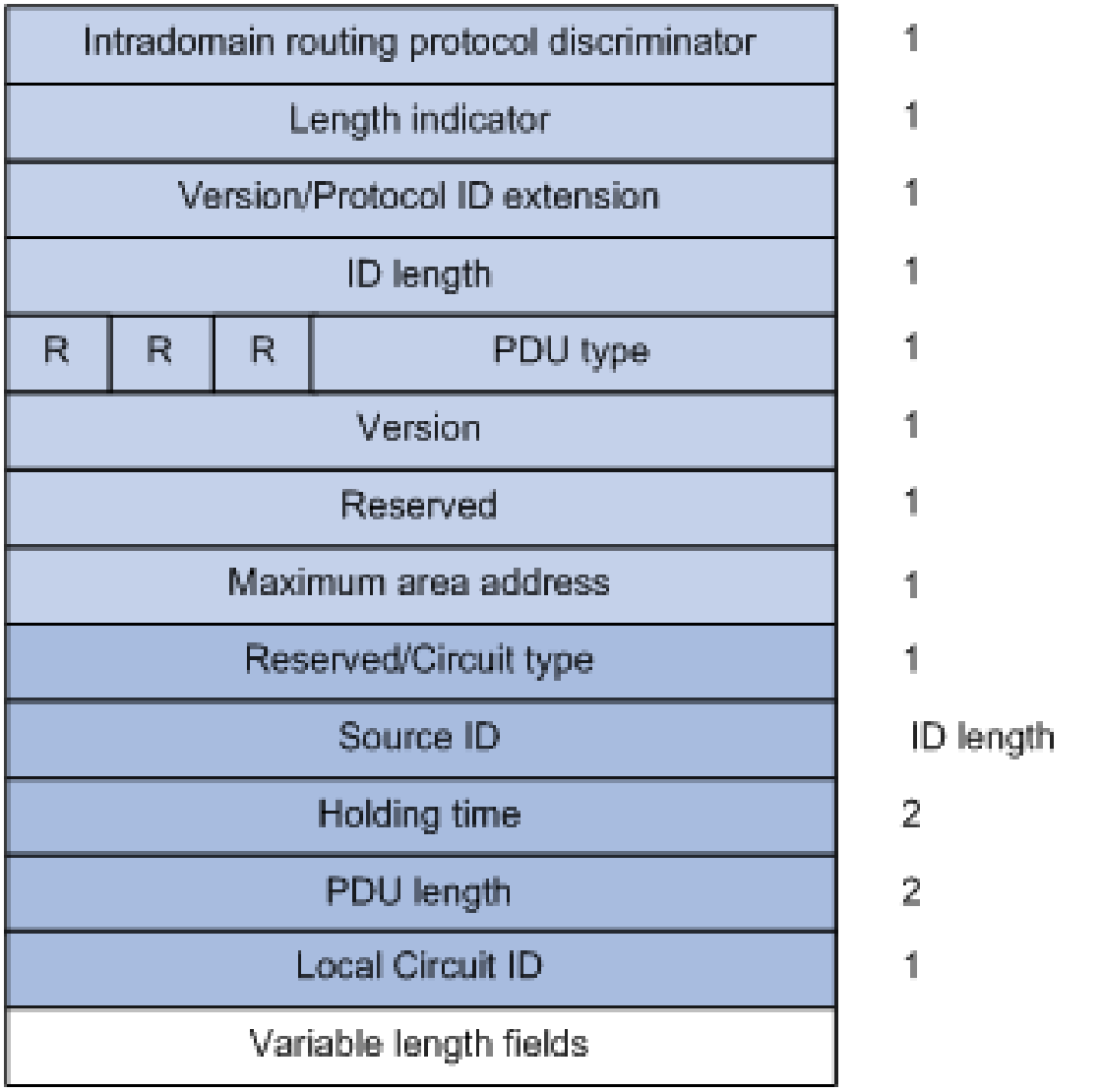
IS-IS L1/L2 LAN IIH Format



IS-IS L1/L2 LAN IIH Format

- **Reserved/Circuit Type**—The first 6 bits are reserved with a value of 0. The last 2 bits indicate the router type. 00 means reserved, 01 indicates L1, 10 indicates L2, and 11 indicates L1/2.
- **Source ID**—System ID of the router advertising the hello packet.
- **Holding Time**—If no hello packets are received from the neighbor within the holding time, the neighbor is considered down.
- **PDU Length**—Total length of the PDU in bytes.
- **Priority**—DIS priority.
- **LAN ID**—Includes the system ID and a one-byte pseudonode ID.

IS-IS P2P IIH Format



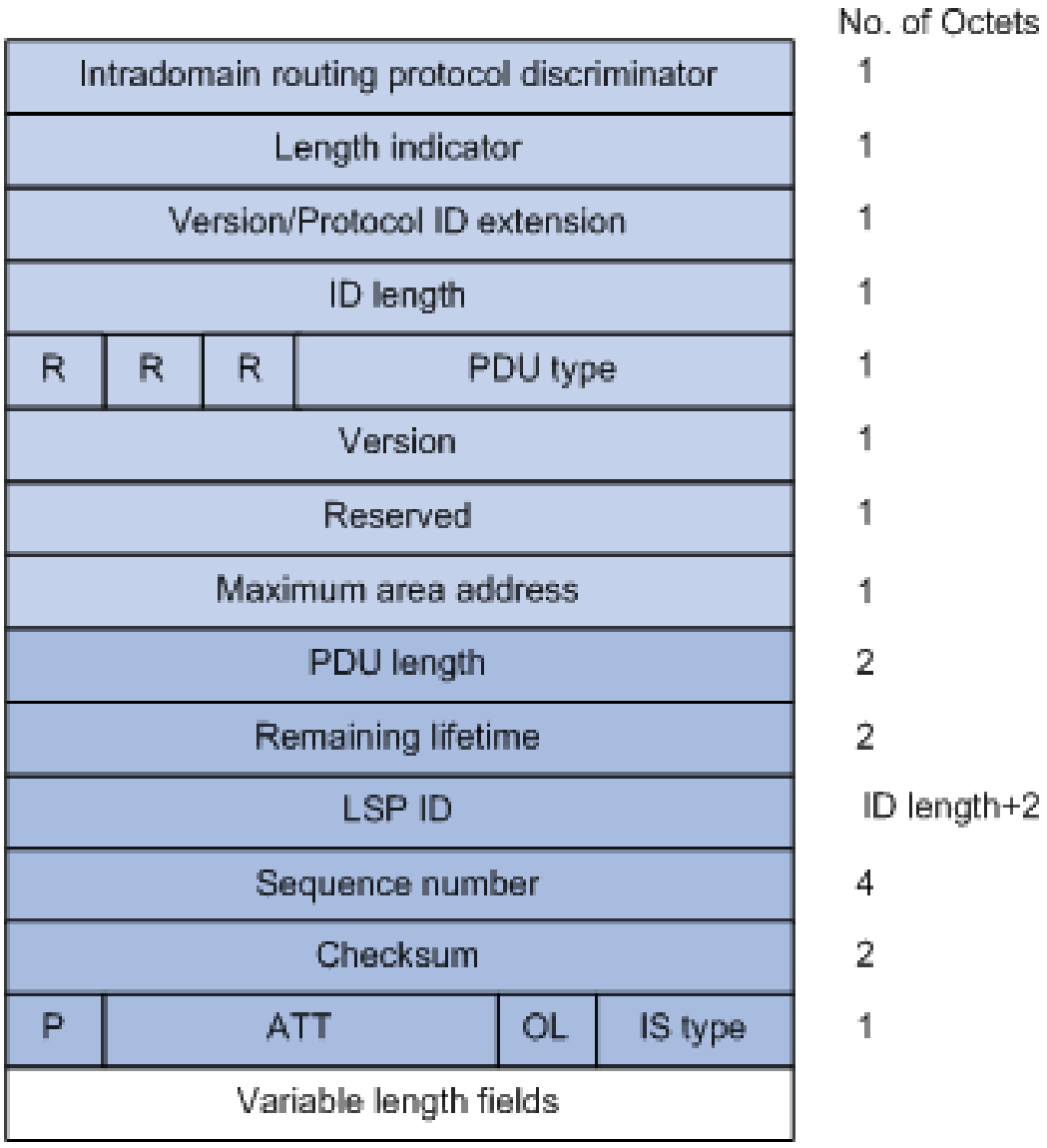
IS-IS P2P IIH Format

- Instead of the priority and LAN ID fields in the LAN IIH, the P2P IIH has a Local Circuit ID field.

IS-IS LSP PDU

- The Link State PDUs (LSP) carry link state information.
- LSP involves two types:
  - Level-1 LSP and Level-2 LSP
- The Level-2 LSPs are sent by the Level-2 routers, and the Level-1 LSPs are sent by the Level-1 routers.
- The level-1-2 router can send both types of LSPs.

IS-IS L1/L2 LSP Format



IS-IS L1/L2 LSP Format

- **PDU Length**—Total length of the PDU in bytes.
- Remaining Lifetime—LSP remaining lifetime in seconds.
- **LSP ID**—Consists of the system ID, the pseudonode ID (one byte) and the LSP fragment number (one byte).
- **Sequence Number**—LSP sequence number.
- Checksum—LSP checksum.
- **P (Partition Repair)**—Only for L2 LSPs. It indicates whether the router supports partition repair.
- **ATT (Attachment)**—Generated by a L1/L1 router for L1 LSPs only. It indicates that the router generating the LSP is connected to multiple areas.
- **OL (LSDB Overload)**—Indicates that the LSDB is not complete because the router has run out of memory. Other routers will not send packets to the overloaded router, except packets destined to the networks directly connected to the router.
- **IS Type**—Type of the router generating the LSP.

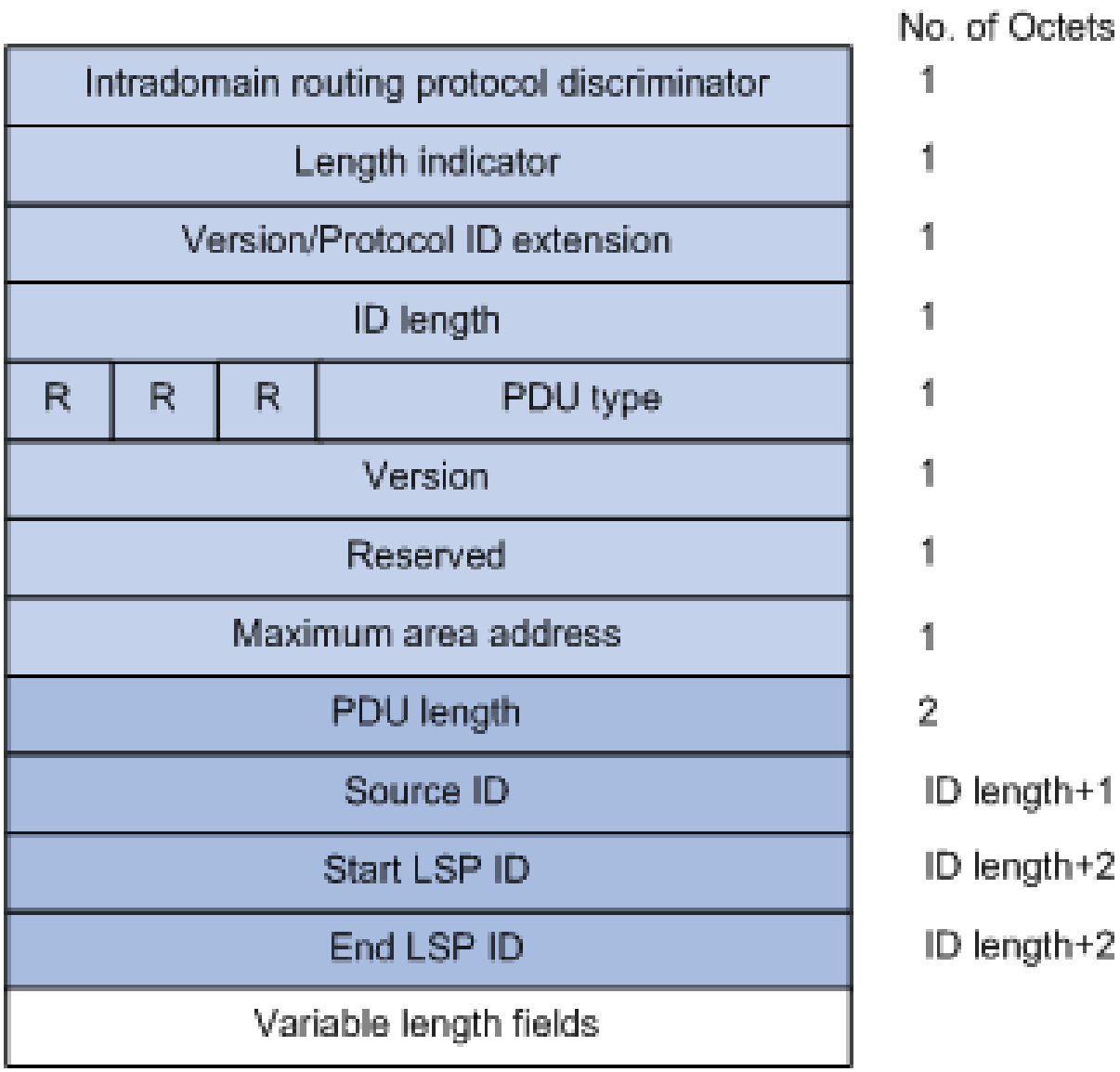
IS-IS SNP PDU

- A sequence number PDU (SNP) acknowledges the latest received LSPs. It is similar to an Acknowledge packet, but more efficient.

- SNP involves **Complete SNP (CSNP)** and **Partial SNP (PSNP)**, which are further divided into Level-1 CSNP, Level-2 CSNP, Level-1 PSNP and Level-2 PSNP.

- **CSNP** covers the summary of all LSPs in the LSDB to synchronize the LSDB between neighboring routers. On broadcast networks, CSNP is sent by the DIS periodically (10s by default). On point-to-point networks, CSNP is only sent during the adjacency establishment.

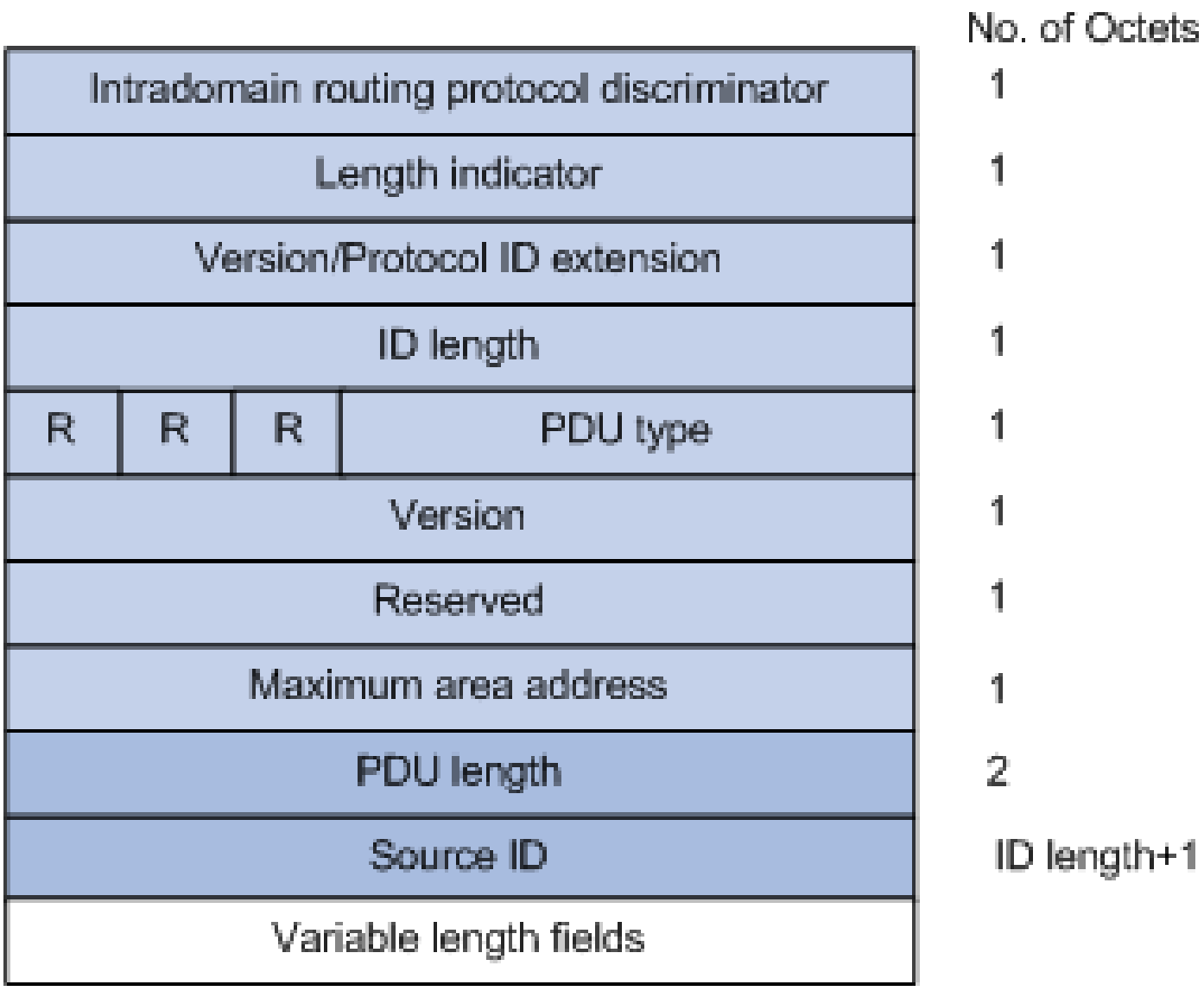
IS-IS L1/L2 CSNP Format



IS-IS SNP PDU

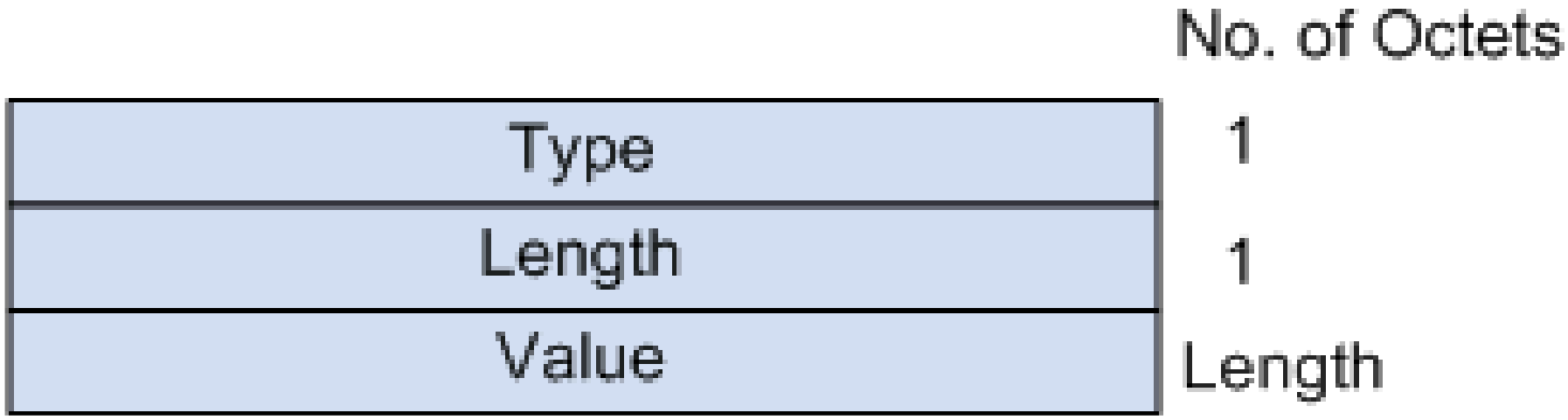
- **PSNP** only contains the sequence numbers of one or multiple latest received LSPs. It can acknowledge multiple LSPs at one time. When LSDBs are not synchronized, a PSNP is used to request new LSPs from neighbors.

IS-IS L1/L2 PSNP Format



IS-IS TLV Format

- The variable fields of PDU comprise multiple Type-Length-Value (TLV) triplets.



IS-IS TLV Format

- Below shows that different PDUs contain different TLVs. Code 1 to 10 of TLV are defined in **ISO 10589** (code 3 and 5 are not shown in the table), and others are defined in RFC 1195.

IS-IS TLV Format

TLV Code	Name	Applied PDU Type
1	Area Addresses	IIH, LSP
2	IS Neighbors (LSP)	LSP
4	Partition Designated Level2 IS	L2 LSP
6	IS Neighbors (MAC Address)	LAN IIH
7	IS Neighbors (SNPA Address)	LAN IIH
8	Padding	IIH
9	LSP Entries	SNP
10	Authentication Information	IIH, LSP, SNP
128	IP Internal Reachability Information	LSP
129	Protocols Supported	IIH, LSP
130	IP External Reachability Information	L2 LSP
131	Inter-Domain Routing Protocol Information	L2 LSP
132	IP Interface Address	IIH, LSP

EIGRP

EIGRP Related RFCs

- **RFC 7868**: Cisco's Enhanced Interior Gateway Routing Protocol (EIGRP)

EIGRP Packet Format

- EIGRP packets are directly encapsulated into IP packets. EIGRP has the IP protocol number **88**.
- The EIGRP packet format is shown below (taking a LSU packet as an example).

Frame Header	IP Header	Protocol Number (EIGRP = 88)	EIGRP Header	EIGRP Message	CRC
On a LAN, the EIGRP packet is encapsulated in an Ethernet frame with a destination multicast MAC address: **01-00-5E-00-00-0A**	The destination IP address is set to the multicast **224.0.0.10** and the EIGRP protocol field is **88**.		The EIGRP header identifies the type of EIGRP packet and autonomous system number.	The EIGRP message consists of the Type / Length / Value (TLV).	

EIGRP Packet Types

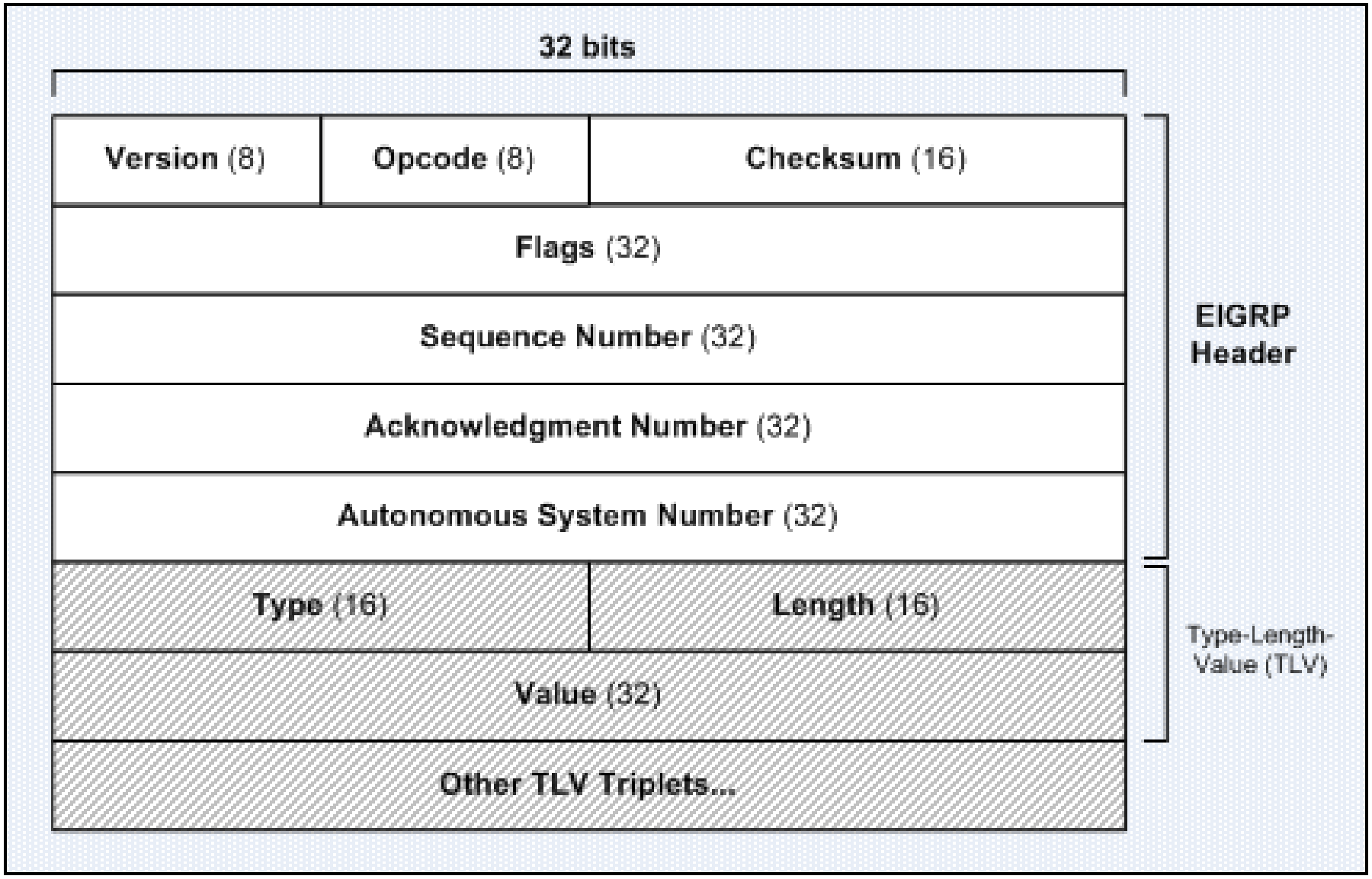
- EIGRP uses these 5 packet types to maintain its various tables and establish complex relationships with neighbor routers:

- **Hello**
- **Acknowledgment**
- **Update**
- **Query**
- **Reply**

EIGRP Packet Types

- **Hello** Used to discovery neighbor before establishing adjacency. EIGRP Hellos are sent as **multicasts** and contain an acknowledgment number of 0. EIGRP routers must form neighbor relationships before exchanging EIGRP updates.
- **Update** Used to communicate the routes that a particular router has used to converge. EIGRP Updates are sent as **multicasts** when a new route is discovered or when convergence is completed (the route becomes passive); and are sent as **unicasts** when synchronizing topology tables with neighbors upon the EIGRP startup. They are sent reliably between EIGRP routers.
- **Query** Used to query other EIGRP neighbors for a feasible successor when DUAL is re-computing a route in which the router does not have a feasible successor. EIGRP Queries are sent reliably as **multicasts**.
- **Reply** Sent as the response to an EIGRP Query packet. EIGRP Replies are sent reliably as **unicasts**.
- **Acknowledge** Used to acknowledge EIGRP Updates, Queries, and Replies; Hello and ACK packets do not require acknowledgment. ACKs are Hello packets that contain no data and a non-zero acknowledgment number and are sent as **unicasts**.

EIGRP Packet Format



EIGRP Packet Format

- **Version** Identifies the EIGRP process version. The current EIGRP version is 2.
- **Opcode** Identifies the EIGRP packet type – **Update (0x01), Query (0x03), Reply (0x04), Hello (0x05)**. It determines the TLVs that follow the EIGRP header.

Note: ACKs are Hello packets that contain a non-zero ACK number.

- **Checksum** The checksum of the entire EIGRP packet, excluding the IP header.
- **Flags** 1st LSB bit (0x00000001) – Init bit, used indicate the first set of routing updates upon establishing a new neighbor relationship. 2nd LSB bit (0x00000002) – Conditional Receive bit, used in the Cisco-proprietary reliable multicast protocol – Reliable Transport Protocol (RTP). Other bits are not being used.
- **SEQ and ACK** Used by RTP for reliable EIGRP message exchange.
- **AS Number** Identifies the autonomous system of an EIGRP packet. An EIGRP process only process EIGRP packets within an EIGRP domain (same AS number).
- **Type / Length / Value (TLV)** TLVs are comprise of a 16-bit Type field, a 16-bit Length field, and a vary number of fields depends on the type of TLV.

EIGRP Packet Format

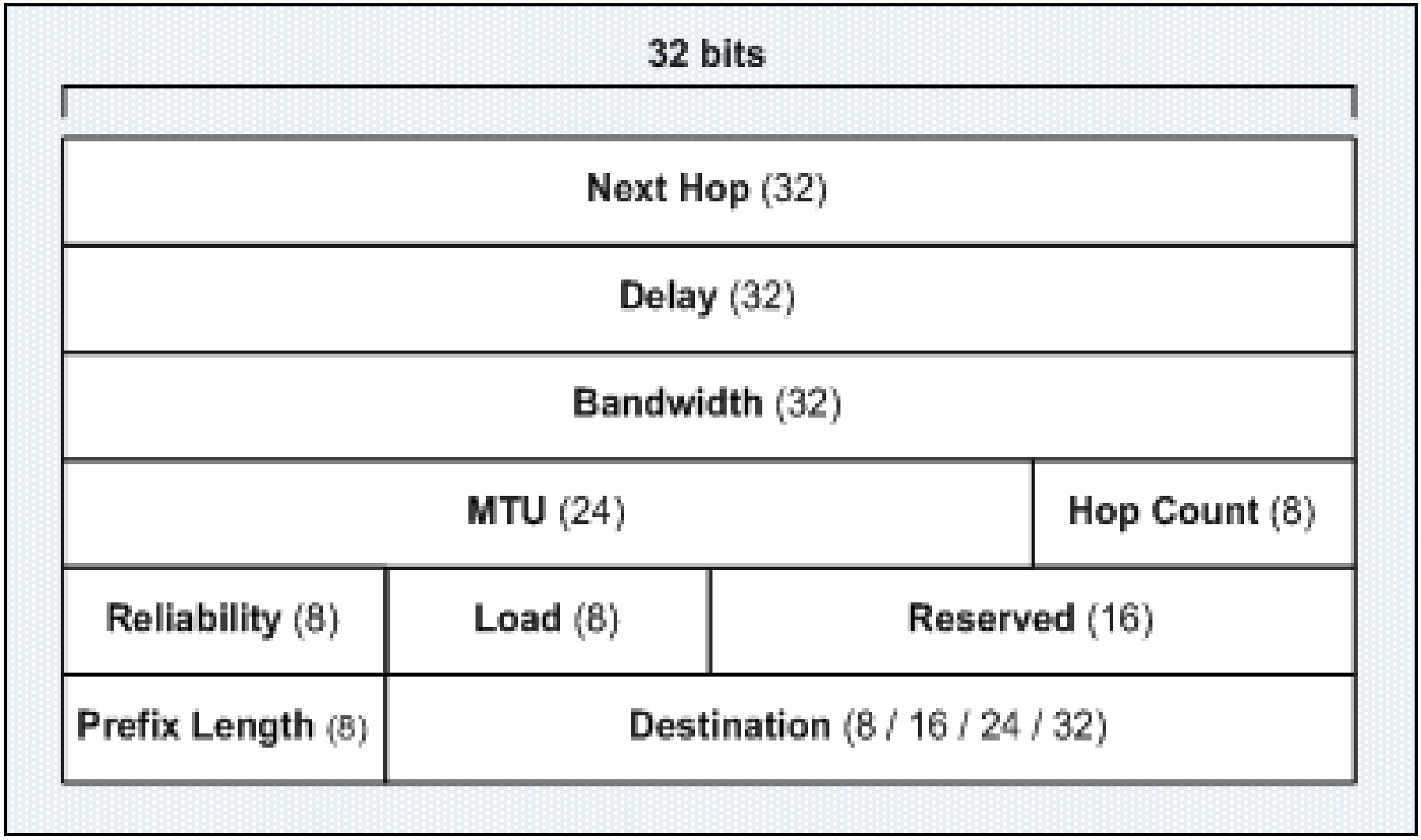
General TLVs:

- **0x0001** – EIGRP parameters – K values and hold time. Size of 12 bytes.
- **0x0002** – Message Digest 5 (MD5) authentication data. Size of 40 bytes.
- **0x0003** – Sequence. Used by RTP.
- **0x0004** – Software versions – IOS and EIGRP release versions. Size 8 bytes.
- **0x0005** – Next Multicast Sequence. Used by RTP.
- **0x0006** – EIGRP stub parameters.

IP TLVs:

- **0x0102** – IP internal route. Size of 28 bytes.
- **0x0103** – IP external route. Size of 48 bytes.

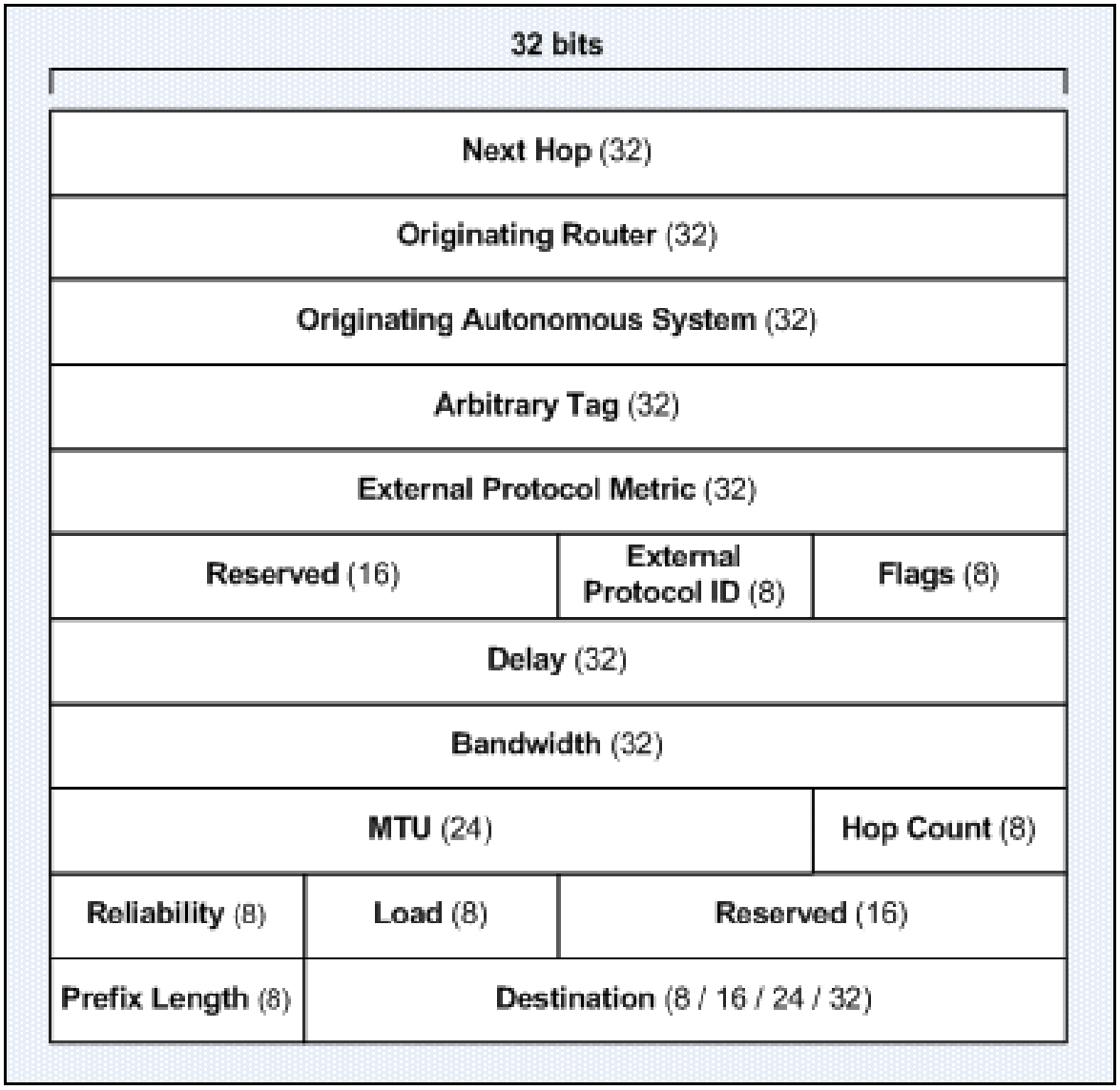
EIGRP IP **Internal** Route Packet Format



EIGRP IP **Internal** Route Packet Format

- **Next Hop** – is the next-hop, neighboring router, IP address.
- **Delay** – is the sum of the configured delays expressed in units of 10 microseconds. A delay of 0xFFFFFFFF, or 256, indicates an unreachable route.
- **Bandwidth** – is 256 x BW(min), or 2,560,000,000 divided by the lowest *configured* bandwidth of any interface along the route.
- **MTU** – is the smallest Maximum Transmission Unit of any link along the route to the destination. This value is not used for the metric calculation.
- **Hop Count** – is a number between 0x01 and 0xFF indicating the number of hops to the destination. A router will advertise a directly connected network with a hop count of 0.
- **Reliability** – is a number between 0x01 and 0xFF that reflects the total outgoing error rates of the interfaces along the route, calculated on a five-minute exponentially weighted average. 0xFF indicates a 100 percent reliable link.
- **Load** – is also a number between 0x01 and 0xFF, reflecting the total outgoing load of the interfaces along the route, calculated on a five-minute exponentially weighted average. 0x01 indicates a minimally loaded link.
- **Reserved** – is an unused field and is always 0x0000
- **Prefix Length** – specifies the number of network bits of the address mask.
- **Destination** – is the destination address of the route.

EIGRP IP **External** Route Packet Format



EIGRP IP **External** Route Packet Format

- **Next Hop** – is the next-hop IP address. On a multiaccess network, the router advertising the route might not be the best next-hop router to the destination. The Next Hop field allows the "bilingual" router to tell its EIGRP neighbors, "Use address A.B.C.D as the next hop instead of using my interface address."
- **Originating Router** – is the IP address or router ID of the router that redistributed the external route into the EIGRP autonomous system.
- **Originating Autonomous System Number** – is the autonomous system number of the router originating the route.
- **Arbitrary Tag** – may be used to carry a tag set by route maps.
- **External Protocol Metric** – is the metric of the external protocol.
- **Reserved** – is an unused field and is always 0x0000.
- **External Protocol ID** – specifies the protocol from which the external route was learned. 0x01 = IGRP, 0x02 = EIGRP, 0x03 = Static Route, 0x04 = RIP, 0x05 = Hello, 0x06 = OSPF, 0x07 = IS-IS, 0x08 = EGP, 0x09 = BGP, 0x0A = IDRP, 0x0B = Connected Link.
- **Flags** – currently constitute just two flags. If the right-most bit of the eight-bit field is set (0x01), the route is an external route. If the second bit is set (0x02), the route is a candidate default route.

Q&A