IPv6 (IP version 6) Essentials
Ch4: ICMPv6

Louis Chuang
Fu Jen Catholic University
EE ENCL

Introductoion (1)

- In IPv4, the ICMP (Internet Control Message Protocol) is very important:
  - It gives important information about the health of the network.
  - Such as ping (uses ICMP Echo Request & Echo Reply messages to test availability of a node).
- ICMPv6 (version 6) is for IPv6:
  - Node sends informational messages about the status of the network back to source node.
    - It reports errors if packets cannot be processed properly.
    - Ex) if a router cannot forward a packet because it is too large to be sent to the network:
      - The router sends back an ICMP message to the source host.
      - The source host can use this ICMP message to determine a better packet size and then resend the data.

Introductoion (2)

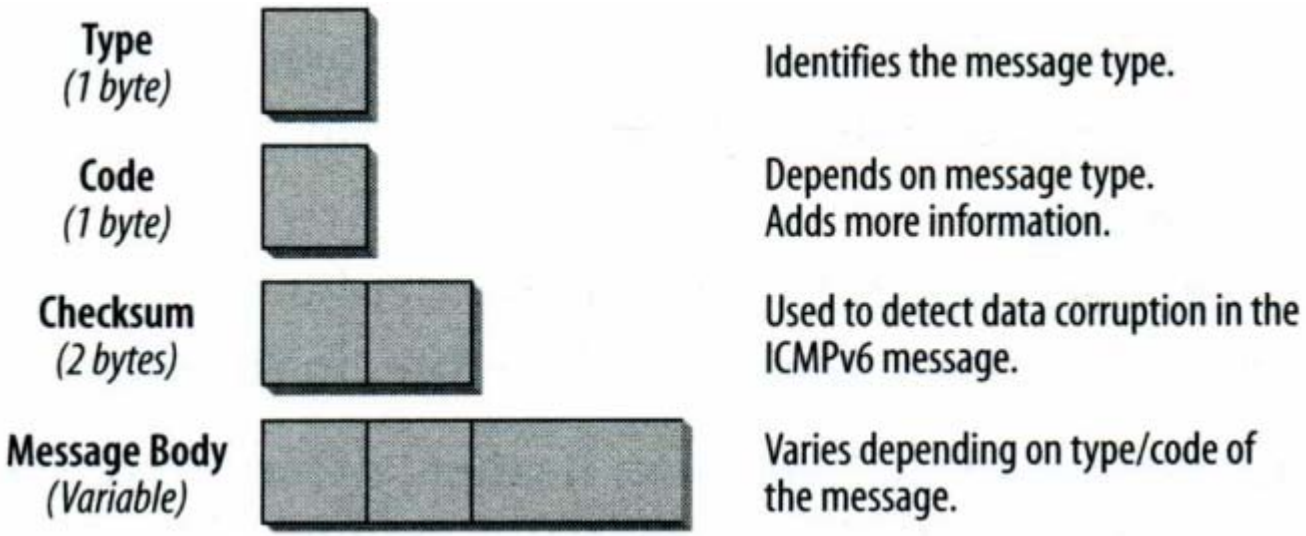
- ICMPv6 is much more powerful than ICMPv4, & contains new functionality:
  - The IGMP (Internet Group Management Protocol) function:
    - In IPv4, it manages multicast group memberships in IPv4.
  - The ARP/RARP (Address Resolution Protocol/Reverse Address Resolution Protocol) function:
    - In IPv4, it is used to map layer 2 addresses (MAC address) to IP addresses (and vice versa).
  - Neighbor discovery (ND) function:
    - It uses ICMPv6 messages to fine the link-layer addresses of its neighbors on the same link.
  - ICMPv6 also supports Mobile IPv6.
- ICMPv6 is defined in RFC 2463 (obsoletes RFC 1885). Neighbor discovery is defined in RFC 2461 (obsoletes RFC 1970).
- The two versions of ICMPv4 & ICMPv6 are not compatible.

General Message Format (1)

- General ICMPv6 header format:
  - Type field (1-byte), code field (1-byte), checksum field (2-byte), & Message Body (Variable Size).
- There are 2 classes of ICMPv6 messages:
  - ICMPv6 error messages.
    - The type field (8-bit) f error messages are 0XXXXXXX. Hence, ICMPv6 error message types are from 0 to 127.
  - ICMPv6 informational messages.
    - The type field (8-bit) of informational messages are 1XXXXXXX. Hence, ICMPv6 informational message types are from 128 to 255.
- All IPv6 extension headers are allocated in front of ICMPv6 messages.
- The following ICMPv6 messages are described in RFC 2463:
  - ICMPv6 error messages:
    - Message type 1: Destination Unreachable.
    - Message type 2: Packet Too Big.

General Message Format (2)

- General ICMPv6 header format.



General Message Format (2)

- Message type 3: Time Exceeded.
- Message type 4: Parameter Problem.

- ICMPv6 informational messages:
  - Message type 128: Echo Request.
  - Message type 129: Echo Reply.

General Message Format (3)

- General ICMPv6 header format:
  - Type (1-byte).
  - Code (1-byte).
    - To provide more explicit information.
  - Checksum (2-byte).
    - It is used to detect/protect the ICMPv6 header & parts of the IPv6 header (source & destination addresses).
  - Message body (variable size).
    - The message body will hold different data.
      - Ex) the error message will contain as much as possible of the error information for troubleshooting.
    - The total size of the ICMPv6 packet should not exceed the minimum IPv6 MTU (1280-byte).

General Message Format (5)

- ICMPv6 error messages: different types & codes.

Message number	Message type	Code field
1	Destination Unreachable	0 = no route to destination 1 = communication with destination administratively prohibited 2 = beyond scope of Source address 3 = address unreachable 4 = port unreachable 5 = Source address failed ingress/egress policy 6 = reject route to destination
2	Packet Too Big	Code field set to 0 by the sender and ignored by the receiver
3	Time Exceeded	0 = hop limit exceeded in transit 1 = fragment reassembly time exceeded

General Message Format (6)

- ICMPv6 error message & code type.

Message number	Message type	Code field
4	Parameter Problem	0 = erroneous header field encountered 1 = unrecognized next header type encountered 2 = unrecognized IPv6 option encountered The pointer field identifies the octet offset within the invoking packet where the error was detected. The pointer points beyond the end of the ICMPv6 packet if the field in error is beyond what can fit in the maximum size of an ICMPv6 error message.
100 and 101	Private experimentation	RFC 4443
127	Reserved for expansion of ICMPv6 error messages	RFC 4443

General Message Format (7)

- ICMPv6 informational messages: different types & codes.

Message number	Message type	Description
128	Echo Request	RFC 4443. Used for the ping command.
129	Echo Reply	
130	Multicast Listener Query	RFC 2710. Used for multicast goup management.
131	Multicast Listener Report	
132	Multicast Listener Done	
133	Router Solicitation	RFC 2461. Used for neighbor discovery and autoconfiguration.
134	Router Advertisement	
135	Neighbor Solicitation	
136	Neighbor Advertisement	
137	Redirect Message	
138	Router Renumbering	RFC 2894 Codes: 0 = Router renumbering command 1 = Router renumbering result 255 = Sequence number reset

General Message Format (8)

- ICMPv6 informational messages.

Message number	Message type	Description
139	ICMP Node Information Query	*draft-ietf-ipngwg-icmp-name-lookups-15.txt*
140	ICMP Node Information Response	
141	Inverse ND Solicitation	RFC 3122
142	Inverse ND Adv Message	RFC 3122
143	Version 2 Multicast Listener Report	RFC 3810
144	ICMP Home Agent Address Discovery Request Message	RFC 3775 ICMPv6 Messages for Mobile IPv6
145	ICMP Home Agent Address Discovery Reply Message	
146	ICMP Mobile Prefix Solicitation Message	
147	ICMP Mobile Prefix Advertise-ment Message	

General Message Format (9)

- ICMPv6 informational messages.

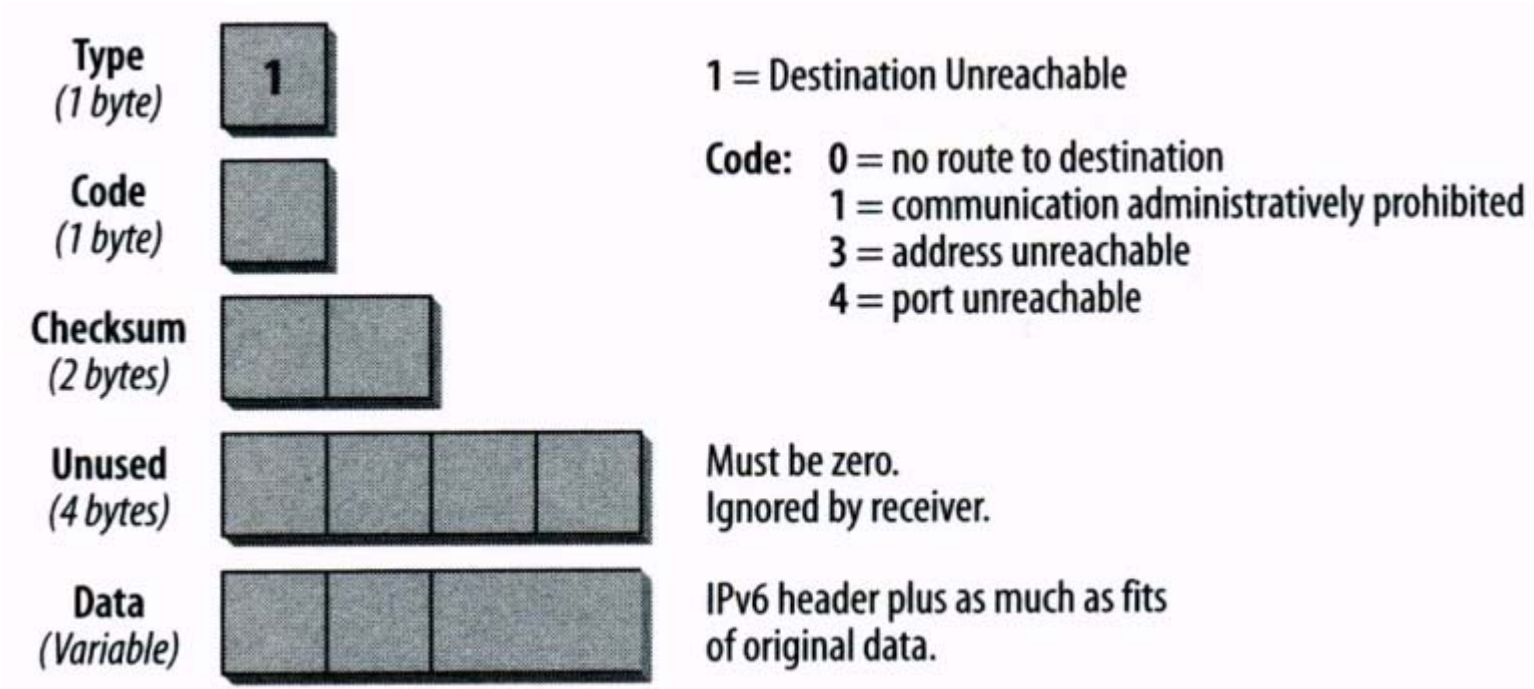
Message number	Message type	Description
148	Certification Path Solicitation Message	RFC 3971 ICMPv6 Messages for SEcure Neighbor Discovery
149	Certification Path Advertisement Message	
151	Multicast Router Advertisement	RFC 4286
152	Multicast Router Solicitation	
153	Multicast Router Termination	
200	Private experimentation	RFC 4443
201		
255	Reserved for expansion of ICMPv6 informational messages	RFC 4443

ICMPv6 Error Messages (1)

- Message type 1: Destination Unreachable:
  - If an IP datagram cannot be delivered, a Destination Unreachable message is generated & sends back to the source node (host).
  - The type field value:1.
  - The data portion of the ICMP message contains parts of the original content of the IP datagram.
  - If the destination is unreachable due to congestion, no ICMP message is generated.

ICMPv6 Error Messages (2)

- Format of the Destination Unreachable message.



ICMPv6 Error Messages (3)

- Code value of the Destination Unreachable message (type 1).

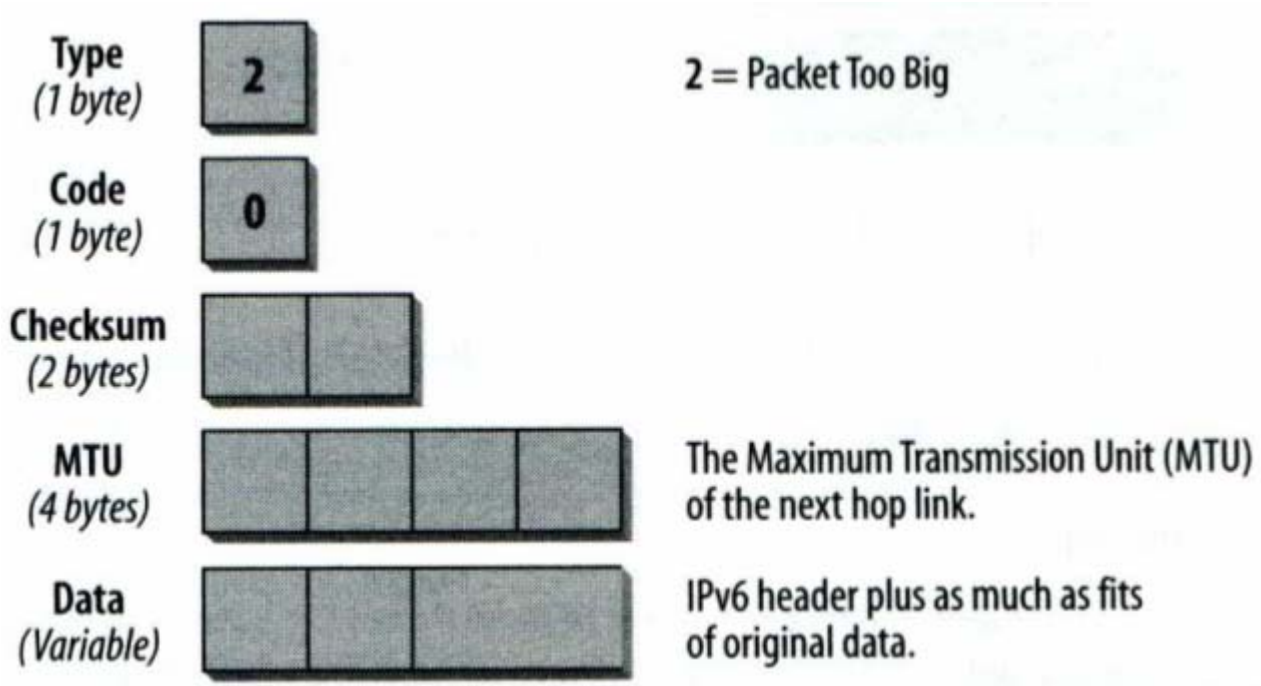
Code	Description
0	"No route to destination." This code is used if a router cannot forward a packet because it does not have a route in its table for a destination network. This can happen only if the router does not have an entry for a default route.
1	"Communication with destination administratively prohibited." This type of message can, for example, be sent by a firewall that cannot forward a packet to a host inside the firewall because of a packet filter. It might also be sent if a node is configured not to accept unauthenticated Echo Requests.
2	"Beyond scope of Source address." This code is used if the Destination address is beyond the scope of the Source address, e.g., if a packet has a link-local Source address and a global Destination address.
3	"Address unreachable." This code is used if a Destination address cannot be resolved into a corresponding network address or if there is a data-link layer problem preventing the node from reaching the destination network.
4	"Port unreachable." This code is used if the transport protocol (e.g., UDP) has no listener and there is no other means to inform the sender. For example, if a Domain Name System (DNS) query is sent to a host and the DNS server is not running, this type of message is generated.
5	"Source address failed ingress/egress policy." This code is used if a packet with this Source address is not allowed due to ingress or egress filtering policies.
6	"Reject route to destination." This code is used if the route to the destination is a reject route.

ICMPv6 Error Messages (4)

- Message type 2: Packet Too Big:
  - If a router cannot forward a packet because it is larger than the MTU of the outgoing link, it will generate a Packet Too Big message & send back to the source host.
  - The type field value: 2.
  - The unused code field value: 0.
  - The MTU field: contains the MTU size of the next hop link.
    - The source host can determine the MTU that it should use for further communication.
  - The data portion of the ICMP message contains parts of the original content of the IP datagram.

ICMPv6 Error Messages (5)

- Format of the Packet Too Big message.

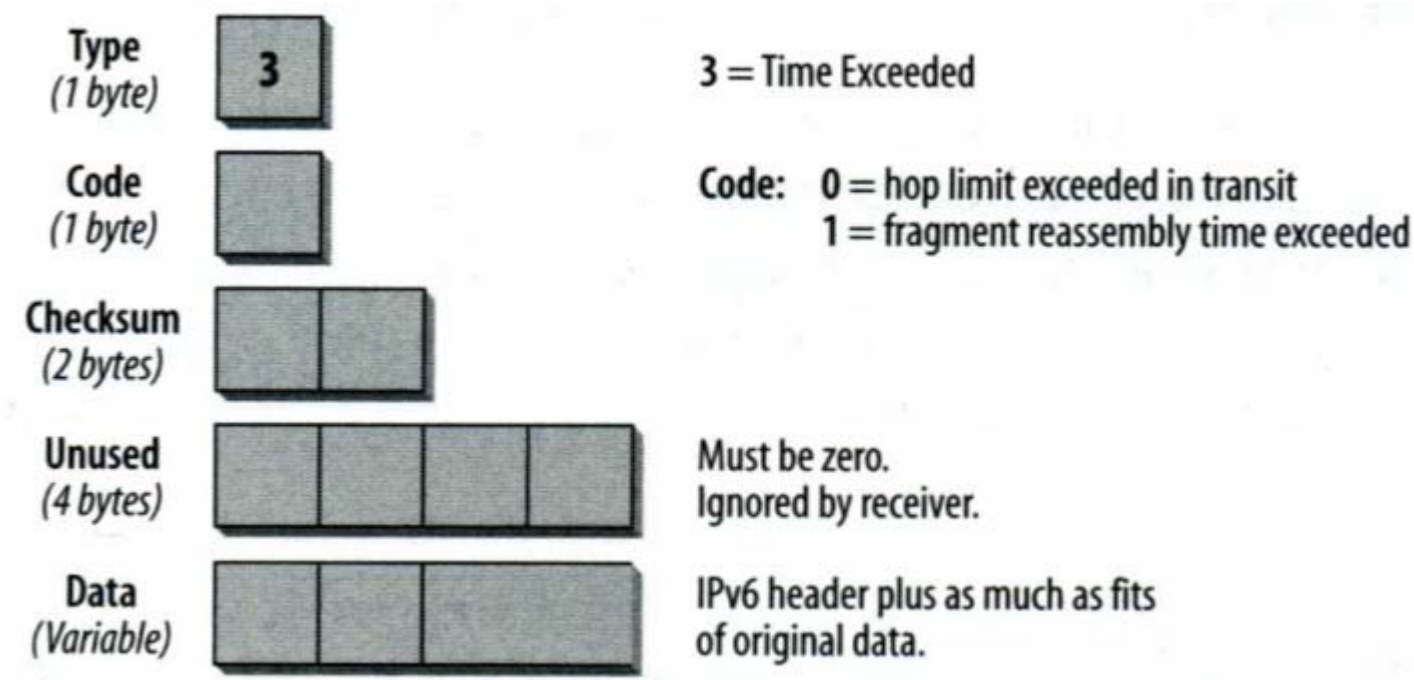


ICMPv6 Error Messages (6)

- Message type 3: Time Exceeded:
  - If a router receives a packet with a hop limit of one, & then decrements the limit to zero. It will discard the packet, generate a Time Exceeded message, & send this message back to the source host.
  - The type field value:3.
  - The code field value:
    - 0 : means that the hop limit exceeded during transmission.
    - 1 : means that the fragment reassembly time is exceeded.
  - The data portion of the ICMP message contains parts of the original content of the IP datagram.

ICMPv6 Error Messages (7)

- Format of the Time Exceeded message.



ICMPv6 Error Messages (8)

- Code values for Time Exceeded message (type 3).

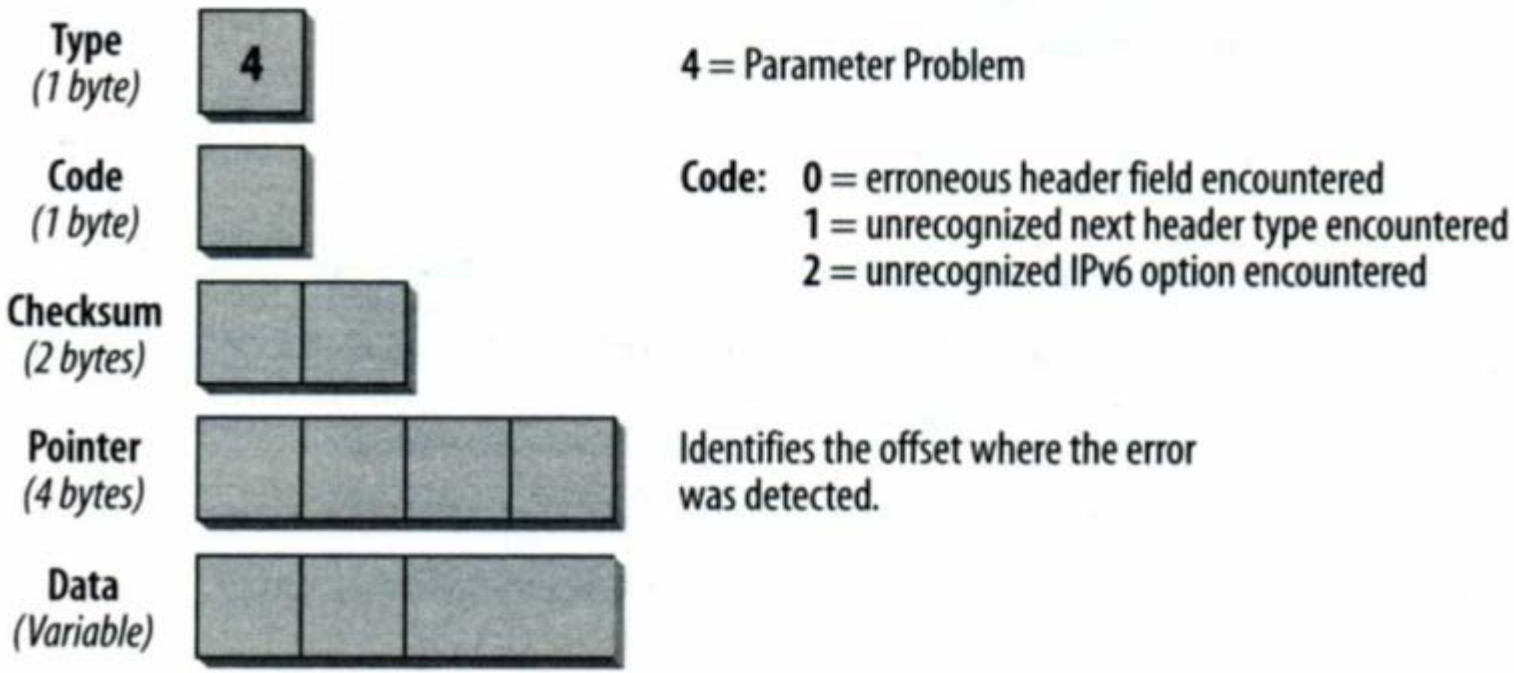
Code	Description
0	"Hop limit exceeded in transit." Possible causes: the initial hop limit value is too low; there are routing loops; or use of the *traceroute* utility.
1	"Fragment reassembly time exceeded." If a fragmented packet is sent by using a fragment header (refer to Chapter 2 for more details) and the receiving host cannot reassemble all packets within a certain time, it notifies the sender by issuing this ICMP message.

ICMPv6 Error Messages (9)

- Message type 4: Parameter Problem:
  - If an IPv6 node (router / host) cannot recognize any part of the IPv6 header or extension headers, it must discard the packet & sends an ICMP Parameter Problem message back to the source host.
  - The type field value:4.
  - The code field value:
    - 0 : means erroneous header field (via checksum).
    - 1 : means unrecognized next header type encountered.
    - 2 : means unrecognized IPv6 option.
  - The Pointer field: points out which byte was detected error in the original packet.
  - The data portion of the ICMP message contains parts of the original content of the IP datagram.

ICMPv6 Error Messages (10)

- Format of the Parameter Problem message:



ICMPv6 Error Messages (11)

- Code values for Parameter Problem (type 4):

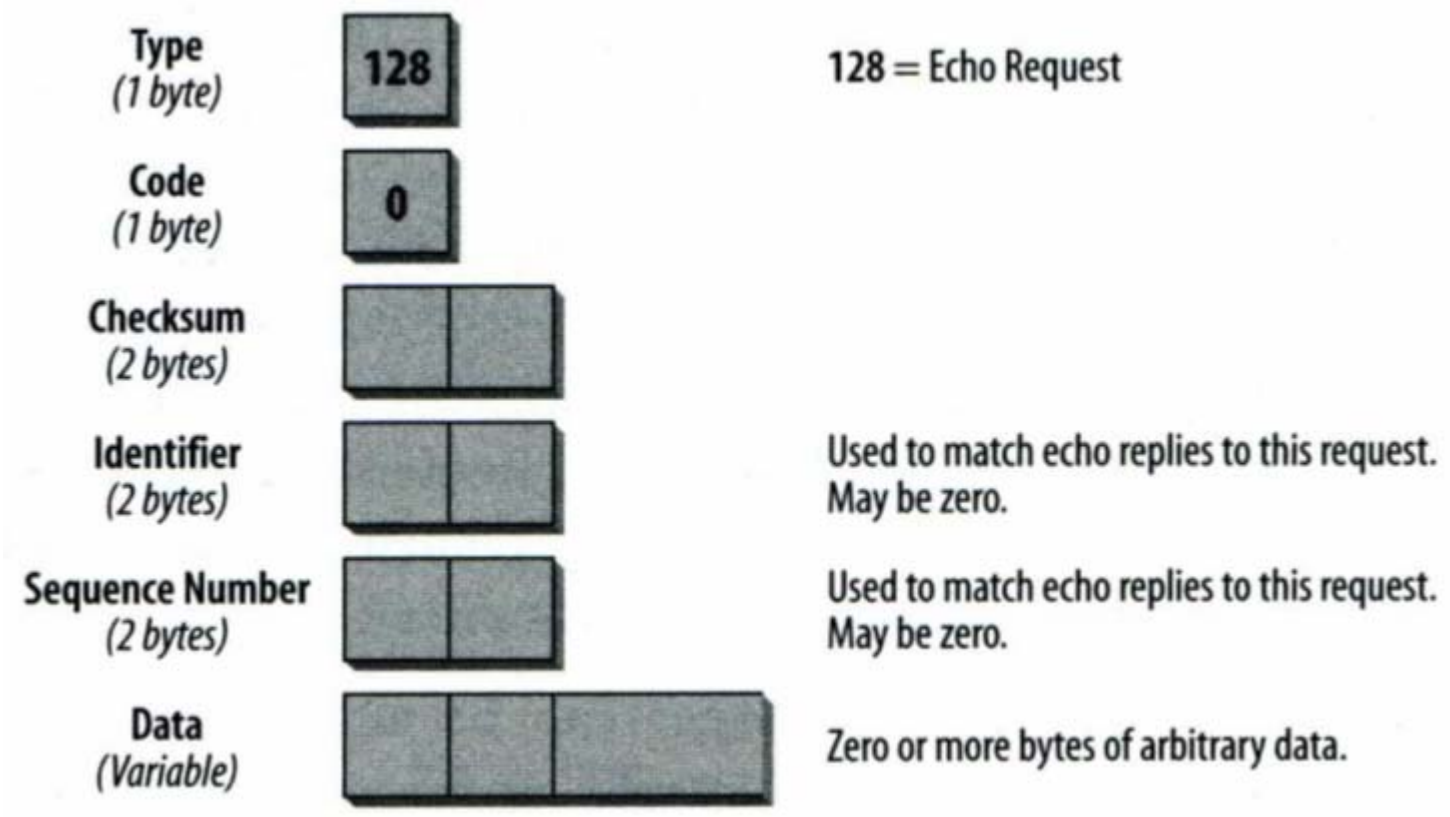
Code	Description
0	Erroneous header field encountered
1	Unrecognized next header type encountered
2	Unrecognized IPv6 option encountered

ICMPv6 Informational Messages (1)

- In RFC 2463, 2 types of informational messages are defined:
  - The Echo Request & the Echo Reply messages.
- Other ICMP informational messages are used for:
  - Path MTU (PMTU) Discovery (RFC 1981) & neighbor discovery (ND) (RFC 2461).
- The Echo Request & Echo Reply messages are often used in IPv4:
  - The Packet INternet Groper (ping).
    - The source host issues an Echo Request message to the destination. The destination host, if available, responds with an Echo Reply message.
- Message type 128: Echo Request message:
  - The type field value: 128.
  - The unused code field value: 0.
  - The Identifier & Sequence Number fields are used to match the requests and the replies.
    - The reply must always contain the same numbers as the request.

ICMPv6 Informational Messages (2)

- Format for the Echo Request message.

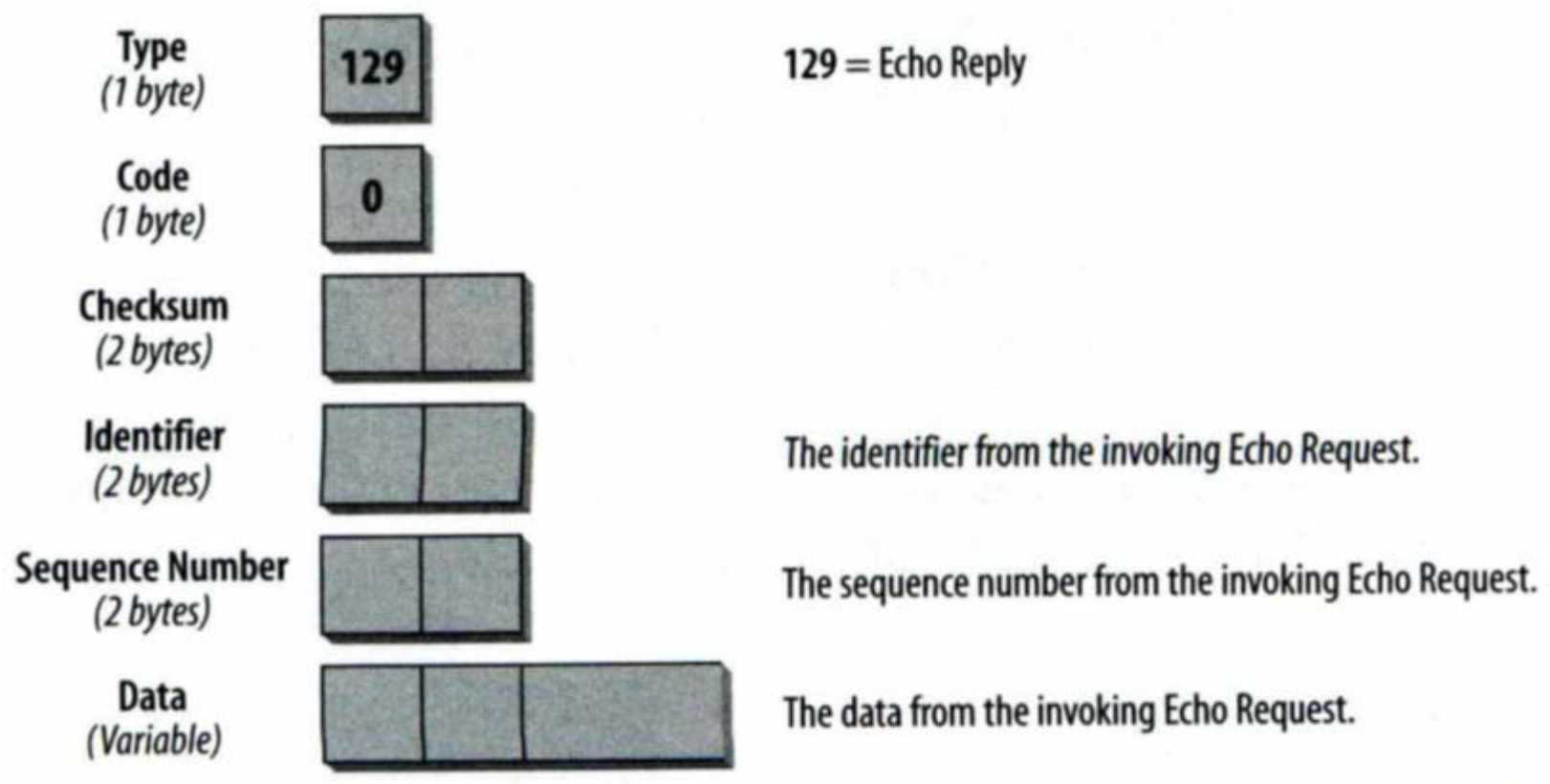


ICMPv6 Informational Messages (3)

- Message type 129: Echo Reply message:
  - The type field value: 129.
  - The unused code field value: 0.
  - The Identifier & Sequence Number fields are used to match the requests and the replies.
    - The reply must always contain the same numbers as the request.
- ICMPv6 Echo Request & Reply messages can be authenticated.
  - Using an IPv6 authentication header.
  - This means that a node can be configured to ignore non-authenticated ICMPv6 pings.
  - To provide protection against different sorts of ICMPv6 attacks.

ICMPv6 Informational Messages (4)

- Format for the Echo Reply message.

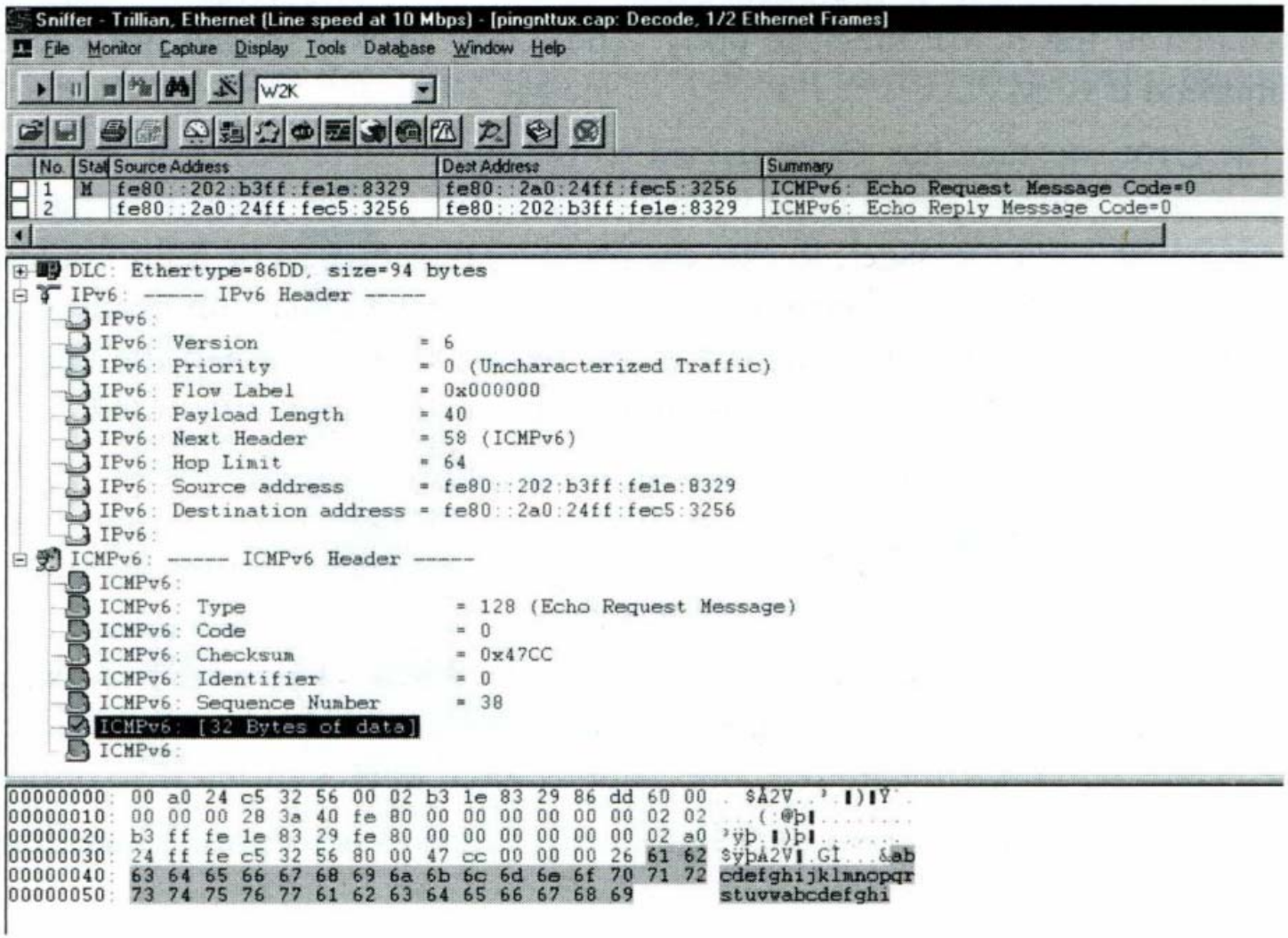


Processing Rules

- There are several rules that govern processing of ICMP packets (RFC 2463).
  - If a node receives a unknown type of ICMPv6 error message, it must pass it to the upper layer.
  - If a node receives a unknown type of ICMPv6 informational message, it must be silently discarded.
  - The content of packet, which caused the ICMP error message, will be included in the ICMP message body.
    - The ICMP packet should not exceed the minimum IPv6 MTU.
- Every IPv6 node must implement a rate-limiting function.
  - It limits the sending rate of ICMPv6 messages.
  - The configurable limit can be based on either timer or bandwidth.
  - Thus, it can protect against denial-of-service attacks.

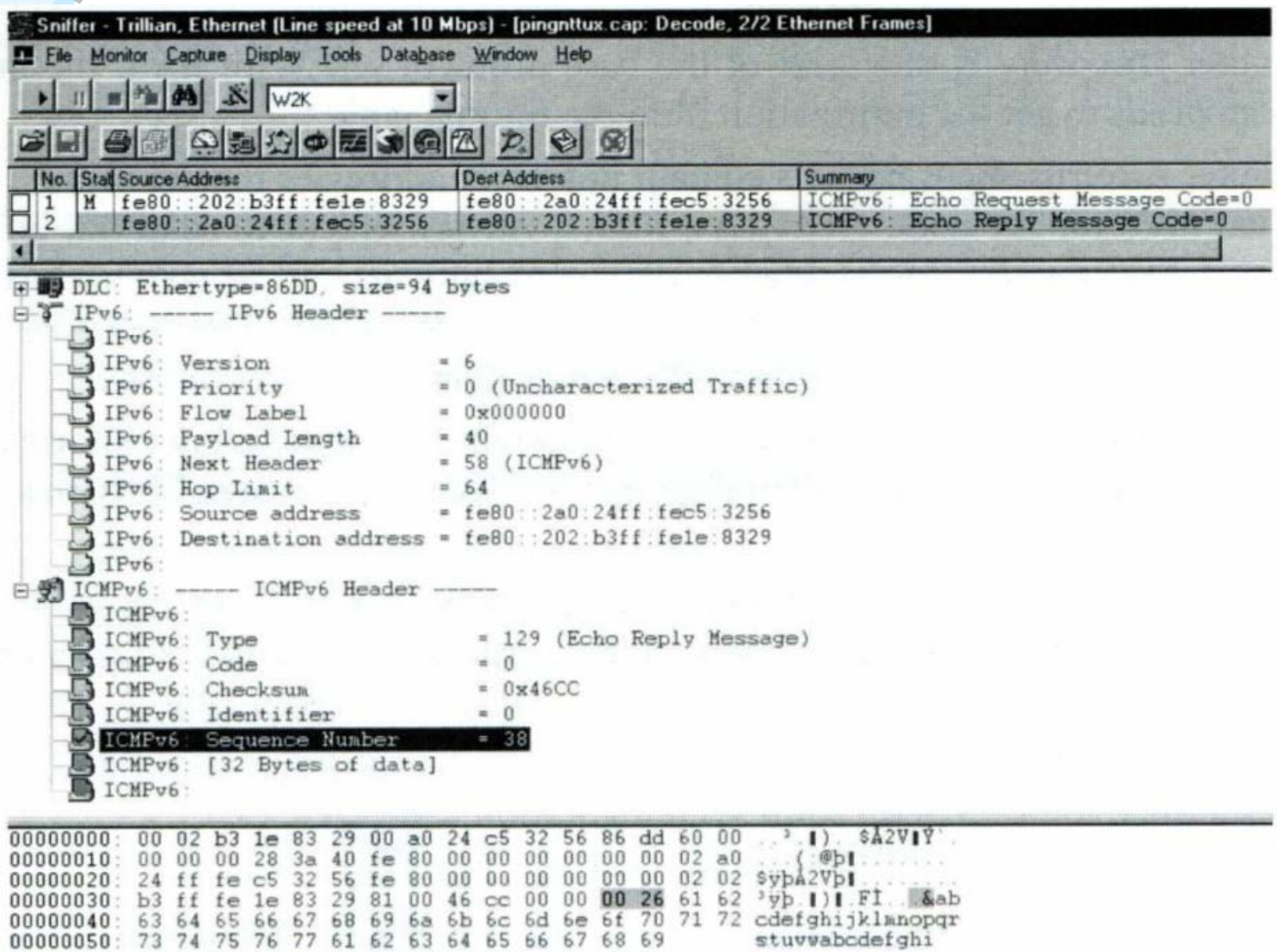
The ICMPv6 Header in a Trace File (1)

- A Windows 2000 host issued a ping command to a Linux host.
  - Echo Request in a trace file.



The ICMPv6 Header in a Trace File (2)

- Echo Reply in a trace file.



Neighbor Discovery (1)

- Neighbor discovery (ND) is specified in RFC 2461.
- It combines:
  - ARP (Address Resolution Protocol) in IPv4.
  - ICMP router discovery.
  - Redirect.
  - DAD (Duplicate Address Detection): duplicate IP address detection has been implemented.
  - Neighbor discovery protocol: a neighbor unreachability detection mechanism.
    - In IPv4, we have no means to detect whether or not a neighbor is reachable.
- The purposes of using IPv6 neighbor discovery:
  - To determine layer 2 (MAC) addresses of other nodes on the same link.
  - To find neighboring routers, which can forward their packets.
  - To keep track of which neighbors are reachable.

Neighbor Discovery (2)

- The improvements from the related IPv4 protocols to IPv6 ND:
  - In IPv6 router discovery is part of the base protocol set.
    - In IPv4, the mechanism needs to get the information from the routing table.
  - Router advertisement (RA) packets contain router link-layer (MAC) addresses.
    - It does not need to send the ARP request (it is required in IPv4) to get the link-layer address of the router interface.
    - It is the same as for ICMPv6 redirect messages.
      - They contain the link-layer address of the new next-hop router.
  - RA packets contain the prefix of a link (the subnet information).
  - ND provides mechanisms to get new prefixes and addresses.
  - RAs enable stateless address auto-configuration.
  - Routers can advertise an MTU of the used link.

Neighbor Discovery (3)

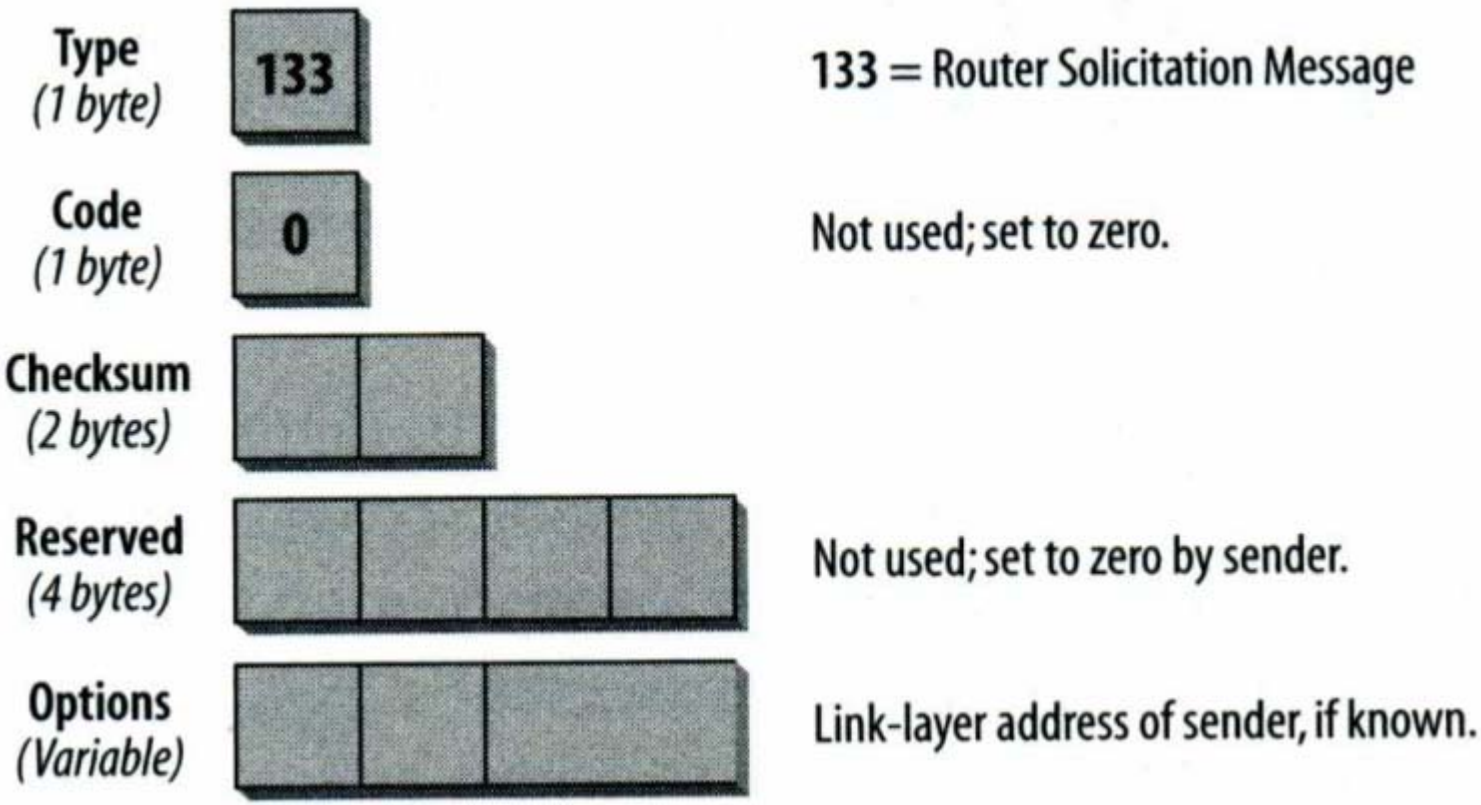
- Multiple prefixes can be assigned to one link.
  - By default, hosts can learn all prefixes from the router.
- Neighbor unreachability detection is part of the base protocol.
  - It solves the issues with outdated ARP caches.
  - The neighbor unreachability detection can detect failed routers & switches to live ones.
- RAs and ICMP redirects use link-local addresses to identify routers.
- ND messages have a hop limit value of 255. A request with a lower hop limit (<255) is not answered.
  - Because it is not a neighbor (<255), it ever went through a router.
- ND protocol is used to detect duplicate IP addresses (DAD) on a link.
- Standard IP authentication & security mechanisms can be applied to ND.

Neighbor Discovery (4)

- The ND protocol consists of five ICMP messages:
  - A pair of Router Solicitation / Router Advertisement messages.
  - A pair of Neighbor Solicitation / Neighbor Advertisement messages.
  - An ICMP redirect message.
- Router Solicitation (RS) & Router Advertisement (RA):
  - Routers send out RA messages in regular intervals.
  - Hosts can also request RAs by issuing a RS message.
  - In the RS message:
    - The destination address is FF02::2 (the all-routers multicast address).
    - The hop limit is set to 255.
    - The ICMP Type field is set to 133.
    - The Code field is unused & set to 0.
    - The following 2 bytes are used for the Checksum.

Neighbor Discovery (5)

- Router Solicitation (RS) message.



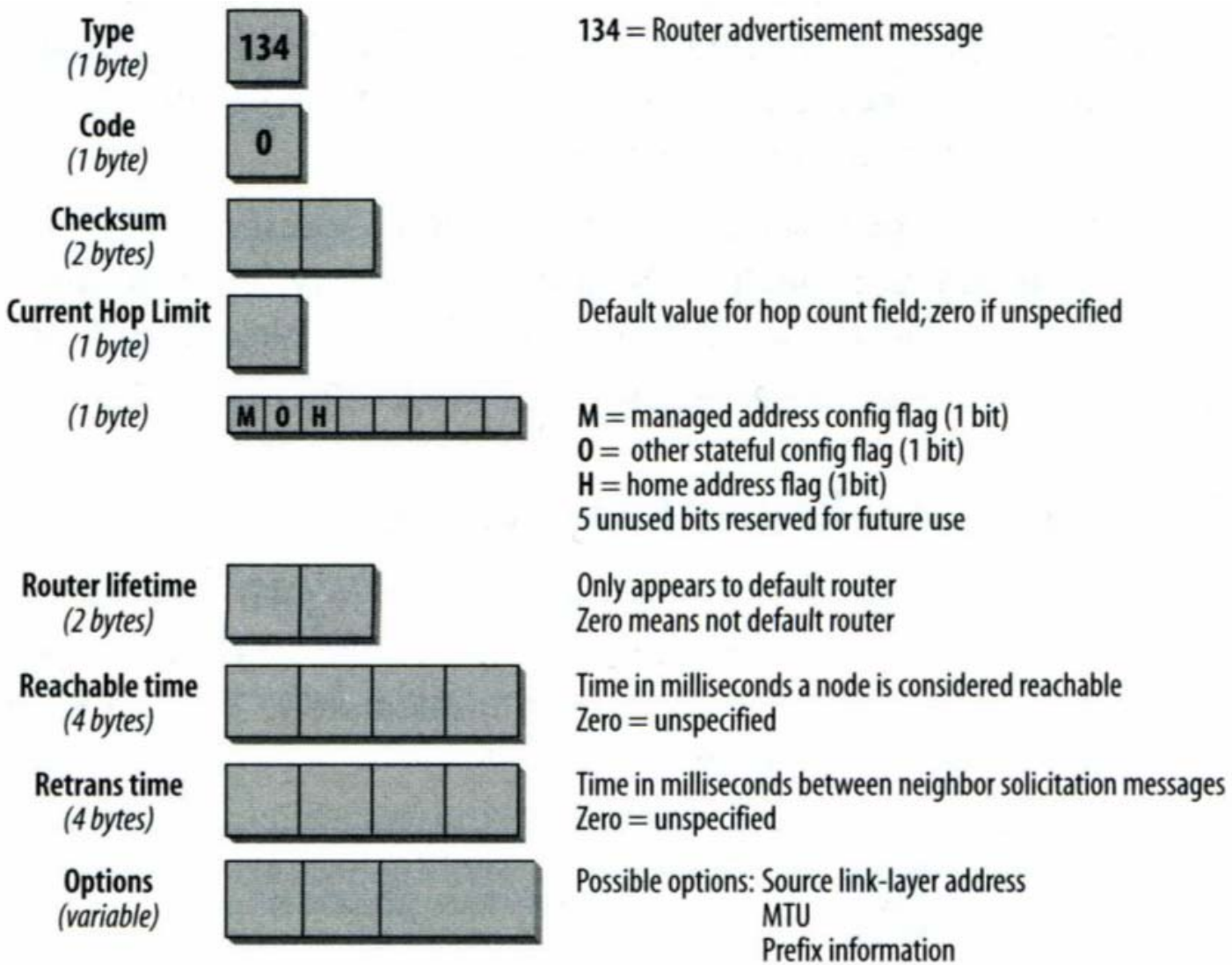
Neighbor Discovery (6)

- The next 4 bytes are unused and reserved for future use.
- The option is for source MAC address, but if the source IP address is :: (all zero), this field is unused.

- Routers that receive this Solicitation message reply with a RA message.
- In the RA message:
  - ICMP Type field: is set to 134.
  - Code field: is unused & set to 0.
  - The Current Hop Limit field: can be set a default hop limit.
    - The hop limit is set to 255.
    - A value of 0 is unused.
  - Destination address:
    - If the RA is periodic: destination address will be the all-nodes multicast address FF02::1.
    - If the RA was sent in reply to a RS message: destination address will be the interface, which sent the solicitation message.

Neighbor Discovery (7)

- Router Advertisement (RA) message.



Neighbor Discovery (8)

- The M flag (1-bit field):
  - 0: stateless address configuration (auto-configuration).
  - 1: stateful address configuration (DHCPv6).
- The O flag (1-bit field):
  - 0: nodes do not use stateful configuration for non-IP address information.
  - 1: nodes use stateful configuration for non-IP address information.
- The remaining 6 bits of this byte: are reserved & must be 0.
- Router Lifetime field (sec):
  - Nodes on a link must have a default router.
  - If the router is not a default router, the value is 0.
  - If the router is a default router, the value is the router lifetime.
  - It counts in seconds, & the maximum value is 18.2 hours.

Neighbor Discovery (9)

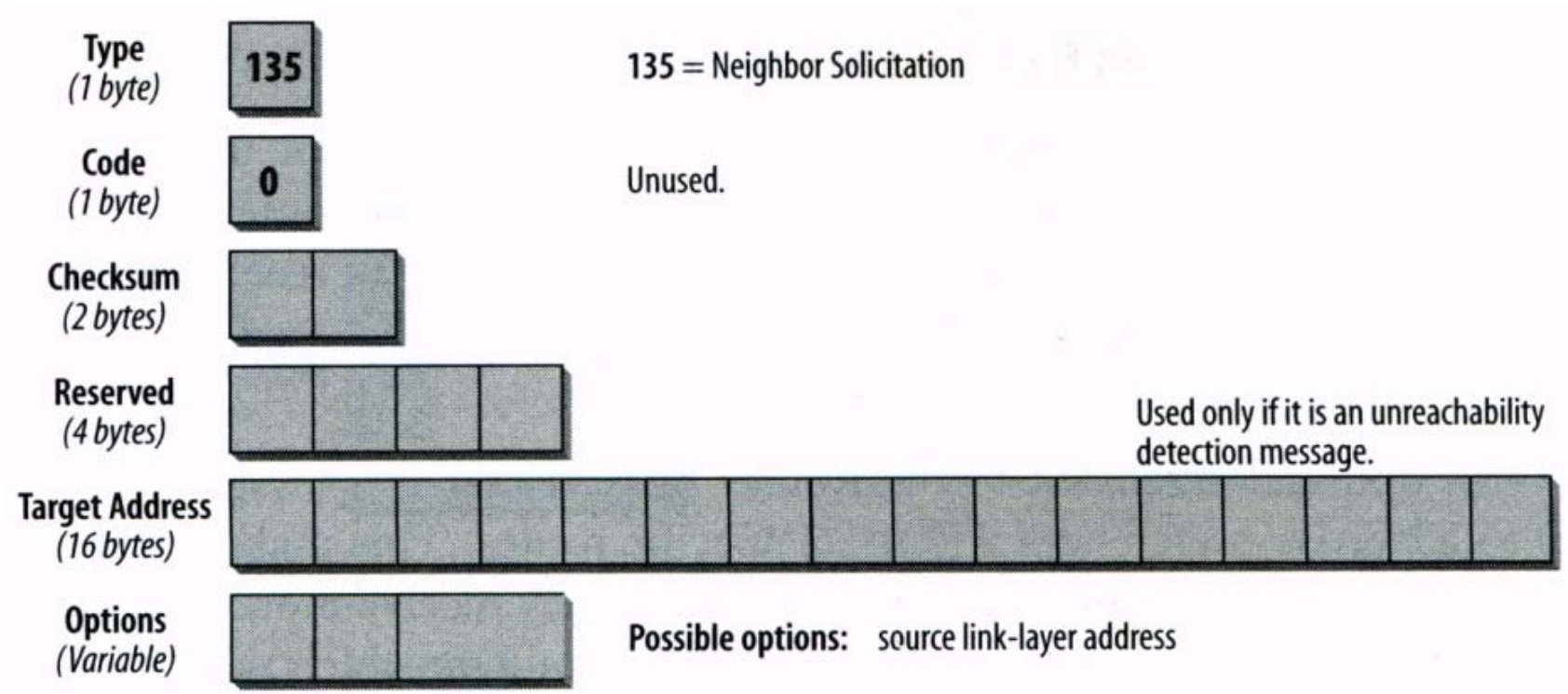
- The Reachable Time field (millisec):
  - It is the time in which a host assumes that neighbors are reachable after having received a reachability confirmation.
  - If the value is 0, which means unspecified.
  - The neighbor unreachability detection algorithm uses this field.
- The Retrans Timer field (millisec):
  - It is used by the address resolution & neighbor unreachability detection mechanisms.
  - It states the time between retransmitted NS messages.
- The Options field (currently 3 possible values):
  - 1: Source link-layer address (MAC address):
  - 3: MTU size:
  - 5: Prefix information: is for auto-configuration.

Neighbor Discovery (10)

- Neighbor Solicitation (NS) & Neighbor Advertisement (NA):
  - This pair of messages have 3 functions:
    - The link-layer (MAC) address resolution:
      - Like ARP in IPv4.
      - The destination address is a multicast address.
    - The neighbor unreachability detection mechanism:
      - The destination address is a unicast address.
    - DAD (duplicate IP address detection).
  - Format of the NS message:
    - The source address: is usual the interface address.
      - In DAD, the source address is all-zeros address.
    - The hop limit: is 255.
    - The Type field: is 135.
    - The Code field: is unused & set to 0.
    - The 2 checksum bytes.

Neighbor Discovery (11)

- Format of the Neighbor Solicitation (NS) message.



Neighbor Discovery (12)

- 4 unused bytes: are reserved & must be set to 0.
- The target address field:
  - Is used for unreachability detection & DAD.
  - It must not be a multicast address.
- The Options field:
  - It usually contains the link-layer (MAC) source address.
    - In DAD, the options field is set to 0.

- NA messages are sent as a reply to NS messages or to propagate new information quickly.
- If the NA message is the answer to a NS message:
  - A solicited advertisement, the destination IP address is the source address of the interface, which sent the solicitation message.
- If the NA message is an unsolicited message:
  - If the message is the reply to a DAD message, the destination IP address is the all-nodes multicast address of FF02::1.

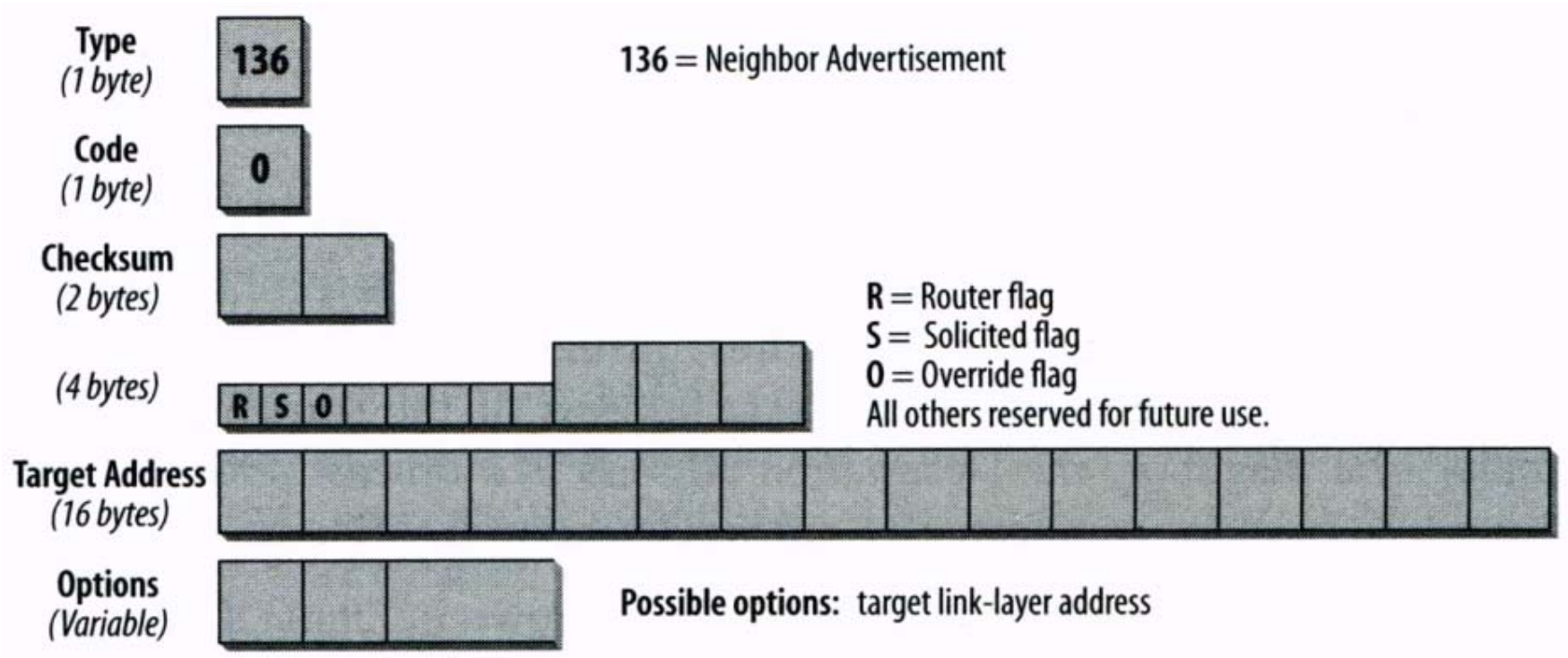
Neighbor Discovery (13)

- The same is true for all unsolicited periodic advertisements.

- Format of the NA message:
  - The Type field: is 136.
  - The Code field: is unused & set to 0.
  - When the Router flag (R bit) is set: the sender is a router.
  - When the Solicited flag (S bit) is set: the message is sent to reply a NS message.
  - The Override flag (O bit) indicates:
    - The information in the NA message should override existing neighbor cache entries.
    - To update any cached link-layer (MAC) addresses.
    - If the O bit is not set, the NA will not update a cached link-layer (MAC) address, but it will update an existing neighbor cache entry for which does not contain the link-layer address.

Neighbor Discovery (14)

- Format of the Neighbor Advertisement (NA) message.



Neighbor Discovery (15)

- Identification of ND message.

Source address	Destination address	Message type
All-zero (::0)	All-routers multicast Solicited node multicast	Stateless autoconfiguration DAD
Unicast	Solicited node multicast Unicast	Resolve link-layer address Unreachability detection

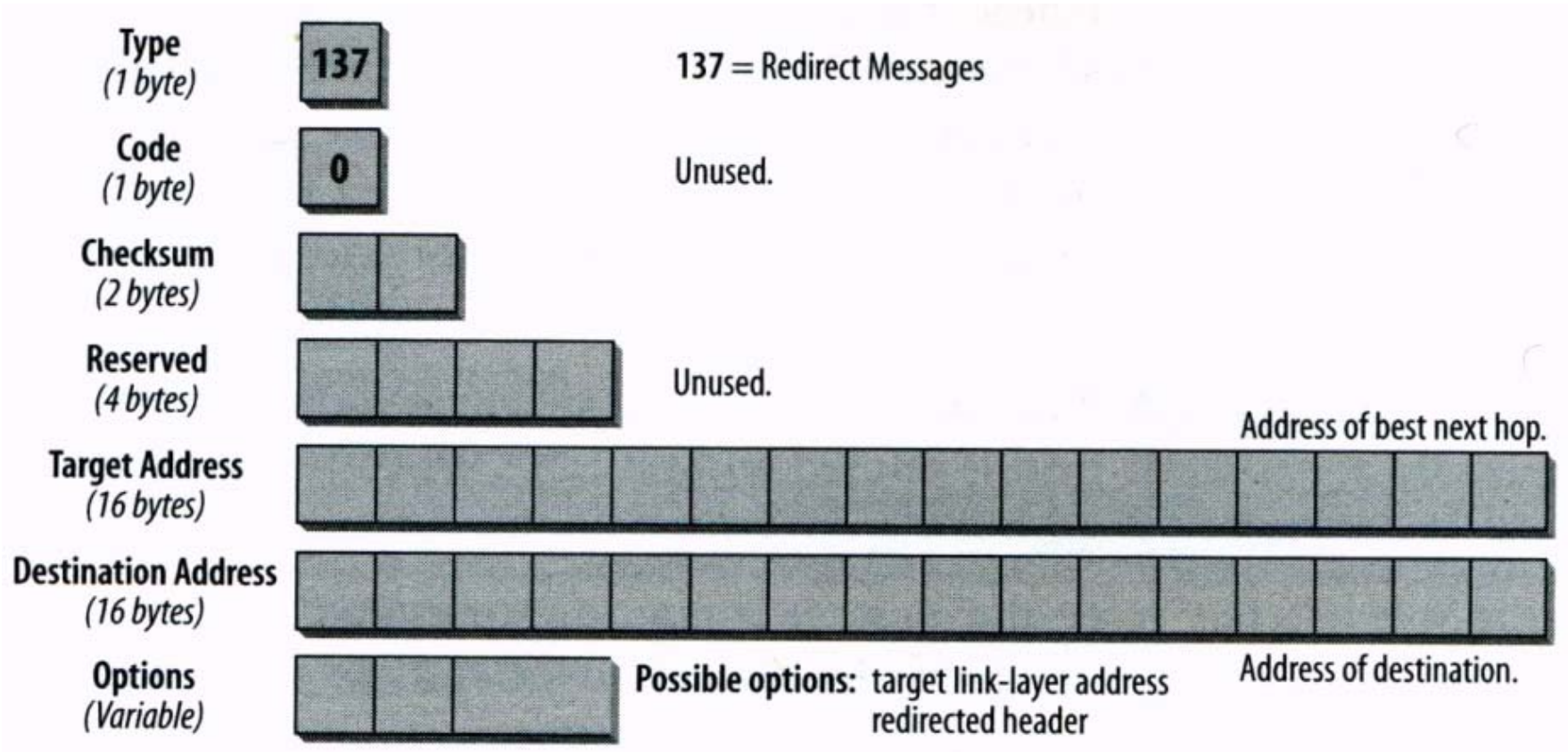
Neighbor Discovery (16)

- The Target Address field:
  - In solicited advertisements: it contains the IP address of the interface, which sent the solicitation message.
  - In unsolicited advertisements: it contains the IP address of the interface, whose link-layer (MAC) address has changed.
- The Options field: is the target link-layer (MAC) address.

- The ICMP Redirect Message:
  - If there is another better path to a destination, routers issue ICMP Redirect messages to inform a node about the first-hop router on the path.
  - A Redirect message can also inform a node that the destination & the node are neighbors on the same link.
  - The format of the ICMPv6 Redirect message:
    - The Type field: is 137.
    - The Code field: is unused & set to 0.

Neighbor Discovery (17)

- Format of the ICMP Redirect message.



Neighbor Discovery (18)

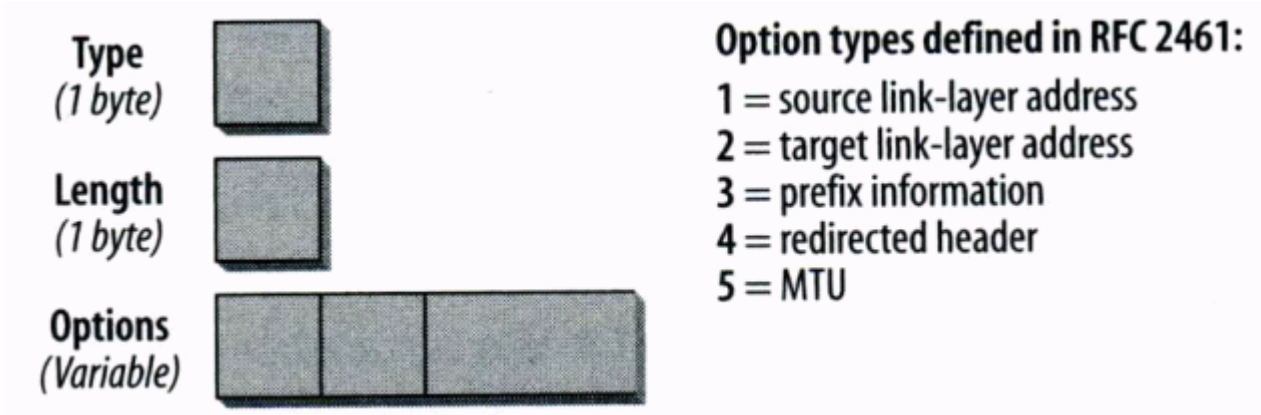
- The source address: must be the link-local IP address of the interface (router) from which the message is sent.
- The destination address: is the node IP address.
- The hop limit: is 255.
- The Target Address field: is the link-local (IP) address of the interface, which is the first next-hop router of the other path.
- The Destination Address field: is the final destination IP address.
  - If the target address is the same as the destination address, the destination is a neighbor on the same link.
- Option field: contains the link-layer (MAC) address of the first next-hop router.
  - In IPv4, the source node needs to send a ARP request to determine the link-layer (MAC) address of the next-hop router.

Neighbor Discovery (19)

- ND Options:
  - 3 parts in the Option field (TLV): type, length, & value (option field).
  - There are 5 type of Options in the ICMPv6 ND messages.
    - The different Type field values represent different value (option) contents (RFC2461).
  - The format of the ND message:
    - Type 1: source link-layer (MAC) address.
    - Type 2: target link-layer (MAC) address.
    - Type 3: prefix information.
    - Type 4: redirected header.
    - Type 5: MTU.
    - Etc..
    - The Length field: is the total length of the Type, Length, & Value (option).
      - Value 0: means this field is invalid, & discards the packet.

Neighbor Discovery (20)

- Format of the Optional field.



Neighbor Discovery (21)

- Overview of ND Options.

Option type	Used in
Type 1 Source link-layer address	Neighbor solicitation Router solicitation Router advertisement IND solicitation/advertisement
Type 2 Target link layer address	Neighbor advertisement Redirect IND solicitation/advertisement
Type 3 Prefix	Router advertisement
Type 4 Redirected header	Redirect
Type 5 MTU	Router advertisement IND solicitation/advertisement
Type 7 Advertisement interval	Router advertisement (defined in Mobile IPv6 specification)
Type 8 Home Agent information	Router advertisement (defined in Mobile IPv6 specification)
Type 9 Source address list	IND Solicitation
Type 10 Target address list	IND Advertisement

Neighbor Discovery (22)

- In figure 4-16: the RA in a trace file.
  - The option 1 is for source link-layer address.
    - Type field is 1.
    - Value field (option) is link-layer (MAC) address.
  - The option 2 is for prefix information.
    - Type field is 3.
    - Value field (option) includes:
      - Prefix length (bits).
      - L bit: is the on-link flag for the prefix.
      - A-bit: is the autonomous address configuration flag for auto-configuration.
      - R bit: router address bit.
      - Valid Lifetime field: specifies how long this prefix is valid. A value of all F means infinity.
      - The Preferred Lifetime: specifies how long the address being configured with this prefix can remain in the preferred state. A value of all F means infinity.

Neighbor Discovery (23)

- The router advertisement in a trace file.

No.	Sta	Source Address	Dest Address	Summary
11		fe80::210:7bff:fe0b:75a0	ff02::1	ICMPv6: Router Advertisement Code=0

```
DLC: Ethertype=86DD, size=110 bytes
IPv6: Flow=0x000000
ICMPv6: ------ ICMPv6 Header ------
ICMPv6:
ICMPv6: Type                    = 134 (Router Advertisement)
ICMPv6: Code                    = 0
ICMPv6: Checksum                = 786B (correct)
ICMPv6: Current Hop Limit       = 64
ICMPv6: M/O/H/Reserved bits       = 00
ICMPv6:                 0... .... = administered protocol not used (address)
ICMPv6:                 .0.. .... = administered protocol not used (non-address)
ICMPv6:                 ..0. .... = Home Agent bit
ICMPv6:                 ...0 0000 = Reserved              = 0x00
ICMPv6: Router Lifetime         = 1800 s
ICMPv6: Reachable Time          = 0 ms (unspecified)
ICMPv6: Retrans Timer           = 0 ms (unspecified)
ICMPv6:
ICMPv6: Options follow
ICMPv6: Type                    = 1 (Source Link-Layer Address)
ICMPv6: Length                  = 1 (units of 8 octets)
ICMPv6: Link Layer Address      = Station Cisco 0B75A0
ICMPv6: Type                    = 3 (Prefix Information)
ICMPv6: Length                  = 4 (units of 8 octets)
ICMPv6: Prefix Length           = 64
ICMPv6: L/A/R bits                = C0
ICMPv6:                 1... .... = on-link determination
ICMPv6:                 .1.. .... = autonomous address configuration
ICMPv6:                 ..0. .... = Router Address bit
ICMPv6:                 ...0 0000 = Reserved                = 0x00
ICMPv6: Valid Lifetime          = 4294967295 s (infinity)
ICMPv6: Preferred Lifetime      = 4294967295 s (infinity)
ICMPv6: Reserved                = 0x00000000
ICMPv6: Prefix                  = caff:ca01:0:56::
ICMPv6:
```

Neighbor Discovery (24)

- Neighbor Cache and Destination Cache:
  - IPv6 nodes (routers & hosts) need to maintain different tables of information.
    - Among these tables, the neighbor cache & destination cache are particularly important.
  - Neighbor cache:
    - The neighbor cache maintains a list of neighbors:
      - Neighbor unicast IP address.
      - Neighbor link-layer (MAC) address.
      - A flag to indicate whether the neighbor is a router or host.
        - This can be compared to the ARP cache in an IPv4 node.
      - Whether there are packets queued to be sent to a destination.
      - Neighbor reachability.
      - The detection time of the next neighbor unreachability detection (if the neighbor is unreachable).

Neighbor Discovery (25)

- Neighbor cache entries on a CISCO router.

```
client#show ipv6 neighbors
IPv6 Address                              Age Link-layer Addr State Interface
FE80::202:B3FF:FE1E:8329                    5 0002.b31e.8329  DELAY Ethernet0/0
CAFF:CA01:0:56:202:B3FF:FE1E:8329           0 0002.b31e.8329  REACH Ethernet0/0
FE80::A00:20FF:FE20:ADC2                    4 0800.2020.adc2  STALE Ethernet0/0
CAFF:CA01:0:56:A00:20FF:FE20:ADC2           4 0800.2020.adc2  STALE Ethernet0/0
```

- A neighbor cache entry has 5 possible states (RFC 2461).
  - Incomplete, Reachable, Stale, Delay, & Probe.

Neighbor Discovery (26)

- States of neighbor cache entries.

State	Description
Incomplete	Address resolution is currently being performed and awaiting either a response or a timeout. Specifically, a Neighbor Solicitation has been sent to the solicited-node multicast address of the target, but the corresponding Neighbor Advertisement has not yet been received.
Reachable	This neighbor is currently reachable, which means positive confirmation was received within the last `ReachableTime` milliseconds that the neighbor was functioning properly.
Stale	More than `ReachableTime` milliseconds have elapsed since the last positive confirmation that the forward path was functioning properly was received. No action will take place regarding this neighbor until a packet is sent.
Delay	This neighbor's Reachable Time has expired, and a packet was sent within the last `DelayFirstProbeTime` seconds. If no confirmation is received within the `DelayFirstProbeTime` seconds, send a Neighbor Solicitation and change the neighbor state to Probe state. The use of `Delay` allows upper-layer protocols additional time to provide reachability confirmation. Without this extra time, possible redundant traffic would be generated.
Probe	A reachability confirmation is being actively attempted by sending Neighbor Solicitations every `RetransTimer` milliseconds until reachability is confirmed.

Neighbor Discovery (27)

- Destination cache:
  - The Destination cache maintains a list of destinations:
    - Recently, traffics ever sent to these destinations, which includes local & remote destinations.
    - In local destination:
      - The neighbor cache can be regarded as a subset of destination cache information.
    - In remote destination:
      - The entry contains the link-layer (MAC) address of the next-hop router.
  - The destination cache is updated when it receives the ICMP Redirect messages.
  - The destination cache can also contain additional information about MTU sizes & roundtrip timers.

Neighbor Discovery (28)

- About neighbor cache, destination cache,& NA message:
  - If the Override flag (O bit) of NA message is set:
    - The neighbor cache & destination cache information are updated.
    - The information in the Advertisement message will override existing neighbor cache entries and update any cached link-layer addresses.
  - If the O bit of NA message is not set:
    - The advertisement will not update a cached link-layer address, but will update an existing neighbor cache entry, which does not contain the link-layer address.

Auto-configuration (1)

- The auto-configuration capability of IPv6 saves network administrators a lot of work.
- It is designed for avoid manually configuring hosts before connecting to the network.
- IPv6 has stateless & stateful auto-configuration.
  - Stateful auto-configuration is to use DHCPv6 server (like in IPv4).
  - Stateless auto-configuration:
    - It is a new function in IPv6, hosts can auto-configure their IPv6 address without any manual configuration, & also does not need a DHCPv6 server to configure hosts.
    - Routers can advertise multiple prefixes, and hosts determine a prefix information from these advertisements.
    - If we change our ISP & the new ISP assigns a new IPv6 prefix, we can configure our routers to advertise this new prefix, & keeping the original SLA.

Auto-configuration (2)

- All hosts attached to those routers will get new global IPv6 address.
- If there is no router present, a host can generate its link-local address only with the prefix of FE80.

- An IPv6 address is leased to a node for a certain lifetime.
  - When the lifetime expires, the address must be released.
  - To make sure an address is unique on a link, a node runs the DAD process (RFC 2462), when it gets a new IPv6 address.
- An IPv6 address of a node can have different states:
  - Tentative address:
    - This is an address that has not yet been assigned.
    - The uniqueness of the address is being verified (in DAD processing).
  - Preferred address:
    - This is the address that has been assigned to an interface.
    - The address can be used without any restrictions.

Auto-configuration (3)

- Deprecated address:
  - The lifetime of a deprecated address might be about to expire.
  - The use of this address is discouraged but not forbidden.
    - It can continue to be used for old connection, but cannot be used for new connection.
- When a node is auto-configured, the following steps are performed:
  - A link-local address is first generated by using the link-local prefix FE80 & the interface identifier.
    - This address is a tentative address.
  - Then the node needs to join the all-nodes multicast group (FF02::1) & the solicited-node multicast group (FF02::1:xx) for the tentative address.
  - A NS message is sent out with the tentative address as the target address (for DAD).
    - The source IP address of this message is the all-zeros address (::).

Auto-configuration (4)

- The IP destination address is the solicited-node multicast address.
- In DAD, if another node on the link already uses this address, that node replies with a NA message and the auto-configuration mechanism stops.
  - In this case, manual configuration of the host is required.
- If there is no answer to the NS, it is safe to use the address.
- The address is assigned to the interface and the state of the address changes to "preferred".

❖ In order to determine which router & prefix are used, the host sends a RS message to the all-routers multicast group (FF02::2).

- All routers on the link reply with RA messages.
  - For each prefix in Router Advertisements, an IPv6 address is generated by combining the prefix with the interface identifier.

Auto-configuration (5)

- These addresses are added to the list of assigned addresses for the interface.

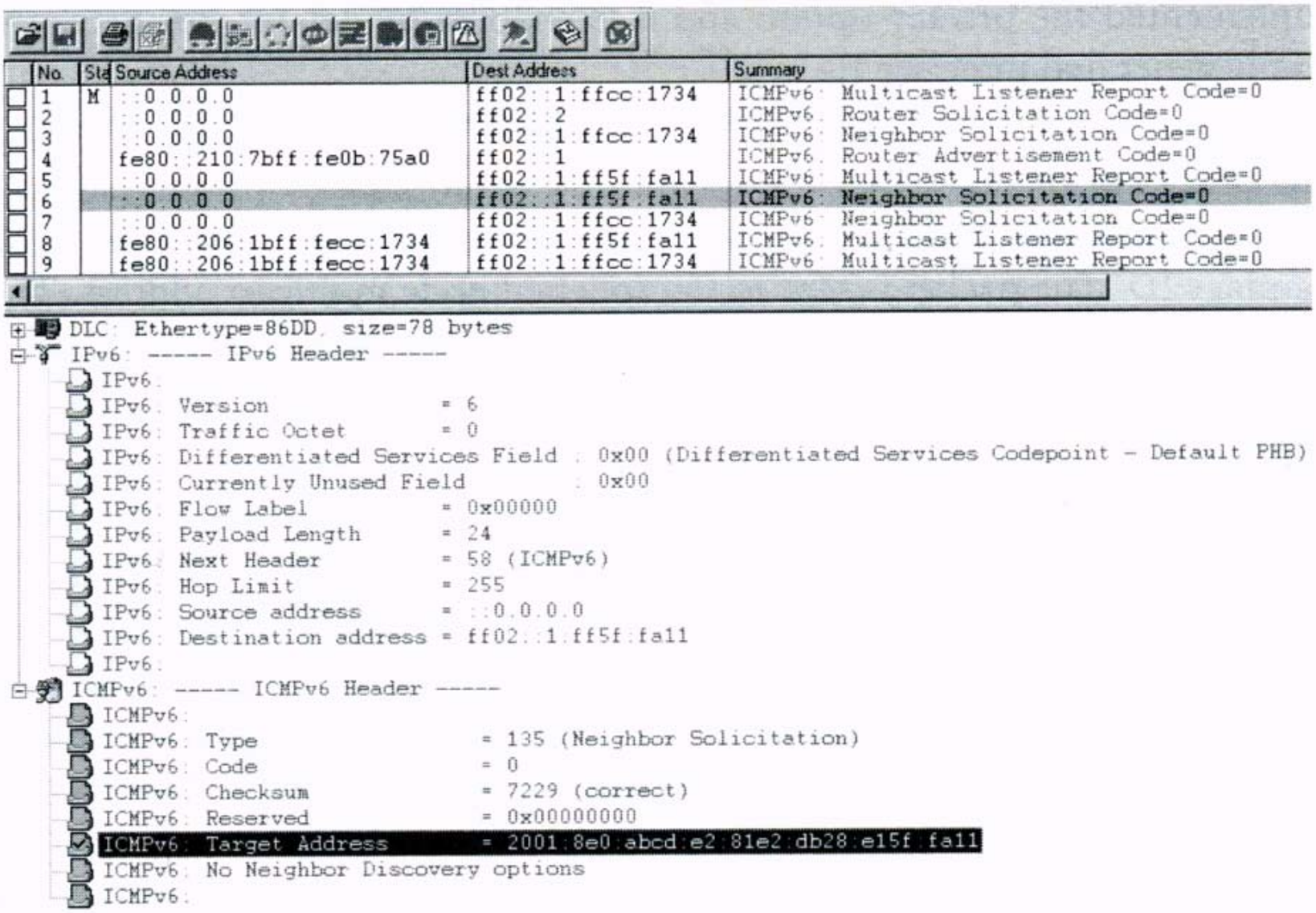
- All other addresses configured manually or through stateful configuration need to be verified individually (using DAD).

Auto-configuration (6)

- Auto-configuration in the trace file.
  - Router Solicitation (RS): it is for request prefix at present.

Auto-configuration (7)

- Neighbor solicitation (NS): it is for DAD at present.

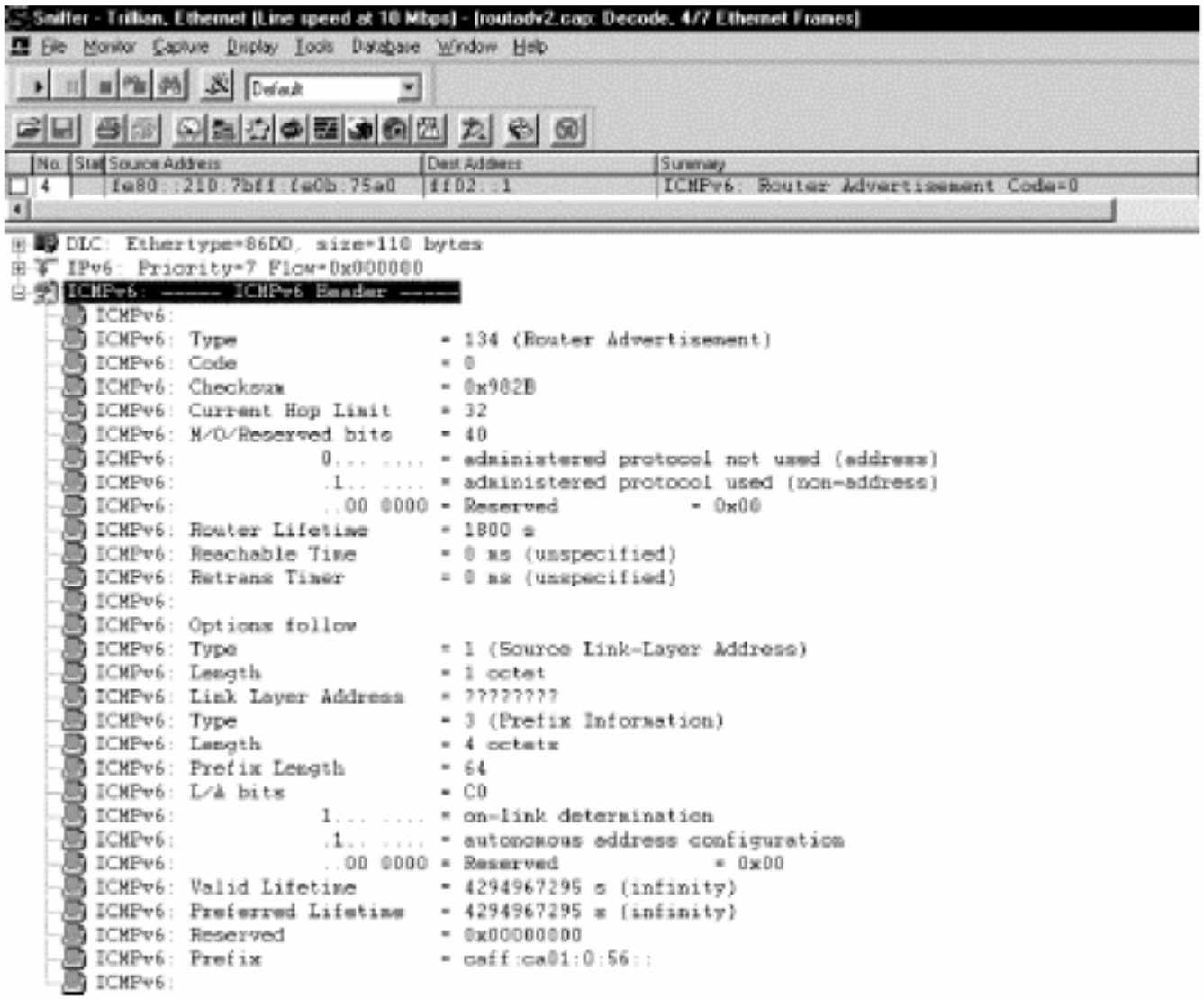


No.	Sta	Source Address	Dest Address	Summary
1	M	::0.0.0.0	ff02::1:ffcc:1734	ICMPv6: Multicast Listener Report Code=0
2		::0.0.0.0	ff02::2	ICMPv6: Router Solicitation Code=0
3		::0.0.0.0	ff02::1:ffcc:1734	ICMPv6: Neighbor Solicitation Code=0
4		fe80::210:7bff:fe0b:75a0	ff02::1	ICMPv6: Router Advertisement Code=0
5		::0.0.0.0	ff02::1:ff5f:fa11	ICMPv6: Multicast Listener Report Code=0
6		::0.0.0.0	ff02::1:ff5f:fa11	ICMPv6: Neighbor Solicitation Code=0
7		::0.0.0.0	ff02::1:ffcc:1734	ICMPv6: Neighbor Solicitation Code=0
8		fe80::206:1bff:fecc:1734	ff02::1:ff5f:fa11	ICMPv6: Multicast Listener Report Code=0
9		fe80::206:1bff:fecc:1734	ff02::1:ffcc:1734	ICMPv6: Multicast Listener Report Code=0

```
DLC: Ethertype=86DD, size=78 bytes
IPv6: ----- IPv6 Header -----
IPv6:
IPv6: Version                  = 6
IPv6: Traffic Octet            = 0
IPv6: Differentiated Services Field : 0x00 (Differentiated Services Codepoint - Default PHB)
IPv6: Currently Unused Field        : 0x00
IPv6: Flow Label               = 0x00000
IPv6: Payload Length           = 24
IPv6: Next Header              = 58 (ICMPv6)
IPv6: Hop Limit                = 255
IPv6: Source address           = ::0.0.0.0
IPv6: Destination address      = ff02::1:ff5f:fa11
IPv6:
ICMPv6: ----- ICMPv6 Header -----
ICMPv6:
ICMPv6: Type                   = 135 (Neighbor Solicitation)
ICMPv6: Code                   = 0
ICMPv6: Checksum               = 7229 (correct)
ICMPv6: Reserved               = 0x00000000
ICMPv6: Target Address         = 2001:8e0:abcd:e2:81e2:db28:e15f:fa11
ICMPv6: No Neighbor Discovery options
ICMPv6:
```

Auto-configuration (8)

- Router advertisement (RA): it is for MAC address & prefix now.



Path MTU Discovery (1)

- In IPv4, every router can fragment packets, if needed.
  - The fragmented packet is then reassembled at the final destination.
- In IPv6, routers do not fragment packets anymore, the sender processes it.
- Path MTU discovery (RFC 1981) tries to ensure that a sent packet uses the largest possible size on a path.
- The Path MTU adopts the smallest link MTU of all links (from source to destination).
- The discovery process works like this:
  - First, a host uses the Path MTU of the first hop link to send its packet.
  - If the packet is too big for the receiving router, the router discards the packet and sends back an ICMPv6 Packet Too Big message.
    - The router refers to the Path MTU of the next link.
  - This replied message includes the MTU size of the next hop link.
  - The host now uses this smaller MTU for sending further packets to the same destination.

Path MTU Discovery (2)

- The process of receiving a Packet Too Big message & reducing the size of the packets can happen more than once, before the packet reaches its destination.
- The PMTU discovery process ends when the packets arrive at the final destination.

- In IPv6, the minimum MTU size is 1280 bytes.
- An IPv6 host will try to increase the MTU size from time to time in order to be able to detect a larger Path MTU.
- Path MTU discovery also supports multicast destinations.
  - The packet size used by the sender is the smallest Path MTU of the whole set of destinations.

Multicast Group Management (1)

- Multicast group addresses are used as an identifier for a group of nodes.
  - They are identified by a high-order byte of FF.
- A protocol is required to manage the efficient routing of packets with multicast group addresses as a destination.
- In IPv4, multicast group management is done via IGMP (Internet Group Management Protocol).
  - IGMP ver2 is defined in RFC 2236.
- IPv6 uses ICMPv6 messages for the same functionality.
  - It was based on IGMPv2 specifications.
  - It is now called MLD (Multicast Listener Discovery) (RFC 2710).
- All MLD messages use link-local IPv6 addresses as source addresses.
  - Their hop limit are set to 1 to make sure they remain in the local network.
- MLD has 3 types of message:
  - Multicast Listener Query message (type: 130).

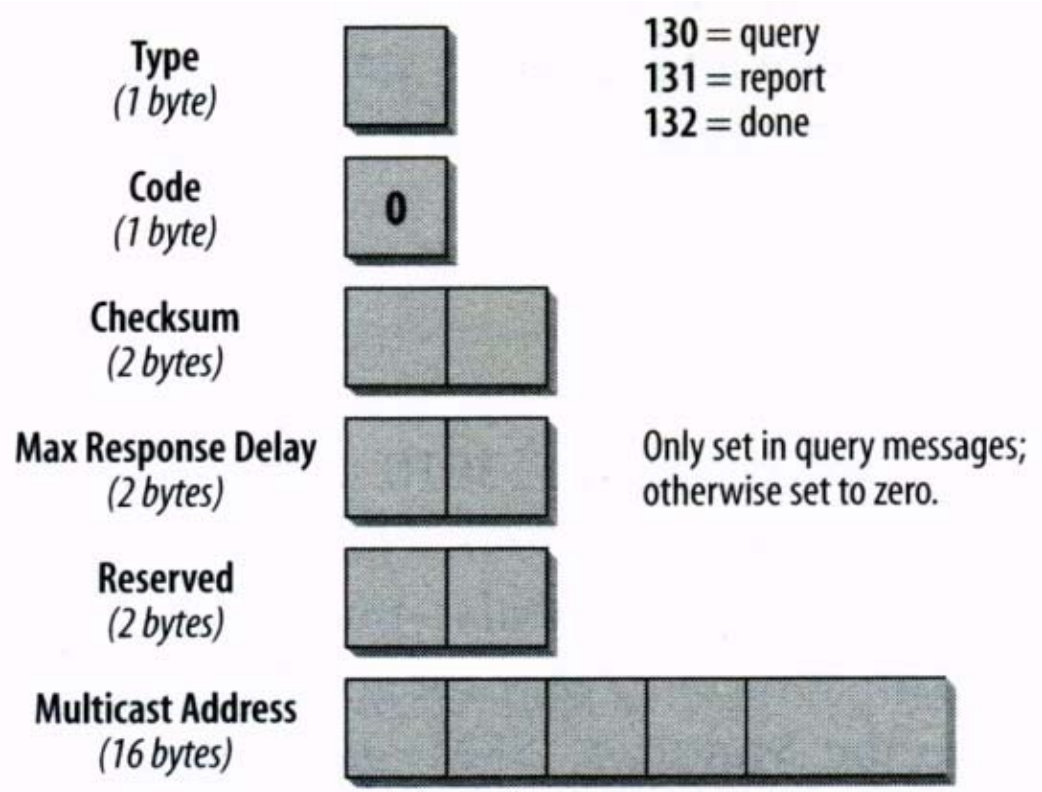
Multicast Group Management (2)

- There are 2 types of query messages:
  - One is a general query: it is used to determine which multicast group addresses have listeners on a link.
  - The other is an address-specific query: it is used to determine whether there are listeners for a specific address on a link.
- Multicast Listener Report message (type: 131).
- Multicast Listener Done message (type: 132).

- These 3 message types have the same format:
  - The Type field: 130, 131, & 132.
  - The Maximum Response Delay field: is used only in query messages.
    - It is the maximum allowed delay (in msec) for a node sends a report if it is a listener.
    - In all other messages (report & done), this field is set to 0.
  - The Multicast Address field:
    - In a general query, it is set to 0.

Multicast Group Management (3)

- MLD message format.



Multicast Group Management (4)

- In an address-specific query, it contains the queried multicast group address.
- In report & done messages, it contains the multicast group to which a member listens or the group it is leaving.

- Router uses MLD to discover which multicast addresses have listeners on each of its links.
  - For each attached link, the router keeps a list of listener addresses.
- General queries are sent to the link-local scope all-nodes with multicast address FF02::1.
- When a station (host) receives the query, it will wait some random delay & then send a report.
  - The maximum delay is specified in the Maximum Response Delay field in the query (like CSMA/CA).
- During the delay time, if the station (host) hears another station sending a report, it will stop the report process.
  - This can avoid bandwidth waste.

Multicast Group Management (5)

- MLD ver2 is already in the works as a draft & is based on IGMPv3.

- Following is the Message types of MLD and their destination addresses.

Message type	IPv6 Destination address
General Query	Link-local scope all-nodes (FF02::1)
Multicast Address–Specific Query	The multicast address being queried
Report	The multicast address being reported
Done	Link-local scope all-routers (FF02::2)